STRUCTURAL ELUCIDATION OF LIGNOCELLULOSE AND WETLAND SOIL BY SOLID-STATE NMR AND DYNAMIC NUCLEAR POLARIZATION

By

Wancheng Zhao

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry – Doctor of Philosophy

2023

ABSTRACT

Solid-state Nuclear Magnetic Resonance (ssNMR) is a powerful analytical technique that has emerged as a versatile tool for high-resolution detection and characterization of biosolids, such as carbohydrates, proteins, nucleic acids, and organic matter, in their native state. By harnessing the principles of NMR spectroscopy, ssNMR provides detailed insights into the atomic-level structure, dynamics, and interactions of these biomolecules, making it an invaluable resource for advancing our understanding of complex biological systems. This work is dedicated to addressing fundamental yet profoundly significant questions in the fields of sustainable bioenergy development and coastal wetland ecosystem preservation by ssNMR.

Firstly, we have unraveled the intricate packing of lignin and carbohydrates within plant cell walls, a critical inquiry that underpins our ability to engineer plants for improved biofuel and materials production. This research holds exceptional significance within the field of bio-renewable energy. Leveraging the power of ssNMR, we have achieved atomic-resolution characterization of the polymorphic structure, dynamical behavior, hydration profiles, and physical arrangement of lignin and polysaccharides. The findings uncover that the extent of glycan-aromatic association increases sequentially across grasses, hardwoods, and softwoods; and lignin principally packs with the xylan in a non-flat conformation via electrostatic interactions and partially binds the junction of flat-ribbon xylan and cellulose surface as a secondary site.

Another exciting development in our research is the integration of the sensitivity-enhancing Dynamic Nuclear Polarization (DNP) technique. This innovative approach holds immense promise for the investigation of natural abundance samples. We have successfully demonstrated the feasibility of combining traditional ssNMR with DNP methods to probe the structure and dynamics of carbohydrates in natural abundance rice stems. This streamlined approach enables the

efficient screening of a diverse array of biomass materials without the need for labeling.

Furthermore, the application of ssNMR and DNP has facilitated the rapid acquisition of 2D $^{13}C/^{1}H$-^{13}C correlation spectra to detail the molecular fingerprint and spatial organization of biopolymers in unlabeled wetland soil obtained from a brackish island situated along the Gulf of Mexico coastline. Surprisingly, the lignocellulosic core identified in the plants grown on this island was found to be preserved in the surface layer of soil formed in the brackish marshes, but with tighter physical packing between molecules. Extending the depth from the surface to 2 m covers a geological timeline of eleven centuries, where we found sophisticated changes in the molecular structure and composition of organic matter as well as the bulk properties of the wetland soil. These findings offer a promising avenue toward addressing pressing environmental challenges, including coastal wetland loss and the impact of sea-level rise.

Additionally, statistical approaches for the analysis of carbohydrate composition and structure using high-resolution ssNMR data have been introduced. We generate simulated spectra (density maps) using data deposited in our developed and maintained Complex Carbohydrates Magnetic Resonance Database (CCMRD, www.ccmrd.org), to demonstrate the chemical shift dispersion and to aid in the fast identification of important fungal and plant polysaccharides in cell walls. By comparing the projected spectra with experimental data, we highlight the challenges of assigning resonances because distinct carbohydrates from different organisms can produce nearly identical signals.

To my beloved wife,
Thanks for your love, encouragement, and unwavering support.

To my dear parents,
Thanks for your values, sacrifices, and unshakable faith.

ACKNOWLEDGEMENTS

In the tapestry of my academic journey, I must first express my profound gratitude to my advisor, Prof. Tuo Wang, who has not only served as my guiding light but also as the very embodiment of the scholar I aspire to become. If asked to name the single individual I held in the highest regard throughout my entire existence, your name would resound without hesitation. Your vast knowledge has been a wellspring of inspiration, continuously guiding us to the forefront of research. Your unwavering patience and serene encouragement have provided solace during our missteps and diversions from our goals. Your generosity in sharing wisdom, resources, and valuable materials has seamlessly integrated us into the academic community. Your passionate support for our participation in conferences has expanded our intellectual horizons, while your guidance transcends research mechanics, instilling in us the values of a proficient researcher and exemplary research habits. Your exceptional support for our scholarships and heartfelt recommendation letters for our careers have been truly extraordinary.

Words alone are inadequate to capture the depth of my gratitude for your contributions. Indeed, it is a rare privilege and a lifetime's honor to have been counted among your students. Only now do I fully appreciate the depth of the saying, “If I have seen a little further, it is by standing on the shoulders of Giants.” To me, you are the very giant whose shoulders I stand upon, affording me a profound view of an expansive world. And it is my fervent hope to be the pillar on which the next generation of scientists can rely, as you have been for me, nurturing their boundless potential.

I would like to express my sincere appreciation to the esteemed members of my committees: Prof. Gary Blanchard, Prof. Charles Hoogstraten, Prof. Josh Vermaas, and Prof. David Weliky. Your invaluable contributions, including critical insights, thoughtful feedback, and constructive suggestions, have been instrumental in enhancing the quality of this work. I am

particularly thankful to Prof. Blanchard for his significant assistance during my transfer process and unwavering support for my Dissertation Completion Fellowship applications. Prof. Hoogstraten, I am grateful for the enlightening discussions we've had at conferences, which have provided me with valuable insights and enriched my academic journey. Prof. Vermaas, your guidance and advice on my research path and future career planning have been invaluable, and I truly appreciate the positive impact you've had on my academic and professional development. Prof. Weliky, your dedication to academic excellence, along with your valuable advice and consistent guidance, has been a cornerstone of my academic journey.

I also want to extend my gratitude to my committee members during my time at LSU: Prof. Bin Chen, Prof. Robert Cook, and Prof. Paul South. Your collective wisdom and guidance have broadened my perspective and significantly contributed to the depth and rigor of this research. Prof. Cook, I am especially thankful for your unwavering support in providing recommendation letters, which have played a pivotal role in shaping my research trajectory. Your mentorship and collaborative efforts in the academic realm have greatly enriched my overall experience.

My heartfelt thanks go out to my lab colleagues, Dr. Qinghui Cheng, Dr. Jayasubba Yarava, Dr. Krishna Bankala, Isha Gautam, Debkumar Debnath, Anand Jacob, Ankur Ankur, Kalpana Singh, and Priya Sahu. I want to extend my heartfelt appreciation to the former members of the Wang group Dr. Xue Kang, Dr. Alex Kirui, Dr. Malitha Chathuranga, Dr. Liyanage Fernando, Dr. Fabien Deligey, Dr. Nader Ghassemi, Dr. Chandra Shekar, Dr. Alexandre Poulhazan, Dr. Liliana Martínez, Arnab Chakraborty, Irina Gomez, Paninga Muiliya, and Jianyu Guan. Your camaraderie, stimulating discussions, and shared experiences have not only created an intellectually stimulating environment but have also provided a strong support system. Special thanks to Fabien, for your guidance. Most of my NMR knowledge has been acquired from you, especially during the initial

stages after the pandemic. I've gained valuable insights from you, and my English proficiency has also significantly improved with your help. Special thanks also to Dev and Isha. We came from LSU to MSU together and went through some of the toughest times as a team. Your support during those moments is something I will cherish throughout my life.

I extend my gratitude to my scientific collaborators who have been essential to our projects and have engaged in invaluable support and discussions. I deeply appreciate the collaboration with Prof. Daniel Cosgrove from Penn State. Thanks for your wholehearted support during my job search, your patient guidance, and the privilege of learning and exchanging ideas within the CLSF center under your leadership. I am profoundly grateful for the collaboration with Dr. Frederic Mentink-Vigier from the National High Magnetic Field Laboratory (Maglab). Our in-depth discussions, your wholehearted support, and your steadfast guidance in the realm of DNP are sincerely appreciated. Additionally, I would like to recognize Dr. Faith Scott, Dr. Zhenhong Gan, Dr. Ivan Hung, and Dr. Riqiang Fu, all from Maglab, for their valuable guidance and assistance in experimental setup. I would like to express gratitude to Prof. John R. White from LSU for our fruitful collaboration and your valuable insights throughout. I would also like to extend my gratitude to Maglab and the Pacific Northwest National Laboratory for their indispensable instrumental support.

I would also like to express my sincere gratitude to the dedicated NMR managers, Dr. Daniel Holmes (MSU), Dr. Li Xie (MSU), Dr. Thomas Weldeghiorghis (LSU), and Dr. Fengli Zhang (LSU). Your unwavering commitment and expertise have been invaluable in ensuring the smooth operation of our NMR facilities. Special thanks to Dr. Holmes and Dr. Xie, your tireless efforts and extensive knowledge in maintaining and optimizing the NMR instruments have greatly contributed to the success of our research endeavors.

I would like to extend my heartfelt appreciation to my friends, who have been an essential source of support and encouragement during my academic pursuit. To Dr. Xiaobing Chen, as my “buddy” at LSU, your unwavering support facilitated my initial adjustment to the new foreign environment. To Dr. Weiwei Xie, your support and encouragement during my toughest moments were profoundly significant, helping me find strength and hope.

I would like to express my deepest appreciation to my family for your endless love, tolerance, and support, which have been instrumental in shaping the person I have become today.

To my wife, Tongjia Fu, your understanding and support have been the bedrock of my daily motivation over the past five years of my academic journey in a foreign place. In moments of uncertainty, you've consistently stood by my side, offering me the strength to persevere. Reflecting on our 18 years together, I can't help but feel incredibly fortunate; having you in my life has been an invaluable blessing. To my parents, Xiufen Lang and Xiliang Zhao, I am profoundly thankful for your care, guidance, understanding, and patience. You've consistently been the steadfast foundation of my academic journey. I am always certain that no matter what, your unwavering support and understanding are there for me. To my parents-in-law, Hongmin Fu and Nairu Fu, your consistent encouragement, trust, and belief in me sustained my motivation during this challenging journey. I'm also grateful that you have entrusted me with such an exceptional daughter.

Heartfelt appreciation to each and every one of you!

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1
REFERENCES 36

CHAPTER 2: SOLID-STATE NMR OF PLANT AND FUNGAL CELL WALLS: A CRITICAL REVIEW 42
REFERENCES 61

CHAPTER 3: CURRENT LIMITATIONS OF SOLID-STATE NMR IN CARBOHYDRATE AND CELL WALL RESEARCH 68
REFERENCES 88

CHAPTER 4: POLYSACCHARIDE ASSEMBLIES IN FUNGAL AND PLANT CELL WALLS EXPLORED BY SOLID-STATE NMR 94
REFERENCES 115

CHAPTER 5: A THREE-DIMENSIONAL 13C-13C-13C CORRELATION EXPERIMENT FOR HIGH-RESOLUTION ANALYSIS OF COMPLEX CARBOHYDRATES USING SOLID-STATE NMR 125
REFERENCES 141
APPENDIX 146

CHAPTER 6: EFFECT OF CROSS POLARIZATION RADIOFREQUENCY PHASES ON SIGNAL PHASE 153
REFERENCES 168

CHAPTER 7: CARBOHYDRATE-AROMATIC INTERFACE AND MOLECULAR ARCHITECTURE OF LIGNOCELLULOSE 173
REFERENCES 200
APPENDIX 210

CHAPTER 8: SOLID-STATE NMR OF UNLABELED PLANT CELL WALLS: HIGH-RESOLUTION STRUCTURAL ANALYSIS WITHOUT ISOTOPIC ENRICHMENT 275
REFERENCES 296
APPENDIX 302

CHAPTER 9: CHRONOLOGICAL MOLECULAR FINGERPRINT OF WETLAND SOIL BY SENSITIVITY-ENHANCED SOLID-STATE NMR 316
REFERENCES 340
APPENDIX 347

CHAPTER 10: CCMRD: A SOLID-STATE NMR DATABASE FOR COMPLEX CARBOHYDRATES 371
REFERENCES 382
APPENDIX 387

CHAPTER 11: CHARTING THE SOLID-STATE NMR SIGNALS OF POLYSACCHARIDES: A DATABASE-DRIVEN ROADMAP 390
REFERENCES 411
APPENDIX 418

CHAPTER 1: INTRODUCTION

1.1 Significance of the Research Endeavor

Lignocellulosic biomass, derived from sources like forest waste, agricultural crop residues, and algae, represents a vast and sustainable resource with the potential to revolutionize the production of biofuels, biomaterials, and high-value bioproducts[1, 2]. The structural foundation of this resource lies in plant cell walls, which are essential for maintaining the mechanical strength and flexibility necessary to preserve cellular integrity and form[3]. The primary cell walls[4, 5] of the growing plant consist of three main classes of carbohydrate polymers: cellulose, hemicellulose, and pectin. Cellulose microfibrils, with diameters ranging from 3 to 5 nanometers, are composed of rigid and semi-crystalline β-(1-4)-glucan chains. Hemicellulose, on the other hand, contains shorter and often branched chains. For example, xyloglucan (XyG), a major hemicellulose in the primary walls of dicots like *Arabidopsis thaliana*, features a β-1,4-glucan backbone adorned with α-xylose (Xyl) sidechains. Pectin, which typically includes homogalacturonan (HG), rhamnogalacturonan-I (RG-I), and RG-II, plays a crucial role in embedding itself within the cellulose-hemicellulose network and performing various functions such as regulating wall porosity, modulating cell expansion, and controlling pH and ionic balance. The secondary cell wall[6, 7], formed inside the primary cell wall when the plant ends growing, comprises cellulose, hemicellulose (specifically xylan and glucomannan), and lignin. In the secondary wall, larger cellulose bundles (ranging from 10 to 20 nanometers) are encased by a xylan-lignin crosslink, with lignin being a diverse and hydrophobic biopolymer rich in aromatic subunits.

Polysaccharides and lignin, as the fundamental biopolymers in plant cell walls, are held together by numerous covalent and physical interactions[8]. Nonetheless, these biopolymers are typically insoluble in water, partially indigestible, and often non-crystalline. This presents a

significant challenge when attempting to achieve high-resolution structural characterization of intact cell walls. Traditional methods primarily focus on the structural analysis of specific components of the cell wall, involving the hydrolysis of biomolecules into small soluble compounds using organic solvents or ionic liquids. The hydrolysis treatments can significantly disrupt and even modify the biopolymers, causing conventional strategies to fall short in their ability to elucidate the packing and motion within native cell walls. Hence, there is a pressing need for alternative methods that allow for undisturbed detection and analysis of plant cell walls, to benefit more sustainable and efficient biomass utilization.

In addition to the more efficient utilization of renewable energy, another critical concern related to global climate issues revolves around mitigating coastal land erosion[9-11]. The escalating loss of both wetland and coastal land, compounded by the growing threat of rising sea levels, represents an urgent and globally significant environmental crisis. Louisiana holds 40% of the coastal and estuarine wetlands in the contiguous United States but has suffered the highest rate of wetland loss, with around 4,900 square kilometers of coastal land vanishing since 1930[12]. Wetlands, encompassing diverse ecosystems such as salt marshes, mangroves, and seagrass beds, play an indispensable role as natural barriers along coastlines. They provide vital services, including shielding against coastal erosion, flood mitigation, and fostering biodiversity. Remarkably, coastal wetlands, known as blue carbon ecosystems, occupy a mere 0.07 to 0.22% of the Earth's surface, yet they sequester more than 50% of the carbon annually buried in the ocean[9, 13]. These wetlands serve as major reservoirs of soil organic matter (SOM) and its associated soil organic carbon (SOC), critical in regulating Earth's climate, water systems, and biodiversity. Unfortunately, subsidence and rising sea levels induce the collapse of wetland soils into adjacent aerobic waters, triggering chemical decomposition of SOM components and the release of stored

carbon as carbon dioxide and methane into the atmosphere. Beyond their role as carbon sinks, the wetlands of Louisiana hold particular significance, as they offer insights into erosion patterns worldwide, driven by accelerated sea-level rise and subsidence. To protect these invaluable ecosystems and mitigate greenhouse gas emissions, it is imperative to comprehensively grasp, model, and predict SOM carbon dynamics at scale. Detailed molecular insights, including functionality, isomerism, and conformations of this carbon pool, are fundamental prerequisites for this understanding. Presently, there is limited research on this advanced level of carbon sequestration within coastal wetlands due to the complexities and insolubility of soil.

Multidimensional ^{13}C–^{13}C/^{1}H correlation ssNMR spectroscopy, with uniformly isotope-labeled (^{13}C) samples, is an emerging technique that unravels the mysteries of solid materials[14, 15]. It breaks free from the constraints of crystallinity, offering high-resolution structural insights into disordered and insoluble samples, capturing dynamic processes, and preserving sample integrity. Its versatility, ability to decipher a multitude of nuclei, and the capability to conduct in situ studies make it invaluable in various fields, from materials science and chemistry to pharmaceuticals and structural biology. In addition, the introduction of cutting-edge Magic-Angle Spinning Dynamic Nuclear Polarization (MAS-DNP) enables the transfer of electron polarization to NMR-active nuclei, leading to one to two orders of NMR sensitivity enhancements[16, 17]. This breakthrough allows for the acquisition of high-resolution 2D spectra even from unlabeled samples containing low-abundance isotopes such as carbon-13 (1.1%) and nitrogen-15 (0.4%).

My thesis research is dedicated to the integration of cutting-edge multidimensional high-resolution ssNMR techniques, complemented by sensitivity-enhanced DNP, to delve deeply into the intricate structural and dynamic attributes of lignocellulosic biomass and Louisiana wetland samples. The outcomes of this research will have a far-reaching impact, fostering the evolution of

next-generation pretreatment techniques and advancing lignin bioengineering for a more sustainable and efficient utilization of biomass. Meanwhile, these advancements will play a crucial role in safeguarding Louisiana's distinctive and fragile coastline, making a substantial and positive contribution towards its preservation.

1.2 SsNMR Methodology

1.2.1 The Basis of NMR Spectroscopy

NMR spectroscopy is a sophisticated analytical technique that has had a profound impact on scientific research across a multitude of disciplines[18-22]. Its origins in the study of nuclear magnetic properties have given rise to its pivotal role as a unifying tool that bridges the gaps between chemistry, physics, biology, and materials science. NMR serves as an indispensable key for unlocking the secrets of the atomic and molecular world, providing profound insights into the structures, dynamics, and interactions at the most fundamental level.

This method capitalizes on the intriguing responses of atomic nuclei possessing. When non-zero spins (I≠0, **Table 1.1**) are subjected to an external magnetic field, they align themselves with this field, much like a compass needle aligns with Earth's magnetic field. The application of a precisely tuned radiofrequency (RF) pulse disrupts this alignment, causing the nuclei to deviate from their equilibrium positions and initiate precession around the magnetic field's direction. This results in an energy difference, denoted as ΔE and defined by the equation:

$$\Delta E = -\hbar\gamma B_0 \qquad \textbf{Eq 1.1}$$

In this equation, γ represents the gyromagnetic ratio, a fundamental property unique to each isotope. B_0 signifies the strength of the static magnetic field, while $\hbar$ (h-bar) represents the reduced Planck constant. In modern NMR spectroscopy, powerful magnetic fields are employed (5 - 28 T), corresponding to the well-known ^{1}H Larmor frequencies (ω_0, **Figure 1.1**)[23].

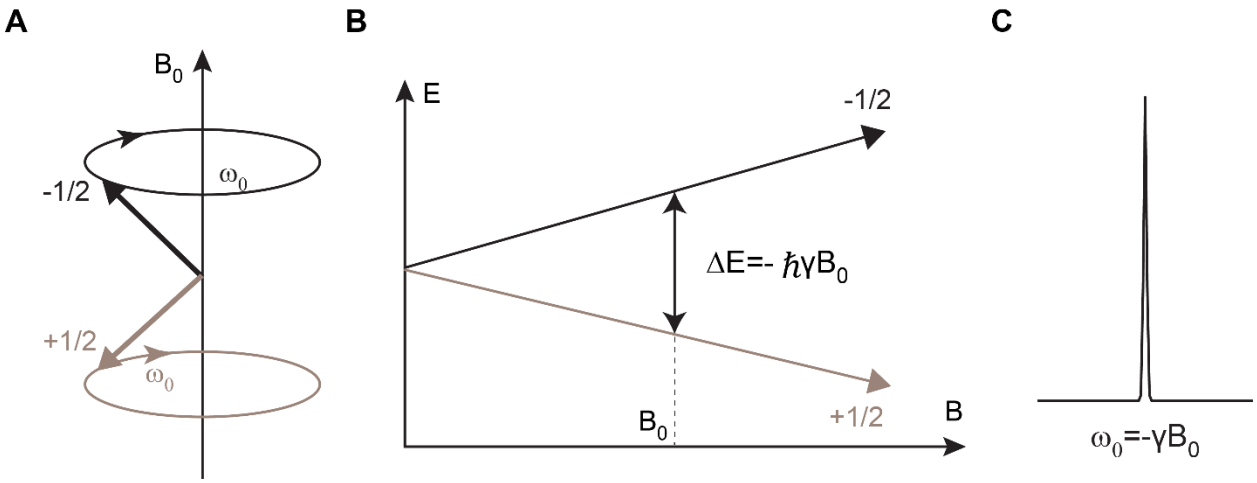

Figure 1.1 Basis of NMR Spectroscopy. **a,** Nuclear spin undergoes precession around a static magnetic field (B_0) at a frequency (ω_0). **b,** The transition frequency between the energy levels of the spins. **c,** The resulted single peak spectrum at the Larmor frequency of the spin.

Table1.1 Frequently examined atomic nuclei in ssNMR.

Nuclei	Spin	γ (10^7·rad/T·s)	ω_0 at 18.8 T	Abundance
^{1}H	1/2	26.75	800 MHz	99.98 %
^{2}H	1	4.11	123 MHz	0.02 %
^{13}C	1/2	6.73	200 MHz	1.11 %
^{14}N	1	1.93	58 MHz	99.63 %
^{15}N	1/2	-2.71	80 MHz	0.37 %
^{17}O	5/2	-3.63	108 MHz	0.04 %
^{19}F	1/2	25.18	753 MHz	100 %
^{31}P	1/2	10.84	324 MHz	100 %
[a] e^-	1/2	-17600	-526 GHz	-

[a] The electron constant is provided for the understanding of DNP mechanism.

In NMR spectroscopy, several fundamental principles and concepts converge, enabling the profound insights it offers into the world of molecules and materials. One of the key elements is the exquisite sensitivity of NMR frequencies to the electron distribution surrounding the nucleus. This electron shielding affects the nucleus's magnetic environment, resulting in slightly different frequencies due to variations in the shielding constant (σ) among nuclei of the same isotope within a molecule. The frequencies, typically reported as chemical shifts (δ) with a unit part per million (ppm), relative to well-established reference compounds, such as tetramethylsilane (TMS) that is frequently utilized for organic molecules, and sodium trimethylsilyl propane sulfonate (DSS), which is a common reference in protein NMR. The variation in chemical shift, for a specific isotope[23, 24], can range from 0 to 10 ppm for ^{1}H, from 0 to 200 ppm for ^{13}C, and up to 900 ppm for ^{15}N. Furthermore, NMR frequencies are subject to the influence of a range of couplings, including spin-spin scalar (J-coupling), spin-spin dipolar, and quadrupolar couplings, which are contingent on factors such as covalent bonding, internuclear distances, and nuclear spin values. These anisotropic interactions[22] (**Figure 1.2**), dependent on the sample's orientation relative to the magnetic field, allow NMR spectra to encode rich three-dimensional structural information, offering valuable insights into the spatial arrangement of atoms within a molecule.

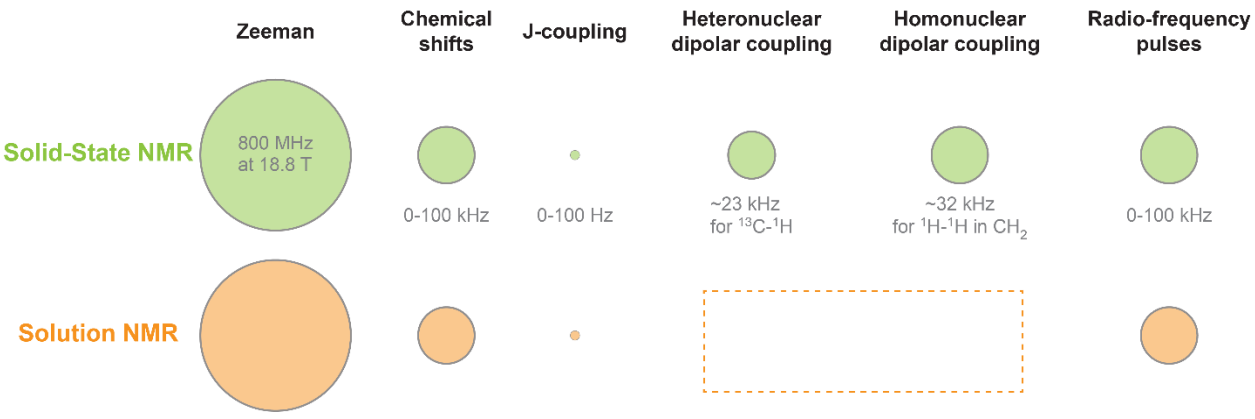

Figure 1.2 NMR spin interactions. Dipolar coupling is absent in solution NMR.

1.2.2 Solid-State NMR Spectroscopy

The stark contrast between ssNMR and solution NMR stems primarily from the absence of rapid molecular tumbling in the former. In solution NMR, the swift tumbling of molecules ensures that the orientation of the shielding/dipolar tensor with respect to the magnetic field (B_0) is rapidly averaged over all conceivable values. Conversely, ssNMR, in a static state, is susceptible to a wide array of orientation-dependent interactions, including chemical shift anisotropy (CSA) and heteronuclear dipolar coupling.

To address this challenge, the adoption of spinning the sample at the magic angle, set at 54.74° relative to the external magnetic field, is a fundamental strategy. This mechanically induced uniaxial rotation, specifically the magic-angle spinning (MAS)[25, 26], is an essential technique extensively utilized in the vast majority of ssNMR experiments (**Figure 1.3**). Its primary function is to nullify the influence of CSA and assist in mitigating heteronuclear dipolar-coupling effects. Furthermore, it serves to narrow spectral lines originating from quadrupolar nuclei and is increasingly preferred for counteracting homonuclear dipolar coupling in NMR spectra.

The chemical shift Hamiltonian can be represented as:

$$\widehat{H}_{cs} = -\sigma_{zz}.\gamma\hbar B_o.\hat{I}_z \qquad \textbf{Eq 1.2}$$

σ_{zz}, along with other two chemical shielding tensors σ_{xx} and σ_{yy}, commonly employed to describe the chemical shift interaction within the principal axis system, which is expressed as:

$$\sigma_{zz}=\sigma_{iso} + \frac{1}{3}\sum_{j=x,y,z}(3cos^2\theta - 1)\,\sigma_{jj} \qquad \textbf{Eq 1.3}$$

Then, to show both isotropic and anisotropic parts, the Hamiltonian can be rephrased as:

$$\widehat{H}_{cs} = \left\{\sigma_{iso}.\gamma\hbar B_0 + \frac{1}{2}\delta[3cos^2\theta - 1 - \eta sin\theta^2\cos(2\phi)]\right\}\hat{I}_z \qquad \textbf{Eq 1.4}$$

When a high-speed uniaxial rotation (η = 0) is employed, it introduces an extra averaging factor, leading to a modification in the frequency of CSA:

$$\omega_{aniso} = \frac{1}{2}\delta[3cos^2\beta - 1 - \eta sin\beta^2 \cos(2\alpha)]\frac{1}{2}(3cos^2\theta_r - 1) \quad \textbf{Eq 1.5}$$

Within the rotor frame (**Figure 1.3**), β characterizes the angle between the principal z-axis of the interaction and the rotor's rotation axis, while θ denotes the angle between B_0 (the external magnetic field) and the primary z-axis of the interaction. When the rotor rotates around an axis tilted at an angle $\theta_r = 54.74°$, then $3cos^2\theta_r - 1 = 0$, resulting in zero CSA. Therefore, the detected frequency matches the isotropic chemical shift in the case of rapid MAS, resulting in well-defined spectral lines instead of the broad powder patterns. In situations where the rotor speed is slow, the chemical shift anisotropy can only be effectively averaged to zero after each rotor period, resulting in spinning sidebands at $\omega = \omega_{iso} \pm n\omega_r$ frequencies.

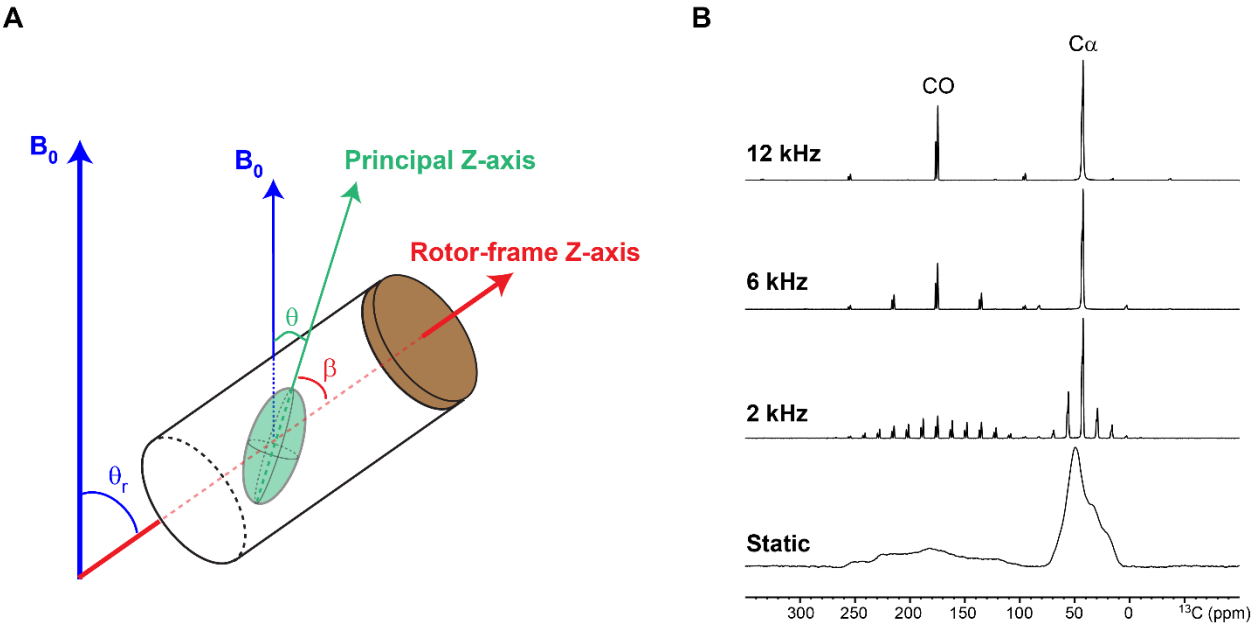

Figure 1.3 Magic-angle spinning. **a,** Illustrated cartoon presenting relationships among magic angle, lab frame, rotor frame, and principal axis system. **b,** SsNMR spectra (14.4 T) of ^{13}C-glycine at various spinning speeds.

The Cross Polarization (CP) technique[26] is also a key method used in ssNMR spectroscopy, which serves as both a standalone experiment and a fundamental component to enhance the sensitivity in numerous ssNMR experiments (**Figure 1.4a**). CP plays a pivotal role in transferring

polarization from NMR-sensitive protons to less sensitive X-nuclei like ^{13}C and ^{15}N (**Table 1.1**), often as the initial step and a common element in an array of 1D, 2D, and 3D ssNMR pulse sequences. CP finds application in both static and MAS conditions and serves diverse purposes. These applications range from reducing experimental time due to the shorter ^{1}H T_1 relaxation, to selectively detecting rigid components within the sample. CP is also employed to extract distance information by estimating dipolar interactions based on the build-up of CP intensities. Remarkably, CPMAS itself becomes part of the strategy to transfer polarization from an even richer source (electrons) in indirect DNP experiments.

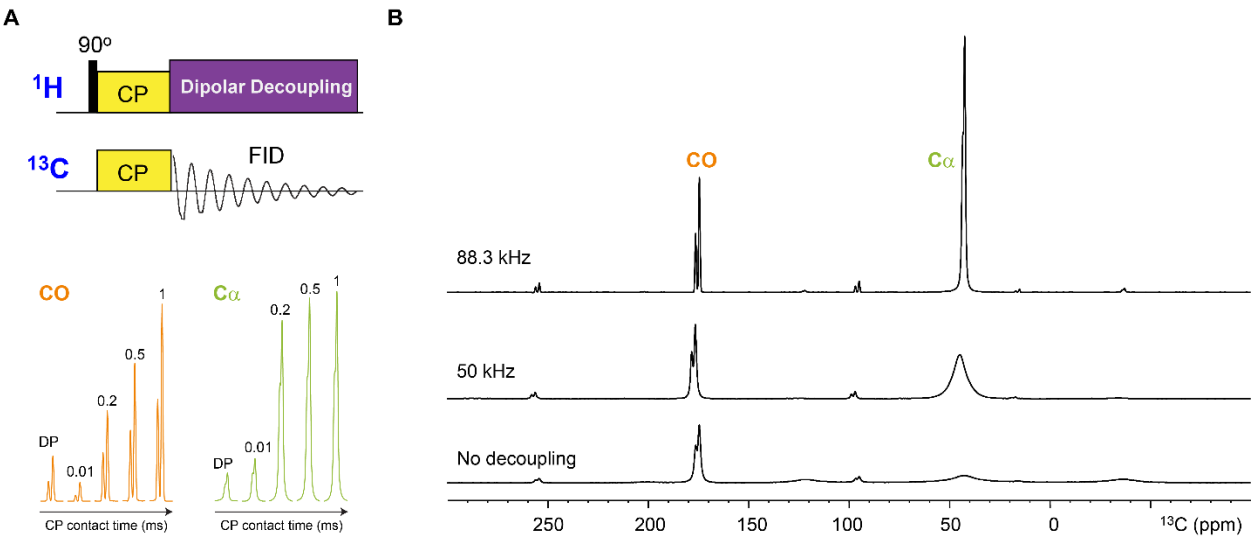

Figure 1.4 Cross polarization. **a,** SsNMR spectra (14.4 T, 12 kHz) of ^{13}C-glycine with various CP contact time, as well as CP pulse sequence. **b,** SsNMR spectra (14.4 T, 12 kHz) of ^{13}C-glycine with various heteronuclear dipolar decoupling.

Dipolar coupling, also known as dipolar interaction or dipole-dipole coupling, plays a significant role in understanding the structural and dynamic properties of molecules, particularly in ssNMR experiments. In ssNMR, the dipolar coupling can lead to homonuclear (between nuclei of the same type) and heteronuclear (between nuclei of different types) dipolar recoupling,

resulting in dipolar-based experiments such as PDSD/DARR, CORD, RFDR, HETCOR, and DIPSHIFT, etc. The dipolar coupling[20, 23] can provide crucial information about the spatial arrangement of atoms in a solid-state sample, as it encodes information about the internuclear distances and orientations. Based on **Eqs. 1.6 and 1.7**, the strength of dipolar coupling is inversely proportional to the cube of the distance (r) between nuclei.

$$\hat{H}_{dd}^{hetero} = -\left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_I\gamma_S\hbar}{r^3}(3cos^2\theta - 1)\hat{I}_z\hat{S}_z \qquad \textbf{Eq 1.6}$$

$$\hat{H}_{dd}^{homo} = -\left(\frac{\mu_0}{4\pi}\right)\frac{\gamma^2\hbar}{r^3}.\frac{1}{2}(3cos^2\theta - 1)(2\hat{I}_{1z}\hat{I}_{2z} - \hat{I}_{1x}\hat{I}_{2x} - \hat{I}_{1y}\hat{I}_{2y}) \qquad \textbf{Eq 1.7}$$

Meanwhile, it comes with challenges related to line broadening, orientation dependence, and data interpretation, which may require specialized techniques and expertise to overcome, such as MAS and decoupling. Dipolar interactions can be effectively nullified through MAS when the spinning frequency greatly exceeds the coupling strength. Nevertheless, heteronuclear dipolar coupling remains significant at moderate MAS frequencies, necessitating the use of ^{1}H dipolar decoupling sequences (continuous irradiation on ^{1}H channel) for achieving sharp lines (**Figure 1.4b**). For ^{1}H-^{1}H homonuclear coupling, besides the application of decoupling techniques, ultra-fast MAS (60-100 kHz) is imperative at present.

1.2.3 Pulse Sequences and Density Operator

A pulse sequence is a structured set of RF pulses, time delays, and gradient pulses used in NMR experiments. It serves as a precise set of instructions for manipulating nuclear spins in a sample and acquiring relevant data. In NMR, pulse sequences are tailored to achieve specific research goals, providing insights into relaxation times, spectral information, and molecular interactions. **Figure 1.5a** shows the pulse sequence of ^{13}C Direct Polarization (DP) experiment with a spin echo (Hahn echo)[27, 28]. As one of the most renowned pulsed NMR experiments and the most fundamental building blocks, the spin echo plays a crucial role by restoring the coherence of

dephased nuclear spins. It can refocus offsets (Ω) and/or chemical shifts and the phase angle by a τ- π- τ sequence, regardless of the initial conditions. As shown in the vector representation of **Figure 1.5b**, at the end, the magnetization of two spins aligns along the same axis, effectively eliminating the impact of offset. Finally, the time-domain Free Induction Decay (FID) signal collected in the acquisition stage is converted into a frequency-domain spectrum via Fourier Transform (FT).

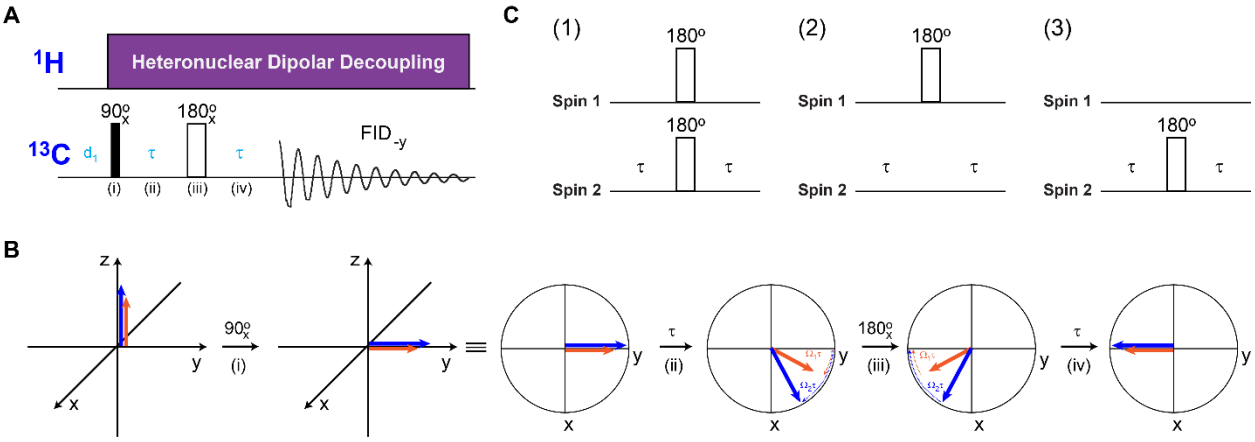

Figure 1.5 Spin echo experiment. **a,** The pulse sequence of ^{13}C DP with echo and decoupling. **b,** The vector representation. **c,** Three distinct spin-echo sequences in heteronuclear spin systems.

The utilization of the density operator (product operator) formalism in dissecting NMR experiments signifies a quantum-mechanical strategy, setting it apart from the semiclassical vector model. By harnessing the concept of density operators, we gain the ability to aptly illustrate significant phenomena, such as spin echo. For example, in **Figure 1.5b**, the original density operator for spin 1 is $\hat{I}_z$. The 90° pulse at (i) make $\hat{I}_z$ rotate along x-axis, giving:

$$\hat{I}_z \xrightarrow{\frac{\pi}{2}\hat{I}_x} cos\frac{\pi}{2}\hat{I}_z + sin\frac{\pi}{2}\hat{I}_y \equiv \hat{I}_y \qquad \textbf{Eq 1.8}$$

Note, in this thesis, we follow the left-hand rotation standard. At (ii), a free precession caused by offset result in a rotation about z-axis:

$$\hat{I}_y \xrightarrow{\Omega_1 \tau \hat{I}_z} cos\Omega_1 \tau \hat{I}_y + sin\Omega_1 \tau \hat{I}_x \quad \textbf{Eq 1.9}$$

Then, at (iii), a 180° pulse is applied about x-axis:

$$cos\Omega_1 \tau \hat{I}_y + sin\Omega_1 \tau \hat{I}_x \xrightarrow{\pi \hat{I}_x} - cos\Omega_1 \tau \hat{I}_y + sin\Omega_1 \tau \hat{I}_x \quad \textbf{Eq 1.10}$$

Only the y-component changes sign. Finally, at (iv), second free precession is along z-axis:

$$-cos\Omega_1 \tau \hat{I}_y + sin\Omega_1 \tau \hat{I}_x \xrightarrow{\Omega_1 \tau \hat{I}_z}$$

$$-cos^2\Omega_1 \tau \hat{I}_y - sin\Omega_1 \tau cos\Omega_1 \tau \hat{I}_x + cos\Omega_1 \tau sin\Omega_1 \tau \hat{I}_x - sin^2\Omega_1 \tau \hat{I}_y$$

$$\equiv -(cos^2\Omega_1 \tau + sin^2\Omega_1 \tau)\hat{I}_y \equiv -\hat{I}_y \quad \textbf{Eq 1.11}$$

Hence, we can concisely summarize the overall result of the spin echo sequence as:

$$\hat{I}_z \xrightarrow{90^\circ_x - \tau - 180^\circ_x - \tau} -\hat{I}_y \quad \textbf{Eq 1.12}$$

The final outcome remains constant regardless of the offset (Ω) or the delay (τ), making the spin echo serves as a mechanism that effectively refocusing chemical shifts.

Nonetheless, it's crucial to emphasize that the evolution of the scalar coupling may not always experience refocusing. **Figure 1.5c** illustrates three distinct scenarios. In the first situation, marked as (1), both chemical shifts of spins are refocused, but the J-coupling is not. In situations (2) and (3), the coupling's impact is refocused, while the chemical shifts of spins without the applied 180° pulse remain unrefocused. These assertions find support in the following density operator representations. J-coupling Hamiltonian can be written as

$$\widehat{H}_J = 2\pi J_{12} \hat{I}_{1z} \hat{I}_{2z} \quad \textbf{Eq 1.13}$$

Then, for spin 1 in situation (1):

$$\hat{I}_{1z} \xrightarrow{-\frac{\pi}{2}\hat{I}_{1y}} \hat{I}_{1x} \xrightarrow{2\pi\tau J_{12}\hat{I}_{1z}\hat{I}_{2z}} cos\pi\tau J_{12}\hat{I}_{1x} - sin\pi\tau J_{12} 2\hat{I}_{1y}\hat{I}_{2z} \xrightarrow{\pi\hat{I}_{1x}+\pi\hat{I}_{2x}}$$

$$cos\pi\tau J_{12}\hat{I}_{1x} - sin\pi\tau J_{12} 2\hat{I}_{1y}\hat{I}_{2z} \xrightarrow{2\pi\tau J_{12}\hat{I}_{1z}\hat{I}_{2z}} cos2\pi\tau J_{12}\hat{I}_{1x} - sin2\pi\tau J_{12} 2\hat{I}_{1y} \quad \textbf{Eq 1.14}$$

Likewise, for spin 2 in scenario (1):

$$\hat{I}_{2z} \xrightarrow{90^{\circ}_{-y} - \tau - 180^{\circ}_{x} - \tau} cos2\pi\tau J_{12}\hat{I}_{2x} - sin2\pi\tau J_{12}2\hat{I}_{2y}\hat{I}_{1z} \qquad \textbf{Eq 1.15}$$

The result shows that the coupling simply evolves for 2τ and generates anti-phase magnetization instead of being refocused.

However, in situations (2) and (3), the 180° pulse ($\pi\hat{I}_{1x}$) is only applied to one spin, causing a refocused J-coupling:

$$\hat{I}_{1z} \xrightarrow{-\frac{\pi}{2}\hat{I}_{1y}} \xrightarrow{2\pi\tau J_{12}\hat{I}_{1z}\hat{I}_{2z}} \xrightarrow{\pi\hat{I}_{1x}} \xrightarrow{2\pi\tau J_{12}\hat{I}_{1z}\hat{I}_{2z}} \hat{I}_{1x} \qquad \textbf{Eq 1.16}$$

Based on the discussion of spin echo, in heteronuclear systems, we are free to decide whether to permit the evolution of the offset and the coupling. This flexibility grants significant latitude in creating and controlling pivotal anti-phase states in multiple-pulse NMR experiments.

1.2.4 One- and Multi-Dimensional Correlation NMR

The progress in ssNMR techniques, including the utilization of high magnetic field spectrometers, high and ultra-high MAS probes, and the implementation of robust pulse sequences, has opened up avenues for acquiring a wide range of atomic-level information. These techniques offer insights into the structure, interactions, dynamics, and polymorphism of cell wall macromolecules.

The primary method for distinguishing magnetically inequivalent units and rapidly screening cell wall samples is one-dimensional (1D) ^{13}C ssNMR. By employing the initial magnetization as a dynamic filter, a diverse range of dynamically heterogeneous polysaccharides present in cell walls can be identified. Firstly, the Dipolar-based ^{1}H-^{13}C CP experiment[29, 30] primarily detects rigidly coupled molecules, and selective detection can be achieved by adjusting CP contact time (**Figure 1.4**). Secondly, the DP experiment is also extensively employed for directly polarizing ^{13}C spins, as illustrated in **Figure 1.5a**. When a long recycle delay, typically set

at 5 times of the T_1 (spin-lattice relaxation time) value, is employed, it allows for complete relaxation of the nuclear spins. It enables the detection of all molecules within the system, leading to quantitative results with both rigid and mobile molecules information. Conversely, the use of a short recycle delay is designed to selectively detect mobile molecules, such as those with shorter T_1 relaxation times. Furthermore, for the analysis of the most dynamic or even solvated molecules within a sample, the scalar-based (J-coupling) ^{1}H-^{13}C refocused-INEPT experiment[31] is a powerful and indispensable tool (**Figure 1.6a**). The ^{1}H magnetization of rigid segments experiences rapid loss, due to fast T_2 (spin-spin) relaxation caused by the strong ^{1}H-^{1}H dipolar coupling. In addition, MultiCP (**Figure 1.6b**) offer a time-efficient alternative to the long recycle delay DP experiment, facilitating quantitative detection, particularly in natural abundance samples[32]. This method accomplishes extended CP from ^{1}H nuclei, surpassing a duration of 6-10 milliseconds, all while minimizing magnetization losses due to relaxation. It operates with a moderate duty cycle of radio-frequency irradiation, achieved through a sequence of multiple 1-ms CP periods, which are interleaved with ^{1}H spin-lattice relaxation periods designed to repolarize the protons.

Multidimensional NMR plays a pivotal role in deciphering complex carbohydrate structures within cell walls, enhancing resolution and simplifying spectral analysis. We can categorize them into two groups: homonuclear and heteronuclear correlation experiments. Heteronuclear resonance assignments are crucial for comprehensive structure determination and intermolecular interactions exploration, such as the two-dimensional (2D) ^{1}H-^{13}C HETCOR[33, 34] (**Figure 1.6c**). During this experiment, ^{1}H homonuclear couplings are removed by Phase-Modulated Lee–Goldburg (PMLG) or frequency-switched Lee-Goldburg (FSLG) decoupling[35], and the resulting narrow lines for rigid-solid protons are observed after the t_1 evolution. By changing CP contact time that adjusts polarization transfer from ^{1}H to ^{13}C, intra- and

intermolecular interactions can be selectively detected. In addition, a mixing period that allows for ^{1}H-^{1}H spin diffusion (SD) can be added before the CP block. The HETCOR method is mainly used in the wetland soil project in **Chapter 9**. 2D ^{13}C-^{13}C homonuclear correlation experiments fall into two categories: through-space via dipolar coupling, such as CHHC, RFDR, PDSD, DARR, and CORD, and through-bond via J coupling (e.g., CP/DP-J-INADEQUATE). These experiments aid in resonance assignments, covalent linkage pattern determination, and estimating the molar percentage of sugar residues in matrix polysaccharides. Particularly, the CHHC experiment[36] is a prime choice for natural-abundance MAS-DNP, due to the ^{13}C–^{1}H–^{1}H–^{13}C pathway for polarization transfer. It effectively resolves the challenge of diagonal peak dominance in the spectrum of unlabeled materials, which dominance arises from the limited probability of observing off-diagonal internuclear correlations at natural ^{13}C abundance levels. The CHHC experiment has been applied to study the unlabeled rice stems in **Chapter 8**.

In order to provide a clearer demonstration of functionality of multidimensional NMR, here I will employ one sugar residue of 3-fold xylan (Xn) and one lignin guaiacyl (G) unit as illustrative examples, which exampled chemical shifts are shown in **Figure 1.6e**. Typically, at the outset of our projects, following the initial 1D screening, we acquire 2D INADEQUATE spectra to assign chemical shifts for each component within the system (**Figure 1.6f**). The 2D ^{13}C-J-INADEQUATE method[37] is a powerful NMR technique that provides double quantum (DQ)-single quantum (SQ) correlations and offers a comprehensive producing diagonal-free asymmetric connectivity map for all carbon atoms within the molecule, making it a definitive tool for structural assignments of biomolecules. In the generated refocused-INADEQUATE spectrum of xylan, which displays unambiguous through-bond correlations via J-coupling, the signals of two directly bonded carbon atoms are located at coordinates (δCa, δCa + δCb) and (δCb, δCa + δCb) on the

plot (**Figure 1.6f**). Here, δCa and δCb represent the SQ chemical shifts of carbon sites a and b, respectively. These values are projected onto the X-axis, corresponding to the direct dimension (ω2) in the experimentally measured NMR spectra. The sum of these two SQ chemical shifts (δCa + δCb) is projected onto the Y-axis, representing the DQ chemical shift for this carbon pair. This corresponds to the indirect dimension (ω1) in the experimentally measured NMR spectra. Each pyranose unit in xylan is depicted by ten scatter dots based on the chemical shifts of seven carbons, such as the spin pairs of the carbonyl group (Ac^{CO} at 174 ppm and Ac^{Me} at 21 ppm).

Via the evolving of density operator, we can prove how the DQ coherence is created. In the INADEQUATE pulse sequence, the transverse ^{13}C-magnetization is generated by the first 90° pulse (or CP). If we set $\tau = 1/(4J_{IS})$, after the first τ - 180° - τ period we can get:

$$\rho(0) \xrightarrow{\tau - 180^\circ - \tau} -2\hat{I}_x\hat{S}_z - \hat{I}_z\hat{S}_x \qquad \textbf{Eq 1.17}$$

The DQ coherence is created during the following 90° pulse for t_1 evolution:

$$-2\hat{I}_x\hat{S}_z - \hat{I}_z\hat{S}_x \xrightarrow{90^\circ_x} -2\hat{I}_x\hat{S}_y - \hat{I}_y\hat{S}_x \qquad \textbf{Eq 1.18}$$

Finally, the third 90° pulse converts the DQ coherence to SQ one, followed by the second τ - 180° - τ sequence that establishes an in-phase coherence for t_2 detection, where $\tau = 1/(4J_{IS})$:

$$-2\hat{I}_x\hat{S}_y - \hat{I}_y\hat{S}_x \xrightarrow{90^\circ_x} 2\hat{I}_x\hat{S}_z + \hat{I}_z\hat{S}_x \xrightarrow{\tau - 180^\circ - \tau} \hat{I}_y + \hat{S}_y \qquad \textbf{Eq 1.19}$$

Similar to 1D experiments, the interplay between direct polarization and J-coupling predominantly highlights mobile segments, making the DP J-INADEQUATE experiment inclined toward dynamic signals. In parallel, the analogous version of the CP J-INADEQUATE experiment is designed to exclusively capture rigid signals.

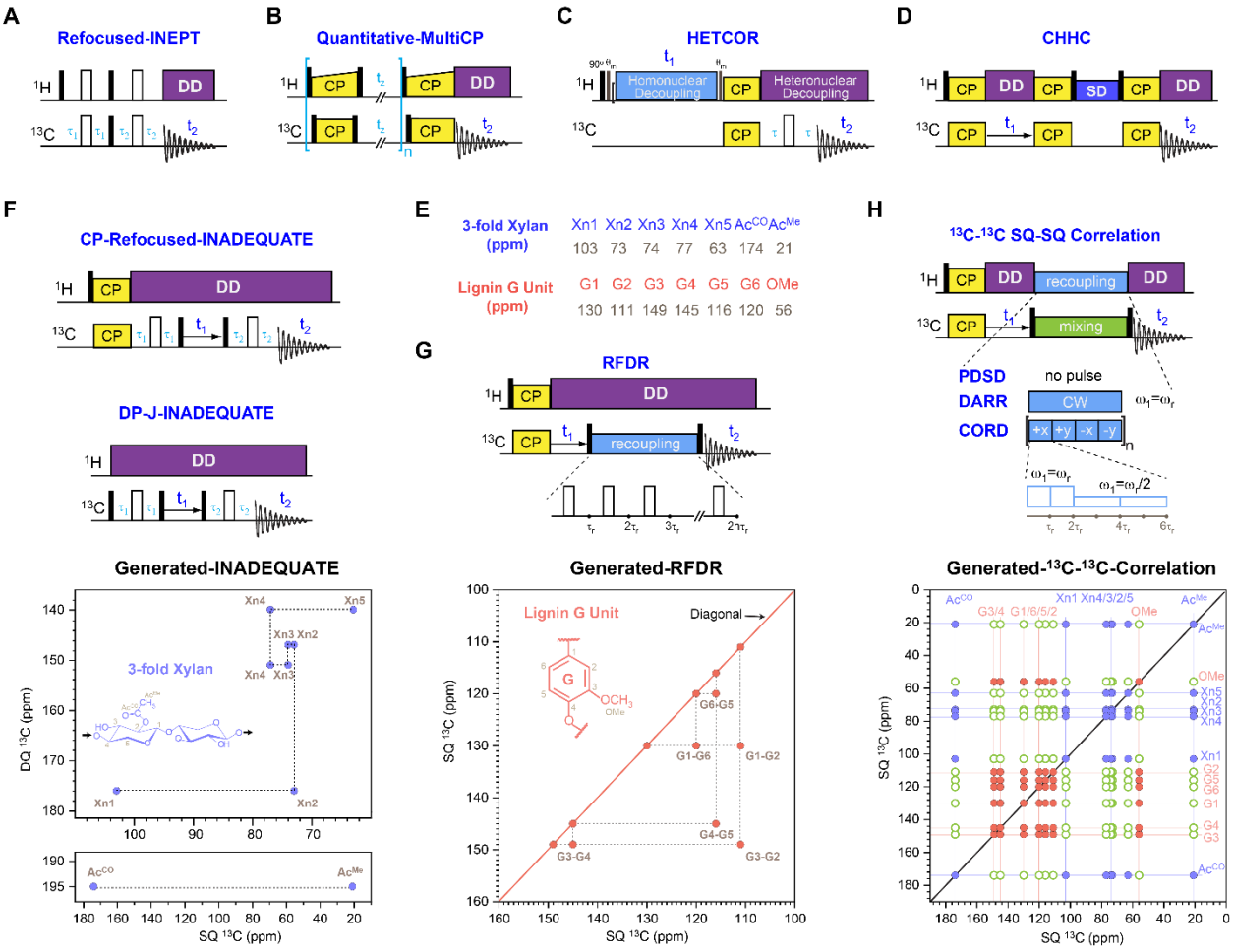

Figure 1.6 Several typical ssNMR experiments. **a,** 1D Insensitive nuclei enhancement by polarization transfer (INEPT) experiment with heteronuclear dipolar decoupling (DD). **b,** 1D MultiCP. **c,** 2D ^{1}H–^{13}C heteronuclear chemical shift correlation (HETCOR). **d,** 2D CHHC. **e,** Chemical shifts of exampled molecules in (f), (g), and (h). **f,** 2D CP/DP based J- incredible natural abundance double quantum transfer experiment (INADEQUATE) and a generated spectrum of one xylan residue. **g,** 2D rotational echo double resonance (REDOR) and a pseudo spectrum of one unit of G lignin. **h,** 2D ^{13}C–^{13}C correlation, such as proton-driven spin diffusion (PDSD), dipolar assisted rotational resonance (DARR), and combined $R2_n^{\upsilon}$-driven (CORD). In the generated spectrum, intramolecular interactions inside of xylan and G lignin are colored blue and magenta, respectively, while intermolecular peaks are denoted by green dots.

Another commonly employed pulse sequence for structural assignments is RFDR[38, 39] (**Figure 1.6g**), which direct recouples ^{13}C-^{13}C dipolar coupling for polarization transfer (commonly referred to as first-order recoupling). Though it is categorized as through-space methods, it is particularly useful for selectively identifying short-range (one or two-bond) correlations by effectively suppressing multi-bond interactions with brief recoupling time (~1-2 ms). This technique utilizes multiple 180° pulses (2n) on the ^{13}C channel to reintroduce ^{13}C-^{13}C homonuclear dipolar couplings through longitudinal polarization transfer. Each pulse should be precisely positioned at the midpoint of a single rotor period. The pseudo-RFDR spectrum illustrates the anticipated one-bond connections within a single G-lignin unit. The graph clearly reveals that every pair of bonded carbon atoms exhibits mutual cross-peaks. However, the OMe group, lacking immediate neighboring carbon atoms, remains devoid of any discernible signals.

Upon obtaining structural assignments, we can delve into the exploration of intra- and inter-molecular interactions using 2D ^{13}C-^{13}C correlation experiments that rely on dipolar recoupling during the mixing phase, as illustrated in **Figure 1.6h**. Short mixing times are valuable for detecting intra-residue correlations, aiding in resonance assignment. In contrast, longer mixing times unveil intermolecular cross peaks, offering insights into sub-nanometer molecular packing. **In Figure 1.6h**, we generate a pseudo ^{13}C-^{13}C correlation spectrum, where intramolecular interactions within xylan and G lignin marked in blue and magenta, respectively. Possible intermolecular peaks between xylan and G lignin are indicated by green dots, applying a longer mixing time.

The foundation of these 2D ^{13}C-^{13}C correlation experiments lies in the reintroduction of ^{13}C-^{1}H dipolar couplings during the mixing phase, often referred to as second-order recoupling. Under moderate MAS conditions and without active recoupling, the residual ^{13}C-^{13}C dipolar

couplings are minimal, leading to slow polarization transfer. However, carbon-to-carbon interactions become viable through their interactions with surrounding protons. This is achieved by temporarily deactivating proton decoupling or employing active recoupling with radiofrequency power, thus reintroducing ^{13}C-^{1}H dipolar couplings (known as recoupling). This action broadens the line shape of carbon peaks and enables spectral overlap during the mixing period. Spectral overlap can occur between two carbon peaks with small chemical shift differences or between a carbon peak and the spinning sidebands of a distant carbon peak, preserving energy for polarization transfer. These experiments are less affected by severe dipolar truncation effects, making them particularly suitable for measuring long-range correlations.

The thesis explores several common 2D ^{13}C-^{13}C correlation experiments, including PDSD, DARR, and CORD[40-42]. PDSD, for instance, offers a straightforward experimental setup and the capability to investigate both short-range and long-range correlations. In PDSD, the mixing period does not require radiofrequency irradiation, allowing for extended mixing times of several seconds without concerns about radiofrequency-induced heating. The rate of spin diffusion buildup serves as a qualitative indicator of internuclear distances, providing semi-quantitative insights into spatial proximities. PDSD is most effective at slower spinning frequencies (around 10 kHz) and lower magnetic fields (less than 14 T). However, its efficiency decreases notably at higher MAS frequencies and magnetic fields. To partially address these limitations, the DARR sequence can be employed by applying RF irradiation on the ^{1}H channel ($\omega 1 = \omega r$). Nevertheless, hardware constraints, such as sample heating by the RF pulse, limit the duration of the mixing period, rendering DARR transfer inefficient when the MAS frequency surpasses 30 kHz. The CORD sequence has been developed to facilitate spin diffusion at faster MAS and features broad-band homonuclear dipolar recoupling. CORD spectra exhibit more consistent cross-peak intensities

across the spectrum, rendering them advantageous for homonuclear correlation studies in biological samples.

The flexibility to combine various sequence blocks makes ssNMR a remarkably versatile technique for structural exploration. Pulse sequences like INADEQUATE, PDSD, and RFDR serve as fundamental building blocks for creating more intricate pulse sequences[43], such as 3D DQ-SQ-SQ experiment[44] introduced in **Chapter 5**. It is comprised of a dipolar-base INADEQUATE block followed by a CORD block, which enables unambiguous resonance assignments of complex carbohydrates in cell walls. Moreover, the incorporation of dipolar coupling, J-coupling, and relaxation filters is instrumental in spectral-editing methods, which enable the selective detection of specific sites within a molecule or specific molecules within a sample.

For instance, the water-edited experiment[4, 45] (**Figure 1.7a**) has become a valuable tool for understanding the role of water in biological systems. These experiments allow for the selective detection of water-proximal sites or interactions, shedding light on the dynamic behavior of water molecules and their influence on the structure and function of macromolecules. Typically, by using relaxation filters, like the ^{1}H-T_2 filter, water signals can be selectively isolated. This is based on the dynamic nature of water's ^{1}H magnetization, which enables it to persist through a moderately long ^{1}H-T_2 filter (in the order of milliseconds). Then, the selected ^{1}H magnetization from water is manipulated and transferred to the neighboring ^{1}H in biomolecules through ^{1}H-^{1}H spin diffusion, further propagated to ^{13}C nuclei through CP. The resulting spectra provide crucial insights into the local environment and interactions of water molecules with biomolecules.

Similarly, the Dipolar-Gated PDSD experiment[7, 14] is employed to enhance the detection of aromatic signals within lignin, which are typically challenging to discern in plant cell walls due

to the substantial CSA in ssNMR. This specialized experiment is designed to selectively capture lignin signals while minimizing interference from proton-rich polysaccharides. The key innovation in this approach involves reintroducing ^{13}C-^{1}H dipolar coupling during an asymmetrical dipolar-dephasing period before t_1 evolution. This period comprises two un-decoupled delays with respect to the 180° pulse in the Hahn echo. This strategic design efficiently targets non-protonated carbons within the aromatic lignin, while simultaneously suppressing signals originating from protonated carbons. As illustrated in **Figure 1.7b**, when compared to conventional 1D CP or 2D DARR experiments, both the 1D and 2D Gated-PDSD experiments exhibit a preference for detecting non-protonated carbons within lignin and alleviate the spectral asymmetry introduced by the heterogeneous proton density within the aromatic motifs.

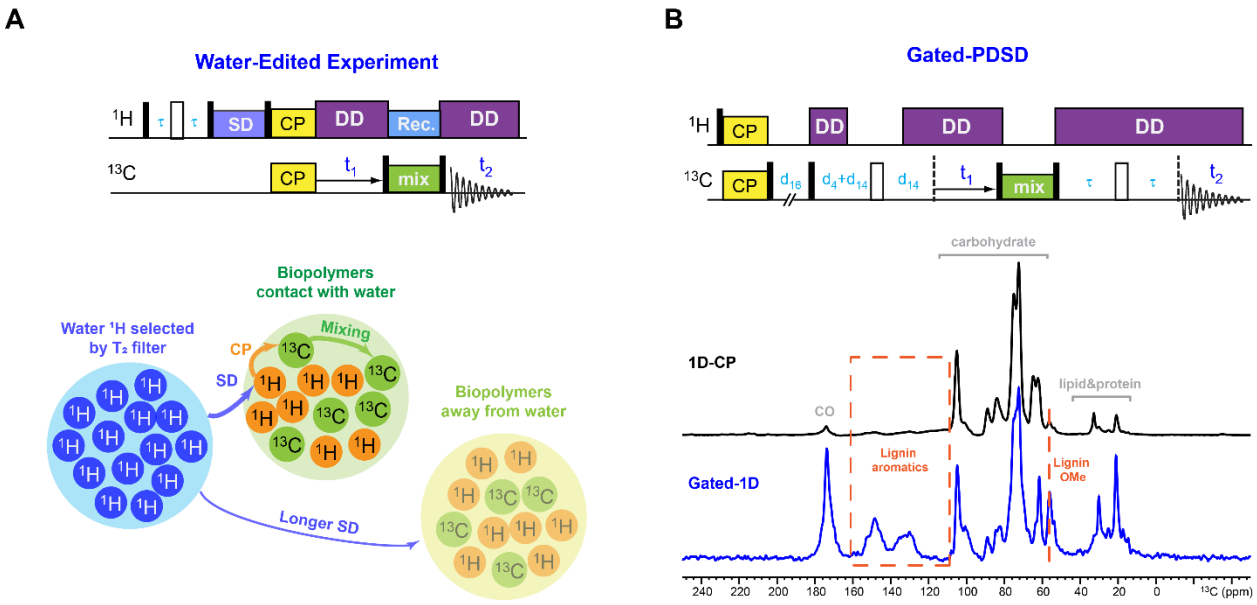

Figure 1.7 Edited-NMR experiments. **a,** Water-edited experiment with ^{1}H-T_2 filter for selective detection of biopolymers in close proximity to water. **b,** Dipolar-gated ^{13}C-^{13}C correlation experiment with dipolar-dephasing designed to selectively detect non-protonated carbons within the aromatic lignin.

1.2.5 Molecular Motions: Rates and Amplitudes

Beyond the fundamental principles, NMR encompasses relaxation processes that govern the behavior of nuclear spins and the properties of NMR signals[42, 46, 47]. These processes are broadly categorized as longitudinal relaxation (T_1) and transverse relaxation (T_2). T1 relaxation, also called spin-lattice relaxation, dictates the rate at which nuclear spins revert to their equilibrium positions along the magnetic field after being perturbed by RF pulses. The T_1 relaxation time offers insights into the dynamics and interactions of molecules, providing information about molecular motion and the correlation times of molecular processes. T_2 Relaxation, also named spin-spin relaxation, accounts for the loss of phase coherence among nuclear spins in the transverse plane. This relaxation process directly influences the width of NMR signals and, consequently, the resolution of NMR spectra. T_2 plays a vital role in NMR applications where high-resolution spectra and well-defined line shapes are desired. Understanding these relaxation processes is critical for tailoring NMR experiments to the specific properties of the sample under investigation.

Various relaxation processes in ssNMR spectroscopy exhibit sensitivity to molecular motions occurring on distinct timescales. ^{13}C-T_1 Relaxation primarily reflects motions occurring on the nanosecond timescale, which is commonly measured by the DP-based inversion recovery experiment (**Figure 1.8a**). In a nutshell, this experiment involves flipping the initial magnetization to the negative z-direction through a 180° pulse. Giving a specific delay period (τ), the magnetization is expected to relax back towards the equilibrium magnetization in the positive z-direction. Then, a following 90° pulse flips the magnetization to the x-y plain for further detection. In the case of a single exponential fitting, the following equation is typically employed to quantitatively find the T_1 relaxation constant:

$$I = 1 - 2e^{-t/T_1} \qquad \textbf{Eq 1.20}$$

^{13}C-T_1 Torchia sequence[48] is also a powerful tool for the measurement of T_1 relaxation times (**Figure 1.8b**). This method leverages CP to generate transverse magnetization in ^{13}C nuclei, which favors the detection of rigid species within the sample, and subsequently flips magnetization to the positive-z direction, allowing spin-spin relaxation via varying delay τ. After the second 90° pulse, we can monitor how rapidly the magnetization returns to equilibrium, using the equation for a single exponential fitting:

$$I = e^{-t/T_1} \qquad \textbf{Eq 1.21}$$

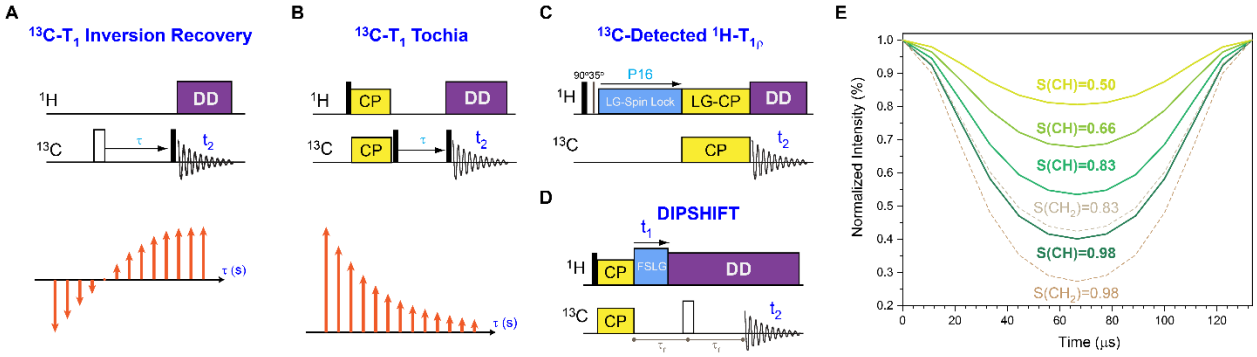

Figure 1.8 Molecular motion measurements. **a,** DP-based ^{13}C-T_1 inversion recovery. **b,** CP-based ^{13}C-T_1 Torchia. **c,** ^{13}C-detected ^{1}H-$T_{1\rho}$ pulse sequence with LG-spin lock. **d,** 2D ^{1}H-^{13}C dipolar chemical-shift (DIPSHIFT) experiment. **e,** representative ^{1}H-^{13}C dipolar dephasing curves under 7.5 kHz MAS. Corresponding order parameters are labeled based on CH or CH_2 group.

^{1}H-$T_{1\rho}$ Relaxation[16], also referred to proton spin-lattice relaxation in the rotating frame, provides insights into motions on the microsecond timescale (**Figure 1.8c**). In the beginning, a 90° pulse followed by a 35° pulse is applied to flip the ^{1}H magnetization to the magic angle. Following, a spin-locking field with a frequency in the range of tens of kHz is applied to the proton channel, which prevents precession in the rotating frame. Then, the LG-CP technique is employed to selectively transfer the ^{1}H polarization to the directly bonded ^{13}C nuclei for the detection. The

measurement of ^{13}C-T_2 relaxation[42] can be accomplished using the CP echo experiment, with the concurrent application of ^{1}H heteronuclear decoupling. The presence of decoupling is essential in this context to prevent the destruction of magnetization due to ^{13}C-^{1}H dipolar coupling. This differs from T_1 measurements, where ^{1}H decoupling is not necessary.

In ssNMR, the determination of motional amplitudes in polysaccharides is often achieved through the employment of the DIPSHIFT experiment[16, 35], as illustrated in **Figure 1.8d**. The order parameters, denoted as S(CH), derived from this experiment offer valuable information regarding the motional amplitudes of the C-H bonds in the biomolecules under investigation (**Figure 1.8e**). Notably, a smaller order parameter corresponds to more significant amplitude in molecular motions, signifying increased flexibility or mobility within the C-H bonds. Conversely, a larger order parameter indicates a rigid molecular structure, such as cellulose crystalline regions. In the DIPSHIFT experiment, to realign chemical shifts, a 180° pulse is strategically inserted after a single rotor period. Moreover, the removal of ^{1}H-^{1}H homonuclear couplings is meticulously accomplished using decoupling sequences like FSLG. This precise setup ensures that the ^{13}C nuclei exclusively evolve under the influence of ^{13}C-^{1}H dipolar couplings. The resultant spectrum in this experiment displays ^{13}C isotropic chemical shifts in the direct dimension, while the indirect dimension reveals ^{13}C-^{1}H dipolar dephasing curves. These dephasing curves are subjected to fitting procedures to extract the apparent ^{13}C-^{1}H dipolar couplings. To obtain the true dipolar coupling, accounting for molecular motion, these apparent couplings are divided by scaling factors associated with the homonuclear decoupling sequences, typically set at 0.577 for FSLG. Subsequently, the true dipolar coupling is further divided by the theoretical rigid-limit ^{13}C-^{1}H one-bond dipolar coupling (~ 22.7 kHz) to get the order parameters. It's worth noting that the CH_2 group has a deeper dephasing curve when compared to that of the CH group, despite having

identical order parameters, as depicted in **Figure 1.8e**.

1.2.6 Resolution and Sensitivity

Two critical aspects of NMR spectroscopy are sensitivity and resolution, which play pivotal roles in the quality and applicability of ssNMR experiments. Resolution in NMR spectroscopy pertains to the ability to distinguish and separate individual resonances in the NMR spectrum. Resolution, often characterized by parameters like the full width at half maximum (FWHM), is inversely related to the apparent T_2 relaxation time. Enhancing resolution can be achieved by promoting crystallinity within the sample, thus slowing down intrinsic T_2 relaxation. Improved coil design and careful sample positioning contribute to reducing field inhomogeneity, resulting in longer apparent T_2 values. Biological samples often face resolution limitations due to peak overlap resulting from the structural similarity of their constituent units. Strategies to address this challenge include selective and sparse labeling of samples. On the spectral side, advanced pulse sequences, including extra dimensions (3D) and spectral editing, offer better resolution and sensitivity, allowing for the analysis of complex samples. Residual anisotropic interactions, such as lingering ^{1}H-^{1}H homonuclear dipolar couplings, can cause peak broadening that standard spinning cannot fully eliminate. Ultra-fast spinning[49-51] at rates ranging from 40 to 120 kHz using small rotors (e.g., 0.7 mm) significantly narrows the ^{1}H linewidth, enabling ^{1}H detection of nanomole to femtomole quantities in deuterated proteins under high magnetic fields.

In the data analysis stage, there are opportunities to further enhance resolution, although these improvements need to scarify sensitivity. Various window functions can be applied to time domain signals to enhance resolution by focusing on the later part of the FID. The choice of window functions, such as Gaussian, exponential, sine, and qsine, should be made thoughtfully to achieve the desired balance between sensitivity and resolution. Another technique involves zero-

filling, which increases digital resolution by appending zeros to the end of the FID.

NMR signals originate from a minor surplus of nuclei occupying the lower energy state, with nuclear spins distributed almost evenly at room temperature. This distribution follows the Boltzmann distribution equation:

$$\frac{N_+}{N_-} = e^{-\Delta E/kT} \quad \textbf{Eq 1.22}$$

Here, k is the Boltzmann constant. At room temperature and within a 10 T magnetic field, the population of the lower energy state (N_+) exceeds that of the higher energy state (N_-) by a factor of only 1 in 10,000. This small difference in population results in inherently weak NMR signals. Consequently, NMR spectra exhibit low signal-to-noise ratios (S/N) due to these weak signals.

$$S/N \propto n\gamma_E T^{-1} f_{CP}\eta(\gamma_d^3 B_0^3 t)^{\frac{1}{2}}(VQ)^{\frac{1}{2}} \quad \textbf{Eq 1.23}$$

This equation[52] offers numerous avenues for enhancing NMR sensitivity. The "n" represents the number of spins. Hence, we can increase the sensitivity by achieving isotope-labeling (^{12}C to ^{13}C), packing more samples into the rotor, or switching to a thin-walled or a larger rotor to contain more spins. Lowering the temperature (T) by employing cryogenic NMR probes and measuring in a higher magnetic field (B_0) are another two straightforward approaches. For example, an 800 MHz spectrometer provides a 2.8-time higher sensitivity than a 400 MHz spectrometer. Additionally, "f_{CP}", "η", "V", and "Q" are CP factor, probe-specific factors, signifying the coil filling factor, coil volume, and coil quality factor, respectively.

Another strategy for sensitivity improvement is to optimize the effectiveness of the experimental time factor "t". 4 times of the number of scans (NS) in experiments provide 2-fold sensitivity enhancement. By introducing paramagnetic agents[53, 54] into the samples, it is possible to reduce the T_1 relaxation times, further, to reduce the recycle delay (typically 5 times T_1), through the paramagnetic relaxation enhancement effect. This approach significantly accelerates data

acquisition cycles, resulting in time savings of approximately 5-20 folds. Non-uniform sampling[55], a technique still under development, holds promise as a method to better allocate the acquisition time during data sampling. This approach enhances sensitivity without sacrificing resolution.

In the equation, "Υ_E" and "Υ_D" stand for the gyromagnetic ratios of the excited and detected nucleus, respectively. It is the reason why a ssNMR pulse sequence always starts with a CP block to enhance sensitivity by exciting ^{1}H instead of ^{13}C. Furthermore, direct ^{1}H-detection contributes to increased sensitivity, owing to the high gyromagnetic ratio of the proton nucleus (Υ_{1H}=4Υ_{13C}). It is the reason that the ^{13}C-^{1}H HETCOR has an 8 times enhancement when compared to ^{1}H-^{13}C HETCOR. DNP is a cutting-edge innovation that facilitates the transfer of polarization from electrons (excitation) to NMR-active nuclei. As a result, sensitivity enhancements of tens to hundreds of folds can be achieved through DNP[56]. This also translates to substantial time savings, reducing experimental durations by a factor of 10^2 to 10^5.

1.3 Sensitivity-Enhancing DNP

In recent developments, the innovative technique of MAS-DNP has demonstrated remarkable potential in addressing a long-standing challenge[21, 57-60]. It is widely acknowledged that NMR spectroscopy suffers from limited sensitivity due to the minute nuclear spin polarization, accounting for less than 0.007% of ^{1}H, resulting from the small gyromagnetic ratio (γ) of the spins. MAS-DNP has the remarkable capability to facilitate polarization transfer from unpaired electrons, which are approximately 660 times larger in γ than ^{1}H (**Table 1.1**), to protons when subjected to microwave irradiation (as illustrated in **Figure 1.9**). This transfer leads to a substantial enhancement in NMR sensitivity, typically ranging from one to two orders of magnitude. The essential microwave resource in this process is the gyrotron, which generates power in the range of 10 to 100 watts. This power is harnessed to induce electron paramagnetic resonance[60] (EPR) in

mono- or bi-radicals, each possessing one or two unpaired electrons[61-63]. To facilitate this process, nitroxide-based radicals (for instance, AMUPol[64]) are commonly dissolved in DNP matrix solutions (such as the mixture of D_2O, H_2O, and ^{13}C-depleted, d_8-glycerol). These solutions are then combined with the solid sample through grinding or vertexing. In order to achieve optimal results, MAS-DNP experiments necessitate cryogenic temperatures within the range of 90–110 K. These low temperatures are essential to circumvent the challenge of the short lifetime of radicals and further improve the S/N ratios by creating a glassy phase within the DNP matrix. This glassy phase efficiently prevents the aggregation of radicals.

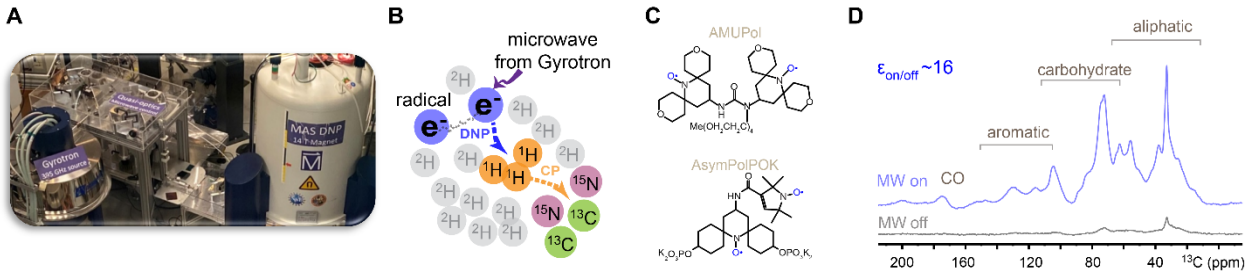

Figure 1.9 Sensitivity-enhancing DNP. **a,** DNP instruments. **b,** DNP mechanism. **c,** DNP radicals. **d,** 16-fold sensitivity enhancement of unlabeled soil samples provided by 600 MHz/395 GHz DNP.

Numerous features can affect the enhancement factor (ε), such as radicals, solvents, and even rotor materials. AsymPolPOK[65], an asymmetric biradical that been reported recently, is capable to significantly shorten DNP build-up time, leading to a larger enhancement factor and a shorter experiment time; TEKPol dissolved in ortho-terphenyl enables DNP experiment conducted at near room temperature, which benefits for the detection of matrix polysaccharides that is limited at cryogenic temperature; TinyPol and TEMTriPol-I have a good performance on high-field (18.8T, 800 MHz/527 GHz) DNP instruments with resolution improvement. The DNP matrix protocols also influence the distribution of radicals in cell wall compounds and impact the

enhancement factor by formation or separation of the glassy phase of solvents. The broadly used protocols in cell wall studies include ^{13}C-depleted, d_8-glycerol/D2O/H2O (60/30/10 vol%), D_2O/H_2O (90/10), d_6-DMSO/D_2O/H_2O (60/30/10), and even a matrix-free protocol (only D_2O), which can effectively avoid radical aggregation[66].

With the remarkable sensitivity enhancement brought about by the DNP technique, it becomes feasible to perform high-resolution structural analysis of unlabeled biological systems using ssNMR. In this dissertation, with assistance of MAS-DNP, (1) a comprehensive toolbox for systematically illustrating structural models in intact natural abundance plant biomass is established (**Chapter 8**); (2) molecular composition of organic carbon within wetland soil samples obtained from a brackish island situated along the Gulf of Mexico coastline is analyzed (**Chapter 9**); (3) unexpected lignin–cellulose interactions in ^{13}C-wood cell walls are confirmed (**Chapter 7**).

1.4 Sample Preparation for ssNMR and DNP Measurements

To conduct ssNMR studies on plant cell wall samples, it is essential to enrich the samples with NMR-active isotopes, such as carbon-13. This can be achieved through two distinct methods, each offering specific advantages based on the research focus.

The first method[67, 68] typically involves the enrichment of young seedlings (aged up to 2 weeks) for the study of carbohydrates in primary cell walls. These seedlings are grown in darkness in a solution medium containing ^{13}C-glucose or in sealed chambers connected to $^{13}CO_2$ resources. After harvesting, the seedlings are rapidly frozen and then ground using liquid nitrogen. Subsequently, a 1.5% (w/v) sodium dodecyl sulfate (SDS) solution is used for a 3-hour wash to remove cell membranes, proteins, and inactivate endogenous wall-degrading enzymes. Starch is eliminated by incubating the sample with α-amylase in a sodium metabisulfite (MES) buffer (pH 6.8). To further break down proteins, Pronase is employed in a sodium MES buffer (pH 7.5).

Throughout these processes, a solution containing 0.02% (w/v) NaN_3 is used to prevent microbial growth. After washing and centrifugation, the resulting never-dried cell walls (approximately 30 mg) can be packed into 3.2-mm rotors for ssNMR experiments.

The second isotope-labeling method[14, 69] provides greater versatility in terms of plant species and developmental stages, including plants like maize and woody species (several months old) for the study of secondary cell walls. In this approach, plants are cultivated in sealed chambers using a hydroponics system to prevent $^{12}CO_2$ contamination from soil microbial activity. The composition of the hydroponic solution is adjusted to suit the specific species, supporting minerals, and pH requirements. A gas mixing system combines 99.8% $^{13}CO_2$ gas with a 20% O_2 in N_2 gas mix, resulting in a final concentration of 600 ppm of $^{13}CO_2$. Ideal LED lights with specific wavelengths and intensity are used to minimize heat generation, and humidity and temperature are carefully controlled. After harvesting, fresh plant tissues are cut into small pieces and evenly packed into rotors for ssNMR analysis.

Sensitivity-enhanced DNP has opened doors to studying wetland soil or plant samples in their natural and unlabeled state (1.11%-^{13}C). This method[57, 70] involves the integration of samples with a suitable DNP solvent, commonly referred to as the "matrix," along with the inclusion of a polarizing agent. These polarizing agents, often based on nitroxides, such as AMUPol, TOTAPOL, and AsymPolPok, are distinguished for their unreactive nature and their affinity for dissolution in various aqueous solvents. In a carefully orchestrated procedure, a concentration typically ranging from 5 to 20 mM of these radicals is expertly dissolved in cryoprotective solvents. These solvents encompass a selection of options, including ^{13}C-depleted d_8-glycerol, D_2O, and H_2O, d_6-DMSO in combination with D_2O and H_2O, d_6-DMSO with H_2O alone, or a matrix-free approach that relies solely on D_2O. The resulting DNP solution, imbued with the distinct properties of these noble

radicals, is artfully blended with unlabeled samples, ensuring a deep integration that permits these radicals to permeate the plant cell walls or allows a homogeneous mixing of radicals with soil. These well-blended samples are subsequently packed into 3.2 mm zirconium or sapphire rotors. These rotors embrace a payload of roughly 30 to 50 milligrams, serving as the staging ground for the forthcoming DNP experiment. The optimization of the enhancement factor ($\varepsilon_{on/off}$) and the precise calibration of the DNP buildup time are exacting and tailored to the unique characteristics of each sample. In particular, soil samples may require an additional treatment with a 2% hydrofluoric acid (HF) solution[30] before mixing with the DNP solution to remove mineral components without altering their chemical structure.

1.5 Thesis Organization

My dissertation primarily comprises published papers and preprint manuscripts during my Ph.D. research, focusing on the development and application of ssNMR, with the assistance of DNP, to investigate plant cell walls and wetland soil samples. Besides sections of abstract, acknowledgment, and table of contents, another 11 chapters are also structured into the thesis. An overview of each chapter is specified below.

Chapter 1 serves as an introductory chapter, presenting background information from four key aspects: (1) the importance of studying plant cell walls and wetland soil; (2) an introduction to ssNMR theory and methodology; (3) sensitivity-enhanced DNP and natural abundance samples. (4) ssNMR sample preparation and isotope labeling. Additionally, subsections of thesis organization and copyright permissions are included.

Chapters 2, 3, and **4** are review papers, as additional introductions, which highlight the recent breakthroughs in using ssNMR to elucidate cell walls and summarize underexplored aspects of cell wall studies. **Chapter 2** reviews the research on carbohydrate and cell walls using ssNMR

and discusses the potential to address unresolved questions by innovative methods, such as proton detection, quadrupolar nuclei, paramagnetic relaxation enhancement, DNP, ultrahigh field, and database. **Chapter 3** focuses on the current limitations of ssNMR in cell wall research, including challenges in resonance assignment due to structural complexity, the need for more effective methods to handle the ever-growing dataset, and the optimization of protocols and sample preparation for the application of high-field DNP. **Chapter 4** summarizes how ssNMR has contributed to our understanding of the structure of extracellular matrices, such as the dynamic structure of living fungal cell walls and macro fibrillar structure in intact plant tissues. The chapter also discusses the combination of NMR with other techniques to address unresolved structural questions.

Chapters 5 and **6** are published research articles, focusing on pulse programs that serve as the foundation of NMR experiments. In **Chapter 5**, we introduce a three-dimensional (3D) ^{13}C-^{13}C-^{13}C correlation experiment designed to enable high-resolution analysis and unambiguous resonance assignments of complex carbohydrates in cell walls. This ssNMR 3D pulse program incorporates a CORD block at the end of an INADEQUATE block, offering an unprecedented resolution for understanding the polymorphic structure of chitin in fungi samples and cellulose in plant cell walls. In **Chapter 6**, we discuss how the phases of radio frequency pulses affect signal phases during cross polarization. Through a first principles approach, we reveal that the NMR signal is simply dependent on arbitrary CP RF phases under both static and magic angle spinning conditions.

Lignin is a polyphenolic biopolymer deposited during the formation of plant secondary cell walls, which provides strength, assists water transporting, and resists microbial attacks. Gaining a comprehensive understanding of the lignin-carbohydrates interactions is pivotal for crop

development and enhancing the quality and durability of wood products. In **Chapter 7**, we achieve the atomic-resolution characterization of the polymorphic structure, dynamical and hydration profile, and physical packing of lignin and polysaccharides using intact plant cells in hardwoods (eucalyptus and poplar) and softwood (spruce). Our findings uncover that the extent of glycan-aromatic association increases sequentially across grasses, hardwoods, and softwoods; and lignin principally packs with the xylan in a non-flat conformation via electrostatic interactions and partially binds the junction of flat-ribbon xylan and cellulose surface as a secondary site.

The emerging sensitivity-enhancing technique DNP holds great promise for investigating natural abundance samples, eliminating the need for isotope labeling. **Chapters 8** and **9** provide two examples showcasing the application of DNP to unlabeled rice and wetland soil samples, respectively. In **Chapter 8**, we demonstrate the feasibility of combining traditional ssNMR with DNP methods to examine the structure and dynamics of carbohydrates in natural abundance rice stems. This time-efficient approach allows for the screening of a large collection of biomass materials without labeling. In **Chapter 9**, DNP enables rapid acquisition of 2D $^{13}C/^{1}H$-^{13}C correlation spectra to detail the molecular fingerprint and spatial organization of biopolymers in unlabeled wetland soil. Surprisingly, the lignocellulosic core identified in the plants grown on this island were found to be preserved in the surface layer of soil formed in the brackish marshes, but with tighter physical packing between molecules. Extending the depth from the surface to 2 m covers a geological timeline of eleven centuries, where we found sophisticated changes in the molecular structure and composition of organic matters as well as the bulk properties of the wetland soil.

Chapters 10 and **11** focus on statistical approaches for the analysis of carbohydrate composition and structure using high-resolution ssNMR data, which are indexed in a carbohydrate

database. **Chapter 10** introduces the Complex Carbohydrates Magnetic Resonance Database (CCMRD, www.ccmrd.org), as the first ssNMR database specifically designed for complex carbohydrates. This chapter also presents the assembly of the database, information search functionalities, and the deposition of data into CCMRD. In **Chapter 11**, We generate simulated spectra (density maps) using data deposited in CCMRD, to demonstrate the chemical shift dispersion and to aid in the fast identification of important fungal and plant polysaccharides in cell walls. By comparing the projected spectra with experimental data, we highlight the challenges of assigning resonances because distinct carbohydrates from different organisms can produce nearly identical signals.

1.6 Copyright Permissions

Chapters 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11 consist of reprinted published papers. Elsevier does not necessitate a license for re-use in a dissertation or thesis. Papers published in Biomedcentral (BMC), Springer Nature (*Nature Communication*), and John Wiley and Sons (*Magnetic Resonance in Chemistry*) are available in open access, and no permission is necessary for their re-use. Detailed copyright permissions have been secured from the following groups:

Chapter 2: Elsevier.

Chapter 3: Elsevier.

Chapter 4: Elsevier.

Chapter 5: Elsevier.

Chapter 6: Elsevier.

Chapter 7: Springer Nature (*Nature Communication*), open access.

Chapter 8: Biomedcentral (BMC), open access.

Chapter 9: Submitted to Nature Geoscience.

Preprinted in ChemRxiv (https://doi.org/10.26434/chemrxiv-2023-p4rvg).

Chapter 10: Springer Nature. License Number: 5663191296494.

Chapter 11: John Wiley and Sons (Magnetic Resonance in Chemistry), open access.

REFERENCES

1. Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D., Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007,** *315* (5813), 804-807.

2. Petridis, L.; Smith, J. C., Molecular-level driving forces in lignocellulosic biomass deconstruction for bioenergy. *Nat. Rev. Chem.* **2018,** *2* (11), 382-389.

3. Zhang, Y.; Yu, J.; Wang, X.; Durachko, D. M.; Zhang, S.; Cosgrove, D. J., Molecular insights into the complex mechanics of plant epidermal cell walls. *Science* **2021,** *372* (6543), 706-711.

4. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water–Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis thaliana from Polarization Transfer Solid-State NMR. *J. Am. Chem. Soc.* **2014,** *136* (29), 10399-10409.

5. Kirui, A.; Du, J.; Zhao, W.; Barnes, W.; Kang, X.; Anderson, C. T.; Xiao, C.; Wang, T., A pectin Methyltransferase modulates polysaccharide dynamics and interactions in *Arabidopsis* primary cell walls: Evidence from solid-state NMR. *Carbohydr. Polym.* **2021,** *270*, 118370.

6. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* **2020,** *11* (1), 6081.

7. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 1-9.

8. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: a critical review. *Solid State Nucl. Magn. Reson.* **2020,** *107*, 101660.

9. Sapkota, Y.; White, J. R., Long-term fate of rapidly eroding carbon stock soil profiles in coastal wetlands. *Sci. Total Environ.* **2021,** *753*, 141913.

10. Haywood, B. J.; Hayes, M. P.; White, J. R.; Cook, R. L., Potential fate of wetland soil carbon in a deltaic coastal wetland subjected to high relative sea level rise. *Sci. Total Environ.* **2020,** *711*, 135185.

11. Roth, V.-N.; Lange, M.; Simon, C.; Hertkorn, N.; Bucher, S.; Goodall, T.; Griffiths, R. I.; Mellado-Vazquez, P. G.; Mommer, L.; Oram, N. J.; Weigelt, A.; Dittmar, T.; Gleixner, G., Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nat. Geosci.* **2019,** *2019*, 755-761.

12. Osland, M. J.; Gabler, C. A.; Grace, J. B.; Day, R. H.; McCoy, M. L.; McLeod, J. L.; From, A. S.; Enwright, N. M.; Feher, L. C.; Stagg, C. L.; Hartley, S. B., Climate and Plant Controls on Soil Organic Matter in Coastal Wetlands. *Glob. Chang. Biol.* **2018,** *24*, 5361-5379.

13. Plaza, C.; Pegoraro, E.; Bracho, R.; Celis, G.; Crummer, K. G.; Hutchings, J. A.; Hicks Pries, C. E.; Mauritz, M.; Natali, S.; Salmon, V. G.; Schadel, C.; Webb, E. E.; Schuur, E. A.

G., Direct observation of permafrost degradation and rapid soil carbon loss in tundra. *Nat. Geosci.* **2019,** *12*, 627-631.

14. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

15. Ghassemi, N.; Poulhazan, A.; Deligey, F.; Mentink-Vigier, F.; Marcotte, I.; Wang, T., Solid-State NMR Investigations of Extracellular Matrices and Cell Walls of Algae, Bacteria, Fungi, and Plants. *Chem. Rev.* **2021**, in press.

16. Zhao, W.; Kirui, A.; Deligey, F.; Mentink-Vigier, F.; Zhou, Y.; Zhang, B.; Wang, T., Solid-state NMR of unlabeled plant cell walls: high-resolution structural analysis without isotopic enrichment. *Biotechnol. Biofuels* **2021,** *14* (1), 14.

17. Kirui, A.; Ling, Z.; Kang, X.; Widanage, M. C. D.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic resolution of cotton cellulose structure enabled by dynamic nuclear polarization solid-state NMR. *Cellulose* **2019,** *26* (1), 329-339.

18. Barna, J. C. J.; Laue, E. D.; Mayger, M. R.; Skilling, J.; Worrall, S. P., Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *J. Magn. Reson.* **1969,** *73* (1), 69-77.

19. Wind, R. A.; Duijvestijn, M. J.; van der Lugt, C.; Manenschijn, A.; Vriend, J., Applications of dynamic nuclear polarization in 13C NMR in solids. *Prog. Nucl. Mag. Res. Sp.* **1985,** *17*, 33-67.

20. Schmidt-Rohr, K.; Spiess, H. W., *Multidimensional Solid-State NMR and Polymers*. Academic Press: San Diego, 1994; p 13-68.

21. Hall, D. A.; Maus, D. C.; Gerfen, G. J.; Inati, S. J.; Becerra, L. R.; Dahlquist, F. W.; Griffin, R. G., Polarization-Enhanced NMR Spectroscopy of Biomolecules in Frozen Solution. *Science* **1997,** *276* (5314), 930-932.

22. Levitt, M., *Spin dynamics: Basics of Nuclear Magnetic Resonance, Second Edition*. Wiley-Blackwell: 2008; Vol. 37.

23. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-state NMR spectroscopy. *Nat. Rev. Methods Primers* **2021,** *1* (1), 2.

24. Duer, M. J., The contribution of solid-state NMR spectroscopy to understanding biomineralization: atomic and molecular structure of bone *J. Magn. Reson.* **2015,** *253*, 98-110.

25. Levitt, M. H.; Grant, D. M.; Harris, R. K., Symmetry-based pulse sequences in magic-angle spinning solid-state NMR. In *Solid State NMR Studies of Biopolymers*, McDermott, A. P., T., Eds, Ed. 2007; p 229.

26. Haw, J. F.; Maciel, G. E.; Schroeder, H. A., Carbon-13 nuclear magnetic resonance

spectrometric study of wood and wood pulping with cross polarization and magic-angle spinning. *Anal. Chem.* **1984,** *56* (8), 1323-1329.

27. Garbow, J. R.; Jacob, G. S.; Stejskal, E. O.; Schaefer, J., Protein dynamics from chemical shift and dipolar rotational spin-echo ^{15}N NMR. *Biochemistry* **1989,** *28* (3), 1362-7.

28. Gullion, T., Introduction to rotational-echo, double-resonance NMR. *Concepts Magn. Reson.* **1998,** *10* (5), 277-289.

29. Yamamoto, H.; Horii, F., CPMAS carbon-13 NMR analysis of the crystal transformation induced for Valonia cellulose by annealing at high temperatures. *Macromolecules* **1993,** *26* (6), 1313-1317.

30. Skjemstad, J.; Clarke, P.; Taylor, J.; Oades, J.; Newman, R., The Removal of Magnetic Materials from Surface Soils - A Solid State ^{13}C CP/MAS NMR Study. *Soil Res.* **1994,** *32*, 1215-1229.

31. Elena, B.; Lesage, A.; Steuernagel, S.; Böckmann, A.; Emsley, L., Proton to Carbon-13 INEPT in Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2005,** *127* (49), 17296-17302.

32. Duan, P.; Schmidt-Rohr, K., Composite-pulse and partially dipolar dephased multiCP for improved quantitative solid-state 13C NMR. *J. Magn. Reson.* **2017,** *285*, 68-78.

33. Chen, L.; Kaiser, J. M.; Lai, J.; Polenova, T.; Yang, J.; Rienstra, C. M.; Mueller, L. J., J-based 2D homonuclear and heteronuclear correlation in solid-state proteins. *Magn. Reson. Chem.* **2007,** *45* (S1), S84-S92.

34. Sengupta, I.; Nadaud, P. S.; Helmus, J. J.; Schwieters, C. D.; Jaroniec, C. P., Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nat. Chem.* **2012,** *4* (5), 410-417.

35. Bielecki, A.; Kolbert, A. C.; De Groot, H.; Griffin, R. G.; Levitt, M. H., Frequency-Switched Lee—Goldburg Sequences in Solids. *Adv. Magn. Reson.* **1990,** *14*, 111-124.

36. Aluas, M.; Tripon, C.; Griffin, J. M.; Filip, X.; Ladizhansky, V.; Griffin, R. G.; Brown, S. P.; Filip, C., CHHC and ^{1}H–^{1}H magnetization exchange: Analysis by experimental solid-state NMR and 11-spin density-matrix simulations. *J. Magn. Reson.* **2009,** *199* (2), 173-187.

37. Cadars, S.; Sein, J.; Duma, L.; Lesage, A.; Pham, T. N.; Baltisberger, J. H.; Brown, S. P.; Emsley, L., The refocused INADEQUATE MAS NMR experiment in multiple spin-systems: Interpreting observed correlation peaks and optimising lineshapes. *J. Magn. Reson.* **2007,** *188* (1), 24-34.

38. Kono, H.; Numata, Y., Structural investigation of cellulose I α and I β by 2D RFDR NMR spectroscopy: determination of sequence of magnetically inequivalent D-glucose units along cellulose chain. *Cellulose* **2006,** *13* (3), 317-326.

39. Gullion, T.; Schaefer, J., Rotational-echo double-resonance NMR. *J. Magn. Reson.* **1989,** *81*

(1), 196-200.

40. Hou, G.; Yan, S.; Sun, S.; Han, Y.; Byeon, I.-J. L.; Ahn, J.; Concel, J.; Samoson, A.; Gronenborn, A. M.; Polenova, T., Spin Diffusion Driven by R-Symmetry Sequences: Applications to Homonuclear Correlation Spectroscopy in MAS NMR of Biological and Organic Solids. *J. Am. Chem. Soc.* **2011,** *133* (11), 3943-3953.

41. Hou, G.; Yan, S.; Trébosc, J.; Amoureux, J.-P.; Polenova, T., Broadband homonuclear correlation spectroscopy driven by combined R2nv sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J. Magn. Reson.* **2013,** *232*, 18-30.

42. Wang, T.; Williams, J. K.; Schmidt-Rohr, K.; Hong, M., Relaxation-compensated difference spin diffusion NMR for detecting 13C–13C long-range correlations in proteins and polysaccharides. *J. Biomol. NMR* **2015,** *61* (2), 97-107.

43. Li, S.; Zhang, Y.; Hong, M., 3D ^{13}C-^{13}C-^{13}C Correlation NMR for De Novo Distance Determination of Solid Proteins and Application to a Human Alpha Defensin. *J. Magn. Reson.* **2010,** *202*, 203-210.

44. Shekar, S. C.; Zhao, W.; Fernando, L. D.; Hung, I.; Wang, T., A ^{13}C three-dimensional DQ-SQ-SQ correlation experiment for high-resolution analysis of complex carbohydrates using solid-state NMR. *J. Magn. Reson.* **2022,** *336*, 107148.

45. Gelenter, M. D.; Wang, T.; Liao, S.-Y.; O'Neill, H.; Hong, M., ^{2}H–^{13}C correlation solid-state NMR for investigating dynamics and water accessibilities of proteins and carbohydrates. *J. Biomol. NMR* **2017,** *68* (4), 257-270.

46. Nadaud, P. S.; Helmus, J. J.; Kall, S. L.; Jaroniec, C. P., Paramagnetic ions enable tuning of nuclear relaxation rates and provide long-range structural restraints in solid-state NMR of proteins. *J. Am. Chem. Soc.* **2009,** *131* (23), 8108-8120.

47. Clore, M. G.; Iwahara, J., Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **2009,** *109* (9), 4108-4139.

48. Torchia, D. A., Measurement of proton-enhanced ^{13}C T_1 values by a method which suppresses artifacts. *J. Magn. Reson.* **1978,** *30*, 613-616.

49. Zhang, R.; Mroue, K. H.; Ramamoorthy, A., Proton-Based Ultrafast Magic Angle Spinning Solid-State NMR Spectroscopy. *Acc. Chem. Res.* **2017,** *50* (4), 1105-1113.

50. Phyo, P.; Hong, M., Fast MAS ^{1}H–^{13}C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73* (12), 661-674.

51. Roos, M.; Wang, T.; Shcherbakov, A. A.; Hong, M., Fast Magic-Angle-Spinning ^{19}F Spin Exchange NMR for Determining Nanometer ^{19}F–^{19}F Distances in Proteins and Pharmaceutical Compounds. *J. Phys. Chem. B* **2018,** *122* (11), 2900-2911.

52. Ishii, Y.; Tycko, R., Sensitivity Enhancement in Solid State ^{15}N NMR by Indirect Detection with High-Speed Magic Angle Spinning. *J. Magn. Reson.* **2000,** *142* (1), 199-204.

53. Oster, C.; Kosol, S.; Hartlmuller, C.; Lamley, J. M.; Luga, D.; Oss, A.; Org, M.-L.; Vanatalu, K.; Samoson, A.; Madl, T.; Lewandowski, J. R., Characterization of Protein–Protein Interfaces in Large Complexes by Solid-State NMR Solvent Paramagnetic Relaxation Enhancements. *J. Am. Chem. Soc.* **2017,** *139* (35), 12165-12174.

54. Zuckerstätter, G.; Terinte, N.; Sixta, H.; Schuster, K. C., Novel insight into cellulose supramolecular structure through ^{13}C CP-MAS NMR spectroscopy and paramagnetic relaxation enhancement. *Carbohydr. Polym.* **2013,** *93* (1), 122-128.

55. Kaur, M.; Lewis, C. M.; Chronister, A.; Phun, G. S.; Mueller, L. J., Non-Uniform Sampling in NMR Spectroscopy and the Preservation of Spectral Knowledge in the Time and Frequency Domains. *J. Phys. Chem. A* **2020,** *124* (26), 5474-5486.

56. Märker, K.; Pingret, M.; Mouesca, J.-M.; Gasparutto, D.; Hediger, S.; De Paëpe, G., A New Tool for NMR Crystallography: Complete ^{13}C/^{15}N Assignment of Organic Molecules at Natural Isotopic Abundance Using DNP-Enhanced Solid-State NMR. *J. Am. Chem. Soc.* **2015,** *137* (43), 13796-13799.

57. Rankin, A. G. M.; Trébosc, J.; Pourpoint, F.; Amoureux, J.-P.; Lafon, O., Recent developments in MAS DNP-NMR of materials. *Solid State Nucl. Magn. Reson.* **2019,** *101*, 116-143.

58. Mentink-Vigier, F.; Akbey, Ü.; Oschkinat, H.; Vega, S.; Feintuch, A., Theoretical aspects of Magic Angle Spinning - Dynamic Nuclear Polarization. *J. Magn. Reson.* **2015,** *258*, 102-120.

59. Mentink-Vigier, F.; Paul, S.; Lee, D.; Feintuch, A.; Hediger, S.; Vega, S.; De Paëpe, G., Nuclear depolarization and absolute sensitivity in magic-angle spinning cross effect dynamic nuclear polarization. *Phys. Chem. Chem. Phys.* **2015,** *17* (34), 21824-21836.

60. Stevanato, G.; Casano, G.; Kubicki, D. J.; Rao, Y.; Esteban Hofer, L.; Menzildjian, G.; Karoui, H.; Siri, D.; Cordova, M.; Yulikov, M.; Jeschke, G.; Lelli, M.; Lesage, A.; Ouari, O.; Emsley, L., Open and Closed Radicals: Local Geometry around Unpaired Electrons Governs Magic-Angle Spinning Dynamic Nuclear Polarization Performance. *J. Am. Chem. Soc.* **2020,** *142* (39), 16587-16599.

61. Mentink-Vigier, F.; Dubroca, T.; Van Tol, J.; Sigurdsson, S. T., The distance between g-tensors of nitroxide biradicals governs MAS-DNP performance: The case of the bTurea family. *J. Magn. Reson.* **2021,** *329*, 107026.

62. Mentink-Vigier, F., Optimizing nitroxide biradicals for cross-effect MAS-DNP: the role of g-tensors' distance. *Phys. Chem. Chem. Phys.* **2020,** *22* (6), 3643-3652.

63. Perras, F. A.; Pruski, M., Large-scale ab initio simulations of MAS DNP enhancements using a Monte Carlo optimization strategy. *J. Chem. Phys.* **2018,** *149* (15), 154202.

64. Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem. Int. Ed.* **2013,** *52* (41), 10858-10861.

65. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

66. Chaudhari, S. R.; Wisser, D.; Pinon, A. C.; Berruyer, P.; Gajan, D.; Tordo, P.; Ouari, O.; Reiter, C.; Engelke, F.; Copéret, C.; Lelli, M.; Lesage, A.; Emsley, L., Dynamic Nuclear Polarization Efficiency Increased by Very Fast Magic Angle Spinning. *J. Am. Chem. Soc.* **2017,** *139* (31), 10609-10612.

67. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis* primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiol.* **2015,** *168* (3), 871-884.

68. Wang, T.; Hong, M., Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **2016,** *67* (2), 503-514.

69. Gao, Y.; Mortimer, J. C., Unlocking the architecture of native plant cell walls via solid-state nuclear magnetic resonance. *Methods Cell Biol.* **2020,** *160*, 121-143.

70. Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Wang, P.; Kang, X.; Wang, T., Preparation of Fungal and Plant Materials for Structural Elucidation Using Dynamic Nuclear Polarization Solid-State NMR. *J. Vis. Exp.* **2019,** (144), e59152.

CHAPTER 2: SOLID-STATE NMR OF PLANT AND FUNGAL CELL WALLS: A CRITICAL REVIEW

Review paper reprinted with permission: *Solid State Nucl. Magn. Reson.* 107, 101660 (2020)

Authors: Wancheng Zhao, Liyanage D. Fernando, Alex Kirui, Fabien Deligey, and Tuo Wang

2.1 Abstract

The cell walls of plants and microbes are a central source for bio-renewable energy and the major targets of antibiotics and antifungal agents. It is highly challenging to determine the molecular structure of complex carbohydrates, protein and lignin, and their supramolecular assembly in intact cell walls. This article selectively highlights the recent breakthroughs that employ $^{13}C/^{15}N$ solid-state NMR techniques to elucidate the architecture of fungal cell walls in *Aspergillus fumigatus* and the primary and secondary cell walls in a large variety of plant species such as *Arabidopsis*, *Brachypodium*, maize, and spruce. Built upon these pioneering studies, we further summarize the underexplored aspects of fungal and plant cell walls. The new research opportunities introduced by innovative methods, such as the detection of proton and quadrupolar nuclei on ultrahigh-field magnets and under fast magic-angle spinning, paramagnetic probes, natural-abundance DNP, and software development, are also critically discussed.

2.2 Cell Walls: a Medically Important and Energy-Relevant Biomaterial

The cell wall is a carbohydrate-rich coating outside the plasma membrane of plants and many microorganisms. The cell walls in photosynthesis systems, such as plants, algae, and green bacteria, are transformed from solar energies and carbon dioxide, with enormous value as a primary source of building materials, textiles, biofuel, nanocomposites, and high-value reagents[1]. Polysaccharides in the cell walls of fungal pathogens and invading bacteria are absent in human cells; these components trigger immune recognition and serve as the major target of antifungal

drugs and antibiotics[2, 3]. Polysaccharides and other biomolecules (such as protein and lignin) are held together by covalent linkages and physical packing interactions to form a mechanically strong composite, which allows the cell to retain integrity and morphology under external stress. Nevertheless, the numerous interactions between biopolymers also pose a challenge for post-harvest processing and utilization of biomass and make it technically difficult to characterize these biomaterials with high resolution.

Recently, magic-angle-spinning (MAS) solid-state NMR spectroscopy has been extensively employed to investigate intact cell walls. Uniformly isotope-labeled samples are produced by feeding the organism of interest with $^{13}CO_2$ or solid/liquid media containing ^{13}C-glucose and ^{15}N-salts[4]. Multidimensional ^{13}C-^{13}C/^{15}N correlation spectra collected on whole cells or isolated cell walls provide the atomic resolution needed for determining the polymorphic structure, intermolecular interaction, water contact, and molecular motions of biomolecules in their cellular environment (**Figure 2.1**). Within the last decade, a large variety of biosystems have been studied: the primary and secondary cell walls of seven plant species, including *Arabidopsis thaliana, Brachypodium distachyon* and *Zea mays* (maize), rice, switchgrass, poplar and spruce[5-14]; the biofilm or cell walls of bacterial and fungal pathogens such as *Aspergillus fumigatus*, *Cryptococcus neoformans*, and *Bacillus subtilis*[15-21]; as well as the carbohydrate components in microalgae *Chlamydomonas reinhardti*[22]. Here we will review the major findings related to plant and fungal cell walls, emphasize the key questions awaiting investigation, and discuss the future directions enabled by the improved instrumentation and methodology, in the attempt to inspire innovative research in carbohydrate and cell wall NMR.

2.3 Recent Advances in Cell Wall Research by Solid-State NMR

2.3.1 Molecular Insight of Plant Primary Cell Walls

Since 2010, Hong and colleagues have been pioneering the investigations of primary plant cell walls, a component synthesized in the growing plants (**Figure 2.2a-2.2c**)[23,24]. The composition is well known from numerous biochemical studies, and three major types of polysaccharides are present[25, 26]. Cellulose microfibrils are formed by 18 or more glucan chains (3-4 nm across) and they are highly rigid and partially crystalline. The backbones of pectin, such as rhamnogalacturonan-I (RG-I) and homogalacturonan (HG), are often acidic and responsible for regulating cell wall hydration. Hemicellulose interacts with cellulose and pectin, with a plant-dependent composition: the major hemicellulose is xyloglucan (XyG) in dicots, such as *Arabidopsis*, but changes to glucuronoarabinoxylan (GAX) and mixed-linkage glucans (MLG) in commelinid monocots (grasses) such as *Brachypodium* and maize. Using ^{13}C-labeled and isolated cell walls, three ground-breaking discoveries were reported, which have revised and substantiated our limited understanding of primary cell wall architecture.

First, cellulose, hemicellulose, and pectin are found to associate noncovalently on the sub-nanometer scale to form an integrated network. In wild-type *Arabidopsis* (**Figure 2.2a**, Sample 1), a large number of cross peaks have been identified between pectin and cellulose, which were previously considered to be phase-separated [5]. The equilibrium intensity of ^{13}C-^{13}C spin diffusion suggests that 25-50% of cellulose are in close proximity to pectin[6, 27]. This polymer interaction is independent of the sample's hydration history[28] and can be fully preserved after partial depectination by CDTA and sodium carbonate (**Figure 2.2a**, Sample 2), which disrupts the calcium crosslinking of HG and consequently removing the interfibrillar HG molecules (40% of all HG) that are not binding cellulose[28-30]. Due to the loss of immobilized water in the depectinated

sample, the rate of ^{1}H-^{1}H polarization transfer from water to polymers have been globally slowed down for all polysaccharides, which can be partially restored by the subsequent digestions of XyG using xyloglucanase and Cel12A enzymes due to the enhanced surface areas of the residual macromolecules (**Figure 2.2a**, Sample 3)[29]. In addition, the removal of XyG using an *xxt1xxt2xxt5* triple knockout line (**Figure 2.2a**, Sample 4) markedly enhances the dynamics of the remaining polysaccharides [5], which echoes with the global alternation of ^{1}H spin diffusion observed in the sequentially digested samples, revealing a single network of all polysaccharides.

It is noteworthy that a weaker pectin-cellulose interaction is often accompanied by the chemical modification of pectin structure, for example, a higher degree of methyl esterification, an increased occurrence of sidechain branching by arabinan or galactan, a reduced extent of calcium-crosslinking, and promoted HG aggregation. These molecular changes macroscopically correlate with faster growth, for example, in the inflorescence stem of *Arabidopsis* with a segmentally increasing rate of elongation from the base to the apical region (the tip)[31], in the *PGX1*^{AT} mutant that produces smaller pectin but larger plants (**Figure 2.2a**, Sample 5)[32], and in a low-pH sample that mimics the acid growth condition (**Figure 2.2a**, Sample 6)[33].

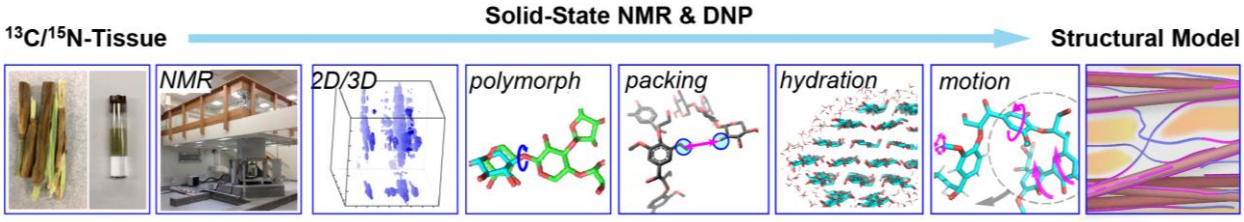

Figure 2.1 Solid-state NMR strategy for investigating cell wall materials.

Second, with the assistance from Dynamic Nuclear Polarization (DNP) and paramagnetic methods, two methods have been developed to reveal how a class of proteins (expansin) unfasten the polysaccharide joints to mediate cell expansion[34, 35]. Expansins lack the lytic activity expected

for wall-loosening enzyme and have been assumed to disrupt the non-covalent contacts between polysaccharides[36]. Solid-state NMR studies have shown that expansins precisely perturb the cellulose-xyloglucan nexus in *Arabidopsis* but disrupt the junctions between the highly and lowly substituted GAX in maize (**Figure 2.2c**); therefore, expansins bind different polysaccharides in the cell walls with distinct composition.

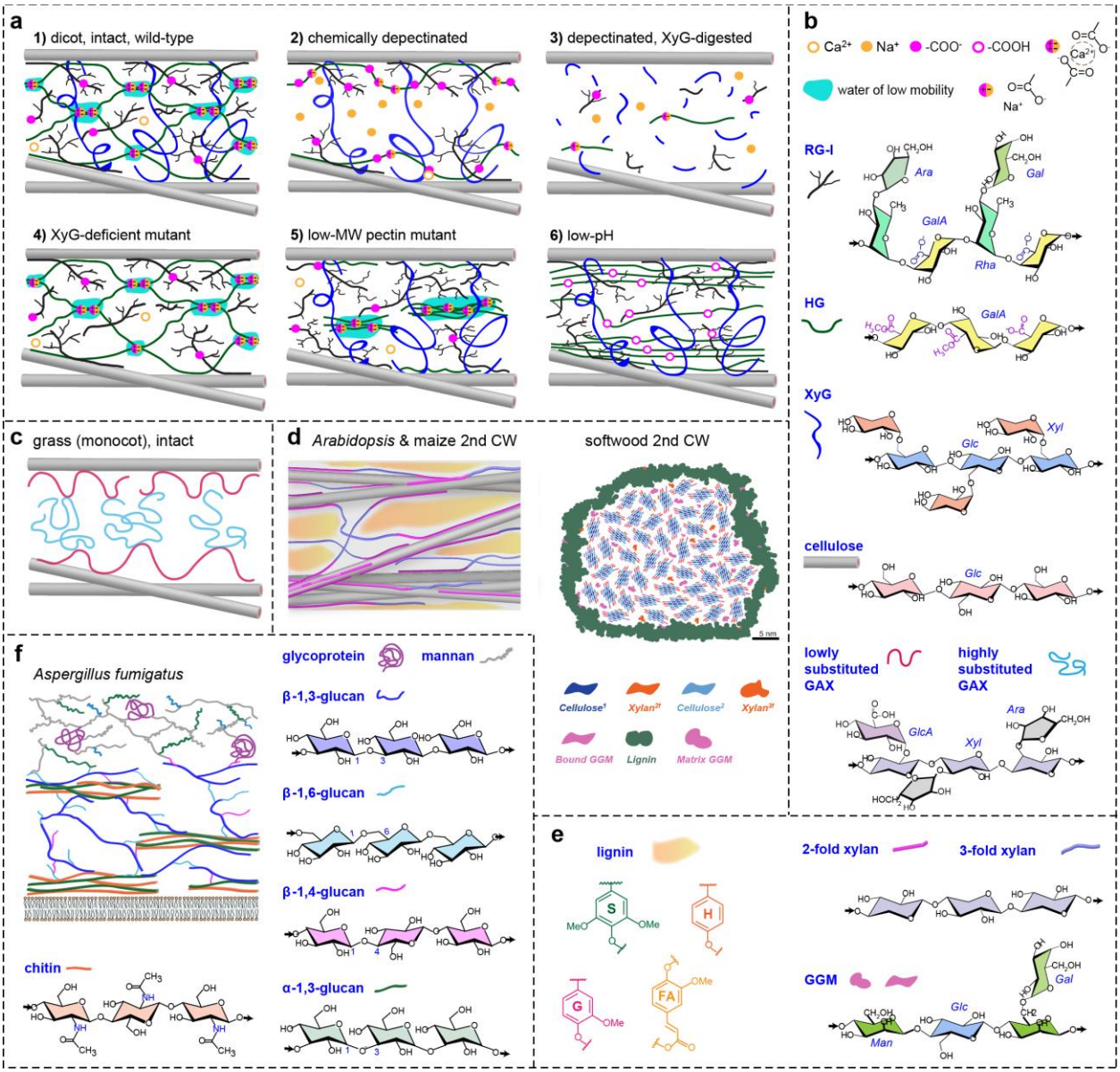

Figure 2.2 Cell walls and biomolecules studied by solid-state NMR. **a,** NMR-derived conceptual

Figure 2.2 (cont'd)

models of primary cell walls of *Arabidopsis* (dicot), including the intact and wild-type cell walls at different pH values, as well as multiple mutants that attenuate the structure and content of matrix polysaccharides (pectin or XyG), sequentially digested cell walls that are chemically depectinated followed by enzymatical removal of XyG. The sample numbers are labeled to facilitate discussion. **b,** Chemical structures of the biomolecular components in primary plant cell walls. **c,** Structural scheme of primary grass cell walls based on data collected using *Brachypodium* and maize. **d,** Plant secondary (2nd) cell wall architecture of *Arabidopsis*, maize, and softwood spruce. **e,** Representative structure of polymers that are unique to secondary plant cell walls. f, the structure and polymers of fungal cell walls in *A. fumigatus*.

Third, with the sharp ^{13}C linewidths on high magnetic fields (0.7-1.0 ppm for cellulose on an 800 MHz NMR) and the chemical shift calculations using Density Functional Theory (DFT), we have resolved seven types of glucose units in the cellulose of *Arabidopsis* and grass primary cell walls, determined their hydroxymethyl conformations via ^{1}H-^{1}H distance measurement, and localized these conformers in the microfibrils[37-39]. These forms deviate noticeably from the crystallographic structures of Iα and Iβ allomorphs obtained using the highly crystalline cellulose from bacteria and tunicates (a marine animal). In addition, these seven types of glucose residues have been consistently observed in the secondary cell walls of *Arabidopsis*, maize, switchgrass, and rice[8], as well as multiple woody plants such as *Eucalyptus*, poplar, and spruce (unpublished results). Therefore, the Iα and Iβ model allomorphs are generally absent in most natural resources. So far, the NMR signals of Iα and Iβ structures have only been observed in cotton, thus a large crystallite is a prerequisite for accommodating the model structures[40].

2.3.2 Lignin-Carbohydrate Packing in Plant Secondary Cell Walls

The secondary cell wall is formed once the cell ceases expansion and it comprises the majority of lignocellulosic biomass. In secondary cell walls, cellulose microfibrils aggregate into larger bundles (10–20 nm across), which are further embedded in a matrix containing the aromatic polymer lignin and hemicellulose such as xylan and glucomannan [41]. Lignin-carbohydrate interactions confer the biomass with recalcitrance to chemical and enzymatical treatments; therefore, it is of broad interest to understand the chemical principles underlying these polymer interactions. Paul Dupree, Ray Dupree, and colleagues have conducted several studies to recognize the functional relevance of xylan polymorphism in *Arabidopsis* secondary cell walls. It is found that only the xylan with a 2-fold helical screw symmetry and a regular pattern of acetate or glucuronate substitutions can bind cellulose microfibrils[10-12].

Stimulated by these discoveries, we have investigated the mature stems of maize, rice, switchgrass, and *Arabidopsis*, using a series of 2D ^{13}C-^{13}C correlation methods specially designed for enhancing the aromatic signals of lignin and detecting the lignin-carbohydrate interface (**Figure 2.3a**)[8]. Hundreds (234) of intermolecular cross-peaks have been identified, which pinpoint six categories of packing interactions between the different functional groups in lignin and carbohydrates as illustrated in **Figure 2.3b-2.3g**. Strikingly, lignin mainly interacts with xylan rather than cellulose. In addition, the number and intensities of these cross peaks statistically correlate with the number of methyl ether substitutions in lignin residues (**Figure 2.3b**), which signposts a prevalent role of electrostatic contacts in stabilizing polymer interface. Integrating the information on polymer packing, dynamics, and hydration has resulted in a molecular view of lignocellulosic materials: lignin self-aggregates to form dynamically unique and hydrophobic nanodomains, with surface contact to the non-flat xylan (3-fold) through abundant electrostatic

interactions[8]. This xylan-lignin interface links to the flat-ribbon domain of xylan that is coating the surface of cellulose microfibrils (**Figure 2.2d**, left).

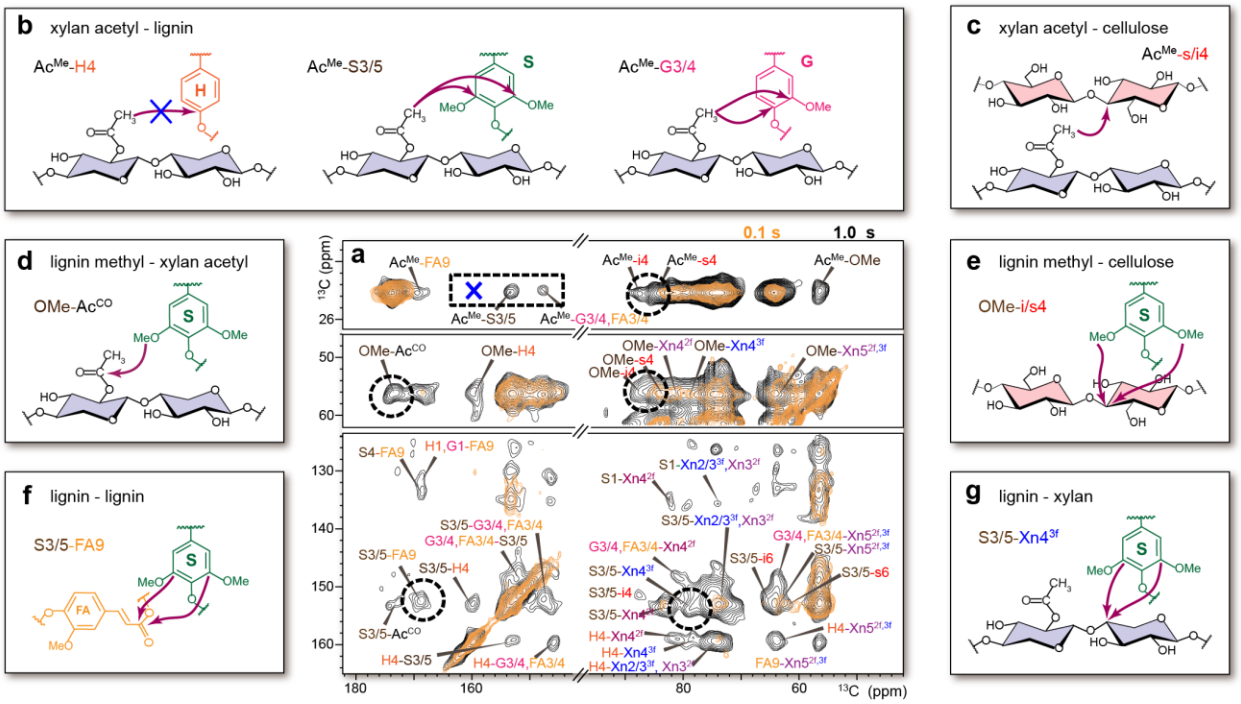

Figure 2.3 Intermolecular interactions in plant secondary cell walls by dipolar-filtered PDSD spectra. **a,** Overlay of 2D ^{13}C-^{13}C correlation spectra measured with a short (0.1 s, yellow) and long (1.0 s, black) mixing time. Six representative regions of intermolecular cross peaks for the structural illustration in panels b-g are highlighted in dashline circles or rectangle. These interactions happen between. **b,** xylan acetyl (Ac) group and lignin aromatics (S, G, or H), **c,** xylan acetyl group and cellulose microfibrils (i: internal glucan chains; s: surface chains), **d,** lignin methyl group (OMe: methyl ether) and xylan acetyl groups, **e,** lignin methyl group and cellulose, **f,** the aromatic carbons of different lignin units, as well as g, lignin aromatics and xylan furanose ring. The blue crosses in panels a and b highlight the missing signal of Ac^{Me}-H4 cross peak between xylan and the H-residue of lignin.

In the softwood spruce, xylan also binds cellulose through its 2-fold conformer while

galactoglucomannan (GGM), a unique hemicellulose in softwoods, binds the surface of cellulose microfibrils in a semi-crystalline manner[13]. Since both GGM and xylan have shown two domains, one coating cellulose and the other filling interfibrillar space, it is proposed that some GGM and xylan bind to the same microfibril and further associate with lignin (**Figure 2.2d**, right).

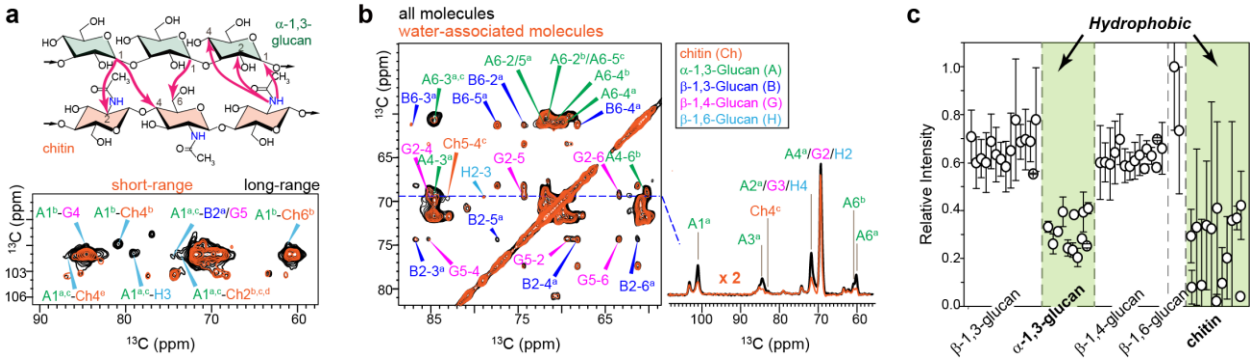

Figure 2.4 Polymer packing and hydration in fungal cell walls of *A. fumigatus*. **a,** Top: Illustration of the observed cross peaks between α-1,3-glucan (A) and chitin (Ch). Bottom: overlay of 2D ^{13}C-^{13}C correlation spectra measured with 100 ms DARR (orange, short-range) and 15 ms PAR (black, long-range). **b,** Overlay of the control (black) and water-edited spectra (orange); the cross section extracted at 69 ppm shows apparent signal dephasing of chitin and α-1,3-glucan in water-edited spectra. **c,** Site-specific hydration map of biopolymers.

2.3.3 Insights into the Fungal Cell Wall Architecture

Recently, we have also initiated a project elucidating the cell wall structure of an airborne fungal pathogen *A. fumigatus*. The samples are measured alive at room-temperature on an 800 MHz NMR; the ^{13}C linewidths are as narrow as 0.4-0.6 ppm for the rigid polysaccharides and 0.3-0.5 ppm for the relatively mobile carbohydrates and proteins. With the remarkable resolution, the ^{13}C and ^{15}N signals of 7 types of polysaccharides, including α-1,3-glucan, chitin (a nitrogenated polysaccharide), mannan, and three types of β-glucans, together with their 23 conformers, have

been identified (**Figure 2.2f**)[15]. Long-range correlation methods, such as ^{13}C-^{13}C and ^{15}N-^{15}N Proton-Assisted Recoupling (PAR)[42] as well as NCACX measured with a variable ^{13}C-^{13}C mixing time, have been employed to identify in total 65 intermolecular interactions. Most of these physical interactions occur between chitin and α-1,3-glucans (**Figure 2.4a**), which also show high hydrophobicity (**Figure 2.4b, c**) and rigidity. These two molecules are complexed to form a mechanical scaffold that is surrounded by a soft matrix of diversely linked β-glucans and capped by an external shell rich in glycoproteins (**Figure 2.2f**). This study has established a preliminary structural frame, which requires systematic validation and encourages structural investigations of individual cell wall molecules and their biomedical relevance.

2.4 Biochemical Perspectives: the Unresolved Questions

In an earlier Trends article published in 2016, several underexplored areas in plant NMR have been summarized, which mainly include the coalescence of multiple cellulose microfibrils, the functional structure of lignin, and the putative interactions between polysaccharides and structural proteins[24]. So far, we have already obtained an in-depth understanding of lignin-carbohydrate interactions in secondary cell walls but also discovered more aspects that remain ambiguous. For instance, the sophisticated patterns of covalent linkages between lignin residues and their impact on lignin's capability of interacting with polysaccharides have not been discussed. Cellulose-lignin interaction is scarce in maize and *Arabidopsis*, but this concept may not hold for the woody plants with a distinct composition of biomolecules and a more compact packing. The conformational relevance of glucomannan in spruce and other softwood species is not yet understood. We also need to understand the structural origin of the abundant electrostatic interactions between carbohydrates and aromatics and how these physical contacts contribute to the mechanical properties and digestibility of lignocellulosic materials. Inevitably, we also need

to figure out a way to integrate solid-state NMR results with the numerous studies using solution-NMR, which are focused on the covalent linkages in extracted residuals.

An unexpected finding in the *A. fumigatus* fungus is the multifaced role of α-1,3-glucans. These molecules are simultaneously in association with chitin for stiffness and existing in the mobile phase[15]. This observation has countered the biochemical results in which α-1,3-glucans are extractable by strong alkali and thus constantly excluded from the structural core of any prevailing models[43]. Also, the amount of this polysaccharide is much lower in many other pathogens such as most yeasts (for example, *Candida albicans*); therefore, it is of great interest to understand the structural and dynamical heterogeneity of polysaccharides across fungal species. Another major polysaccharide, mannan, is found to coexist with proteins in the mobile domain of *A. fumigatus* cell walls, likely constituting the mannoproteins in the outermost layer as depicted in biochemical studies. Further evidence assessing the structural role of mannan and their covalent linkages with structural proteins are crucial to the understanding of this structurally dynamic shell that regulates cellular recognition and fungal pathogenesis[44]. It is also important to understand how the microbe re-structures its cell wall in response to antifungal agents, which will explain the origin of drug resistance from a structural perspective.

2.5 Technical Outlook: Opportunities Beyond Conventional Methods

The past decade has seen the rapid development of solid-state NMR techniques. Here, we have selectively summarized a few technical advances that could potentially revolutionize the field and establish new research directions. These highlights have extended beyond the conventional $^{13}C/^{15}N$-methods by involving other NMR-active nuclei or electrons, with assistance from ultra-fast MAS, ultrahigh magnetic field, DNP, as well as database and software coding.

2.5.1 ^{1}H and ^{19}F under Fast Spinning: Carbohydrate Structure and Interactions

Direct detection of proton resonance provides high-sensitivity due to the high isotope abundance (99.985%) and four-fold higher gyromagnetic ratio over ^{13}C. The strong homonuclear couplings enable distance measurement beyond 1 nm and facilitate structural determination. Although proton detection has already been widely employed to study perdeuterated or protonated proteins[45-48], it is rarely applied to the carbohydrates that are rich in protons with complex chemical environments. Recently, Hong and Phyo have conducted a set of proton-detection experiments, such as the 2D CH INEPT, 3D CHH INEPT-TOCSY, and 2D hChH RFDR techniques[49-51], to assign the ^{1}H resonances of polysaccharides and to determine their intermolecular packing, for example, through cross peaks between cellulose carbons and matrix polysaccharide protons, in ^{13}C-labeled *Arabidopsis* primary cell walls[52]. The protonated material has been back exchanged in D_2O, which suppresses the water intensities and reduces the contribution of hydroxyl protons. A moderately fast MAS frequency (30-50 kHz) is chosen to simultaneously enable proton-detection of the mobile matrix and filter out the signals of rigid microfibrils[52]. The narrow ^{1}H linewidth (0.06 ppm on an 800 MHz spectrometer, **Figure 2.5a**) and the excellent agreement between solid-state ^{1}H chemical shifts and solution-NMR observables consistently confirmed that the observed matrix polysaccharides are intrinsically mobile in cell walls. In addition, Simorre, Schanda, and coworkers have assigned the resonances of peptidoglycan in intact *Bacillus subtilis* under 100 kHz MAS, with representative ^{1}H linewidths of 50-120 Hz (0.05-0.13 ppm) on a 950 MHz spectrometer[53, 54]. Using the ^{1}H-^{1}H RFDR scheme, the authors have identified multiple inter-residue cross peaks, including unambiguous cross peaks between the GlcNAc sugar and the L-alanine residue on the peptide stem, crossing a long distance of at least 5 Å[53]. These studies have presented a novel strategy for investigating complex biosystems and landed the stage for pursuing

^{1}H investigations without labeling.

Similarly, ^{19}F has a high gyromagnetic ratio and 100% natural abundance. Adding to these merits is a large range of chemical shifts for resolving various chemical motifs. As demonstrated on pharmaceutical compounds, GB1 protein, and HIV-1 capsid protein, ^{19}F-^{19}F/^{1}H distances can be measured on the nanometer scale (1-2 nm)[55-57], which is a major extension from the reach of ^{13}C and ^{15}N methods. Typically, site-specific fluorination causes minimal perturbation to the structures of many proteins and materials[56, 58], but may substantially disrupt the hydrogen bonds in carbohydrate polymers. An appropriate labeling scheme is needed to sparsely fluorinate carbohydrates without eliminating their functional structures and assembly[59].

2.5.2 ^{17}O at Ultrahigh-Field: a New Biochemical Probe

Oxygen is another core element that determines the hydrogen bonding and chemical properties of biomolecules. Carbohydrates are particularly rich in oxygen atoms, with at least one oxygen covalently linked to each single carbon site. Recently, the materialization of a world-record 1.5 GHz (35 Tesla) series-connected hybrid (SCH) NMR magnet[60] and the commercial ultrahigh field instruments have presented a unique opportunity for high-resolution ^{17}O studies. Griffin and colleagues have revealed the markedly improved resolution of ^{17}O spectra on ultrahigh fields where the line-broadening by second-order quadrupole coupling is attenuated[61]. They have also collected 2D ^{13}C/^{15}N/^{1}H-^{17}O correlation spectra and determined internuclear distances through recoupling methods such as ZF-TEDOR and REAPDOR[61, 62]. Back in 2007, Grandinetti and coworkers have already pioneered the measurement of ^{17}O MAS patterns for monosaccharides and disaccharides that are site-specifically labeled at either the hydroxyl or glycosidic oxygen sites (**Figure 2.5b**) [63]. The C–O–H angle and C–O distances, instead of the O–H distances, are found to affect ^{17}O quadrupolar couplings in carbohydrates. In addition, many other quadrupolar nuclei

may benefit from the availability of ultrahigh-field magnets. For example, ^{33}S NMR could help characterize many sulfurated carbohydrates (such as the ulvan, carrageenan, and rhamnan sulfate) in marine species[64]. Another popular molecule is heparin, a sulfated glycosaminoglycan that prevents blood clotting as an anticoagulant agent and induces filament assembly of tau proteins [65]. Combining quadrupolar NMR with ultrahigh field magnets provides a novel probe to the biochemically important sites in these carbohydrates.

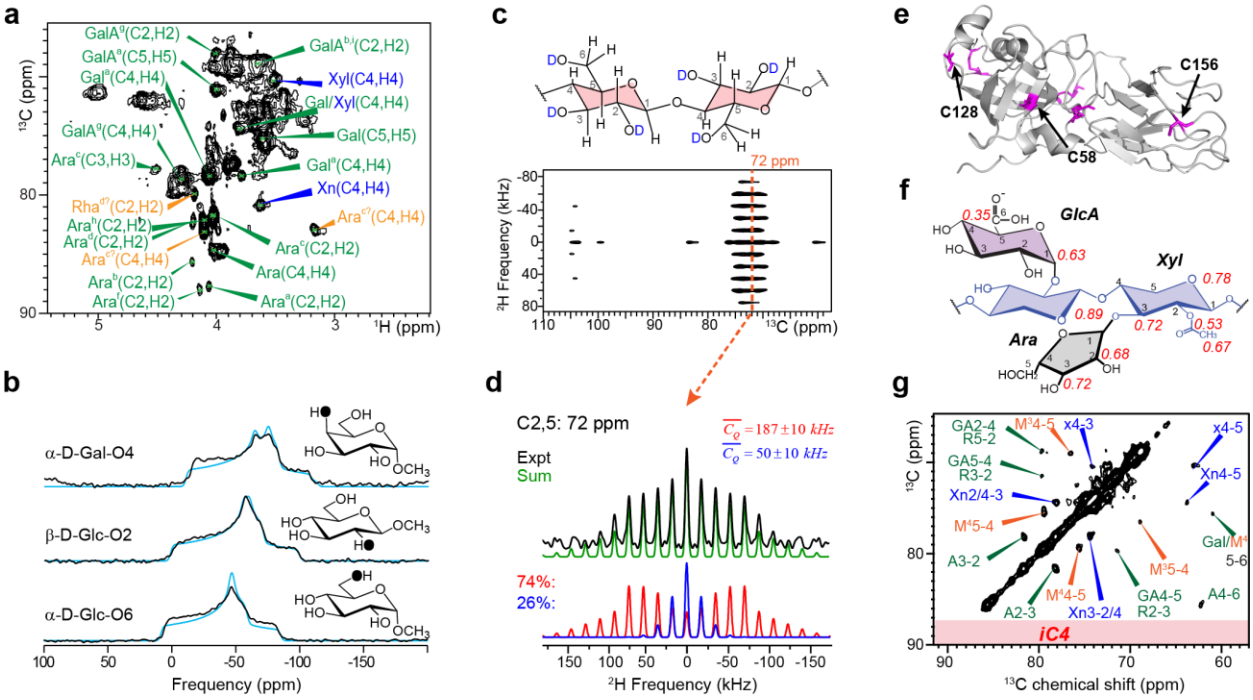

Figure 2.5 Advanced techniques that promote carbohydrate and cell wall research. **a,** 2D CH INEPT spectrum of ^{13}C labeled *Arabidopsis* cell walls measured at 50kHz MAS on an 800 MHz NMR. Superscripts denote the different subtypes of monosaccharides and ambiguous assignments are shown in orange. **b,** Experimental (black) and simulated (blue) ^{17}O MAS spectra at 19.6 T of methyl α-D-galactopyranoside (4-^{17}O), methyl β-D-glucopyranoside (2-^{17}O), and methyl α-D-glucopyranoside (6-^{17}O) at 12 kHz MAS. **c,** 2D ^{2}H-^{13}C correlation spectrum of H/D exchanged and ^{13}C-labeled *Arabidopsis* cell wall. The structure of cellulose and ^{2}H-exchanged OD sites are given. **d,** The 72-ppm ^{2}H cross-sections (black) and best-fit simulation (green) that is a

Figure 2.5 (cont'd)

supercomposition of components with C_Q values of 187 and 50 kHz with weighting factors of 74 and 26%, respectively. **e,** The structure of β-expansin (EXPB1) with 9 Cys residues shown in magenta. The 3 solvent-accessible cysteines that could be tagged with EDTA-Mn(II) are annotated. f, The PRE effect of Mn(II)-tagged expansin on glucuronoarabinoxylan (GAX). The intensities ratios of 2 spectra measured on the Mn-containing sample and the control sample are shown in red: GAX has significant signal dephasing due to binding to the protein. g, Difference PDSD spectrum of two parent spectra measured on Mn-containing or Mn-free samples: showing no cellulose intensities (iC4) but only the signals of matrix polysaccharides that bind proteins.

2.5.3 ^{2}H: Dynamics and Water Accessibility

In cell walls, carbohydrate dynamics were primarily evaluated by measuring NMR relaxation and dipolar couplings[66], and water-polymer contacts were mainly investigated using 1D/2D ^{13}C-detected, ^{1}H spin diffusion methods and dipolar-filtered heteronuclear 2D correlation techniques like MELODI-HETCOR[28, 29, 67] Recently, Hong and coworkers have employed the Rotor Echo Short Pulse IRrAdiaTION mediated cross-polarization ($^{\text{RESPIRATION}}$CP) technique[68, 69] to achieve multi-bond, broadband ^{2}H–^{13}C polarization using an affordable ^{2}H radiofrequency field of ~50 kHz and a short contact time below 1.7 ms[70]. A rapid trans-gauche isomerization is identified in perdeuterated bacterial cellulose. This hydroxymethyl motion around the C5–C6 bond is absent in the interior glucan chains of cellulose but occurs to the surface chains as revealed by their motionally averaged C6-^{2}H quadrupolar couplings. In H/D exchanged *Arabidopsis* cell walls, 2D ^{13}C-^{2}H correlation spectra (**Figure 2.5c**) have shown a mixed quadrupolar pattern that can be best deconvoluted into two components: the quadrupolar coupling constant is 50 kHz for the mobile matrix polysaccharides and 187 kHz for the rigid cellulose, which is a value

approaching the hydrogen-bonded rigid deuteroxyl quadrupolar coupling (**Figure 2.5d**). This robust method can be applied to evaluate the dynamics and water-accessible surface of carbohydrates in various organisms.

2.5.4 PRE: Carbohydrate-Ligand Binding in Cellular Environment

In structural biology, carbohydrates are often treated as small ligands attached to large protein complexes, but this concept has been inverted in cell wall studies. In plant cell walls, functional proteins are usually present at low concentrations, but with the capabilities of perturbing polymer nexuses or chemically modifying/digesting certain structural motifs. For the β-expansins that cannot be produced recombinantly, the extracted proteins from grass pollens are tagged with paramagnetic Mn(II) labels via their solvent-accessible Cys residues (**Figure 2.5e**), and mixed with the ^{13}C-labeled cell walls in maize[35]. Upon binding to expansin, the hemicellulose glucuronoarabinoxylan (GAX) has shown strong ^{1}H and ^{13}C Paramagnetic Relaxation Enhancement (PRE) effects (**Figure 2.5f, g**), and its stiff and mobile fractions have become more rigid and dynamic, respectively. Therefore, β-expansins have released the connections between the highly substituted GAX (mobile) that forms the interfibrillar matrix and the rarely branched GAX (rigid) that are packed with cellulose microfibrils. The optimized protocols for incorporating paramagnetic sites and the PRE-enabled distance determination[71-74] have made it feasible for revealing the interactions between carbohydrates and many proteins or enzymes that contain carbohydrate-binding modules[75].

2.5.5 Natural-Abundance DNP and Database: Accommodate the Growing Field

As an emerging technique, natural-abundance MAS-DNP has enabled the measurement of 2D ^{13}C–^{13}C/^{15}N/^{1}H spectra on unlabeled biomolecules (**Figure 2.6a**). When applied to organic molecules and small peptides, this technique could substantially facilitate NMR crystallography

by enabling the determination of ^{13}N-^{13}C distance up to 7 Å and the measurement of ^{15}N-^{13}C correlation at natural isotope abundance[76-78]. Applications of this method to complicated biosystems allow us to extract long-range distance constraints in polyglutamine (polyQ) amyloid fibrils and nano-assemblies of cyclic peptides[77, 79], validate imino acid-aromatic interactions in native collagens[80], and identify the compositional and conformational differences of cellulose, hemicellulose, and lignin in various plant species (cotton, rice, and poplar)[9, 40, 81]. These studies were conducted on medium magnetic fields, the 400 MHz/263 GHz or 600 MHz/395 GHz DNP instruments; due to the limited resolution, only highly ordered systems, such as cellulose microfibrils and amyloid fibrils, or a selected component within whole-cell sample could be studied. Because high-field DNP is still inefficient at this stage but has become a necessity for providing sufficient resolution for studying complex samples, the efforts trying to improve the polarization mechanism and radicals at high fields could substantially strengthen the capability of natural-abundance DNP[82-84].

A rate-limiting process associated with natural-abundance DNP is to interpret the large number of NMR-observables into structural information. We have recently demonstrated that a heatmap comparing the chemical shifts measured on the cotton cellulose and reported in literature allows us to quickly identify the relevant structures (**Figure 2.6b**)[40]. This application benefits from the implementation of Complex Carbohydrates Magnetic Resonance Database (CCMRD) that supports the storage and sharing of information on chemical shifts, dynamics, and structure. CCMRD is freely available to the public at www.ccmrd.org and supports data deposition and data search by NMR chemical shifts, carbohydrate name, and compound class (**Figure 2.6c, d**)[85]. By the time of this article, 450 compounds from plants, fungi, bacteria, algae, and engineered biomaterials are indexed by CCMRD, and this platform will accommodate the rapid expansion of

the dataset and facilitate the development of statistics-based software[40]. My vision for carbohydrate ssNMR is to enable high-throughput and semi-automatic analysis of spectra and structure, which requires dedicated efforts in method and software development.

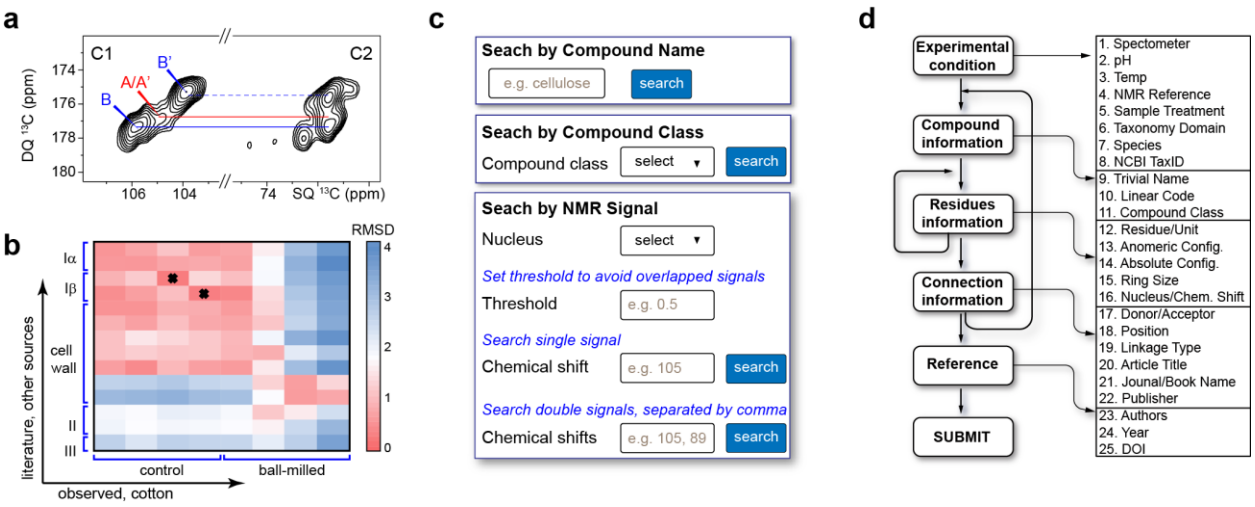

Figure 2.6 Natural-abundance DNP of unlabeled material assisted by database development. **a,** Natural abundance 2D ^{13}C-^{13}C INADEQUATE spectrum of unlabeled cotton. A and A': glucose units in Iα cellulose allomorph; B and B': glucose units in Iβ allomorph. **b,** ^{13}C chemical shift RMSD map for comparisons between cotton and other cellulose sources. The color scale of RMSD (ppm) is shown. **c,** Search interface of CCMRD database that supports data search by compound name, class, and signal. d, Flowchart of data deposition and the 25 types of entries included for each compound.

2.6 Conclusion

Solid-state NMR and DNP have demonstrated their unique capability in understanding the nanoscale assembly of fungal and plant cell walls. The rapid advances in NMR instrumentation and technology have made it possible to address biochemical and structural questions that were previously impossible to answer. The studies of plant and fungal cell walls, combined with the

many investigations of other complex biosystems, such as the bacterial cell walls and biofilm, algal polysaccharides, and mammalian carbohydrates, have formed an emerging and unique research direction, which is of high significance to the development of biorenewable energy, biomedical therapies, and high-value products based on carbohydrate polymers.

2.7 Acknowledgments

This work was supported by the National Science Foundation (CAREER award MCB-1942665) and the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under award number DE-SC0001090.

REFERENCES

1. Somerville, C.; Youngs, H.; Taylor, C.; Davis, S. C.; Long, S. P., Feedstocks for Lignocellulosic Biofuels. *Science* **2010,** *329* (5993), 790-792.

2. Odds, F. C.; Brown, A. J. P.; Gow, N. A. R., Antifungal agents: mechanisms of action. *Trends Microbiol.* **2003,** *11* (6), 272-279.

3. Romaniuk, J. A.; Cegelski, L., Bacterial cell wall composition and the influence of antibiotics by cell-wall and whole-cell NMR. *Phil. Trans. R. Soc. B* **2015,** *370* (1679).

4. Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Wang, P.; Kang, X.; Wang, T., Preparation of Fungal and Plant Materials for Structural Elucidation Using Dynamic Nuclear Polarization Solid-State NMR. *J. Vis. Exp.* **2019,** (144), e59152.

5. Dick-Perez, M.; Zhang, Y. A.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry-Us* **2011,** *50* (6), 989-1000.

6. Wang, T.; Zabotina, O.; Hong, M., Pectin-cellulose interactions in the *Arabidopsis* primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry-Us* **2012,** *51* (49), 9846-9856.

7. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and dynamics of *Brachypodium* primary cell wall polysaccharides from two-dimensional ^{13}C solid-state nuclear magnetic resonance spectroscopy. *Biochemistry-Us* **2014,** *53* (17), 2840-2854.

8. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 347.

9. Perras, F. A.; Luo, H.; Zhang, X.; Mosier, N. S.; Pruski, M.; Abu-Omar, M. M., Atomic-Level Structure Characterization of Biomass Pre- and Post-Lignin Treatment by Dynamic Nuclear Polarization-Enhanced Solid-State NMR. *J. Phys. Chem. A* **2017,** *121* (3), 623-630.

10. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

11. Grantham, N. J.; Wurman-Rodrich, J.; Terrett, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **2017,** *3* (11), 859-865.

12. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry-Us* **2015,** *54* (14), 2335-2345.

13. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 4978.

14. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Barder, M.; Luterbacher, J., Topology of Pretreated Wood Fibers Using Dynamic Nuclear Polarization. *J. Phys. Chem. C* **2019**, in press.

15. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

16. Chatterjee, S.; Prados-Rosales, R.; Frases, S.; Itin, B.; Casadevall, A.; Stark, R. E., Using Solid-State NMR To Monitor the Molecular Consequences of Cryptococcus neoformans Melanization with Different Catecholamine Precursors. *Biochemistry-Us* **2012,** *51* (31), 6080-6088.

17. Chatterjee, S.; Prados-Rosales, R.; Itin, B.; Casadevall, A.; Stark, R. E., Solid-state NMR Reveals the Carbon-based Molecular Architecture of *Cryptococcus neoformans* Fungal Eumelanins in the Cell Wall. *J. Biol. Chem.* **2015,** *290* (22), 13779-13790.

18. Nygaard, R.; Romaniuk, J. A. H.; Rice, D. M.; Cegelski, L., Spectral snapshots of bacterial cell-wall composition and the influence of antibiotics by whole-cell NMR. *Biophys. J.* **2015,** *108* (6), 1380-1389.

19. Takahashi, H.; Ayala, I.; Bardet, M.; De Paëpe, G.; Simorre, J. P.; Hediger, S., Solid-state NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. *J. Am. Chem. Soc.* **2013,** *135* (13), 5105-10.

20. Kern, T.; Giffard, M.; Hediger, S.; Amoroso, A.; Giustini, C.; Bui, N. K.; Joris, B.; Bougault, C.; Vollmer, W.; Simorre, J. P., Dynamics characterization of fully hydrated bacterial cell walls by solid-state NMR: evidence for cooperative binding of metal ions. *J. Am. Chem. Soc.* **2010,** *132* (31), 10911-9.

21. Thongsomboon, W.; Serra, D. O.; Possling, A.; Hadjineophytou, C.; Hengge, R.; Cegelski, L., Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. *Science* **2018,** *359* (6373), 334-338.

22. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. *J. Biomol. NMR* **2018,** *70* (2), 123-131.

23. Wang, T.; Hong, M., Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **2016,** *67*, 503-514.

24. Wang, T.; Phyo, P.; Hong, M., Multidimensional solid-state NMR spectroscopy of plant cell walls. *Solid State Nucl. Magn. Reson.* **2016,** *78*, 56-63.

25. Carpita, N. C., Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Phys.* **1996,** *47*, 445-476.

26. Carpita, N. C.; Gibeaut, D. M., Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *The Plant journal : for cell and molecular biology* **1993,** *3* (1), 1-30.

27. Wang, T.; Williams, J. K.; Schmidt-Rohr, K.; Hong, M., Relaxation-compensated difference spin diffusion NMR for detecting 13C-13C long-range correlations in proteins and polysaccharides. *J. Biomol. NMR* **2015,** *61* (2), 97-107.

28. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis thaliana Primary Cell Walls: Evidence from Solid-State NMR. *Plant Physiol.* **2015,** *168* (3), 871-884.

29. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water-polysaccharide interactions in the primary cell wall of *Arabidopsis thaliana* from polarization transfer solid-state NMR. *J. Am. Chem. Soc.* **2014,** *136* (29), 10399-10409.

30. Dick-Perez, M.; Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Multidimensional solid-state NMR studies of the structure and dynamics of pectic polysaccharides in uniformly ^{13}C-labeled Arabidopsis primary cell walls. *Magn. Reson. Chem.* **2012,** *50* (8), 539-550.

31. Phyo, P.; Wang, T.; Kiemle, S. N.; O'Neill, H.; Pingali, S. V.; Hong, M.; Cosgrove, D. J., Gradients in Wall Mechanics and Polysaccharides along Growing Inflorescence Stems. *Plant Physiol.* **2017,** *175* (4), 1593-1607.

32. Phyo, P.; Wang, T.; Xiao, C. W.; Anderson, C. T.; Hong, M., Effects of Pectin Molecular Weight Changes on the Structure, Dynamics, and Polysaccharide Interactions of Primary Cell Walls of Arabidopsis thaliana: Insights from Solid-State NMR. *Biomacromolecules* **2017,** *18* (9), 2937-2950.

33. Phyo, P.; Gu, Y.; Hong, M., Impact of acidic pH on plant cell wall polysaccharide structure and dynamics: insights into the mechanism of acid growth in plants from solid-state NMR. *Cellulose* **2019,** *26* (1), 291-304.

34. Wang, T.; Park, Y. B.; Caporini, M. A.; Rosay, M.; Zhong, L. H.; Cosgrove, D. J.; Hong, M., Sensitivity-enhanced solid-state NMR detection of expansin's target in plant cell walls. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (41), 16444-16449.

35. Wang, T.; Chen, Y. N.; Tabuchi, A.; Cosgrove, D. J.; Hong, M., The Target of beta-Expansin EXPB1 in Maize Cell Walls from Binding and Solid-State NMR Studies. *Plant Physiol.* **2016,** *172* (4), 2107-2119.

36. Cosgrove, D. J., Loosening of plant cell walls by expansins. *Nature* **2000,** *407* (6802), 321-326.

37. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant

Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

38. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

39. Yang, H.; Wang, T.; Oehme, D.; Petridis, L.; Hong, M.; Kubicki, J. D., Structural factors affecting ^{13}C NMR chemical shifts of cellulose: a computational study. *Cellulose* **2018,** *25* (1), 23-36.

40. Kirui, A.; Ling, Z.; Kang, X.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic Resolution of Cotton Cellulose Structure Enabled by Dynamic Nuclear Polarization Solid-State NMR. *Cellulose* **2019,** *26*, 329-339.

41. Cosgrove, D. J.; Jarvis, M. C., Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant. Sci.* **2012,** *3*.

42. De Paepe, G.; Lewandowski, J. R.; Loquet, A.; Bockmann, A.; Griffin, R. G., Proton assisted recoupling and protein structure determination. *The Journal of chemical physics* **2008,** *129* (24).

43. Latge, J. P., The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.* **2007,** *66* (2), 279-290.

44. Stahl, P. D.; Ezekowitz, R. A. B., The mannose receptor is a pattern recognition receptor involved in host defense. *Curr. Opin. Immunol.* **1998,** *10* (1), 50-55.

45. Andreas, L. B.; Jaudzems, K.; Stanek, J.; Lalli, D.; Bertarello, A.; Le Marchand, T.; Cala-De Paepe, D.; Kotelovica, S.; Akopjana, I.; Knott, B.; Wegner, S.; Engelke, F.; Lesage, A.; Emsley, L.; Tars, K.; Herrmann, T.; Pintacuda, G., Structure of fully protonated proteins by proton-detected magic-angle spinning NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113* (33), 9187-92.

46. Bertini, I.; Emsley, L.; Lelli, M.; Luchinat, C.; Mao, J. F.; Pintacuda, G., Ultrafast MAS Solid-State NMR Permits Extensive ^{13}C and ^{1}H Detection in Paramagnetic Metalloproteins. *J. Am. Chem. Soc.* **2010,** *132* (16), 5558-5559.

47. Zhou, D. H.; Shah, G.; Cormos, M.; Mullen, C.; Sandoz, D.; Rienstra, C. M., Proton-detected solid-state NMR Spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J. Am. Chem. Soc.* **2007,** *129* (38), 11791-11801.

48. Struppe, J.; Quinn, C. M.; Lu, M. M.; Wang, M. Z.; Hou, G. J.; Lu, X. Y.; Kraus, J.; Andreas, L. B.; Stanek, J.; Lalli, D.; Lesage, A.; Pintacuda, G.; Maas, W.; Gronenborn, A. M.; Polenova, T., Expanding the horizons for structural analysis of fully protonated protein assemblies by NMR spectroscopy at MAS frequencies above 100 kHz. *Solid State Nucl. Magn. Reson.* **2017,** *87*, 117-125.

49. Elena, B.; Lesage, A.; Steuernagel, S.; Bockmann, A.; Emsley, L., Proton to carbon-13

INEPT in solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2005,** *127* (49), 17296-17302.

50. Agarwal, V.; Reif, B., Residual methyl protonation in perdeuterated proteins for multi-dimensional correlation experiments in MAS solid-state NMR spectroscopy. *J Magn Reson* **2008,** *194* (1), 16-24.

51. Baldus, M.; Meier, B. H., Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J. Magn. Reson. A* **1996,** *121* (1), 65-69.

52. Phyo, P.; Hong, M., Fast MAS ^{1}H-^{13}C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73* (12), 661-674.

53. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206* (1), 66-72.

54. Schanda, P.; Triboulet, S.; Laguri, C.; Bougault, C. M.; Ayala, I.; Callon, M.; Arthur, M.; Simorre, J. P., Atomic Model of a Cell-Wall Cross-Linking Enzyme in Complex with an Intact Bacterial Peptidoglycan. *J. Am. Chem. Soc.* **2014,** *136* (51), 17852-17860.

55. Roos, M.; Mandala, V. S.; Hong, M., Determination of Long-Range Distances by Fast Magic-Angle-Spinning Radiofrequency-Driven F-19-F-19 Dipolar Recoupling NMR. *The journal of physical chemistry. B* **2018,** *122* (40), 9302-9313.

56. Shcherbakov, A. A.; Mandala, V. S.; Hong, M., High-Sensitivity Detection of Nanometer 1H-19F Distances for Protein Structure Determination by H-1-Detected Fast MAS NMR. *J. Phys. Chem. B.* **2019,** *123* (20), 4387-4391.

57. Wang, M.; Lu, M.; Fritz, M. P.; Quinn, C. M.; Byeon, I. L.; Byeon, C. H.; Struppe, J.; Maas, W.; Gronenborn, A. M.; Polenova, T., Fast Magic-Angle Spinning 19F NMR Spectroscopy of HIV-1 Capsid Protein Assemblies. *Angew. Chem. Int. Ed. Engl.* **2018,** *57* (50), 16375-16379.

58. Fry, R. A.; Tsomaia, N.; Pantano, C. G.; Mueller, K. T., ^{19}F MAS NMR quantification of accessible hydroxyl sites on fiberglass surfaces. *J. Am. Chem. Soc.* **2003,** *125* (9), 2378-2379.

59. Meier, B. H., Polarization Transfer and Spin Diffusion in Solid-State NMR. Adv. Magn. Opt. Reson. , 1994.

60. Gan, Z. H.; Hung, I.; Wang, X. L.; Paulino, J.; Wu, G.; Litvak, I. M.; Gor'kov, P. L.; Brey, W. W.; Lendi, P.; Schiano, J. L.; Bird, M. D.; Dixon, L. R.; Toth, J.; Boebinger, G. S.; Cross, T. A., NMR spectroscopy up to 35.2 T using a series-connected hybrid magnet. *J Magn Reson* **2017,** *284*, 125-136.

61. Keeler, E. G.; Michaelis, V. K.; Colvin, M. T.; Hung, I.; Gor'kov, P. L.; Cross, T. A.; Gan, Z. H.; Griffin, R. G., ^{17}O MAS NMR Correlation Spectroscopy at High Magnetic Fields. *J. Am. Chem. Soc.* **2017,** *139* (49), 17953-17963.

62. Gullion, T., Measurement of Dipolar Interactions between Spin-1/2 and Quadrupolar Nuclei

by Rotational-Echo, Adiabatic-Passage, Double-Resonance Nmr. *Chem Phys Lett* **1995,** *246* (3), 325-330.

63. Sefzik, T. H.; Houseknecht, J. B.; Clark, T. M.; Prasad, S.; Lowary, T. L.; Gan, Z.; Grandinetti, P. J., Solid-state 17O NMR in carbohydrates. *Chem. Phys. Lett.* **2007,** *434* (4), 312-315.

64. Jiao, G. L.; Yu, G. L.; Zhang, J. Z.; Ewart, H. S., Chemical Structures and Bioactivities of Sulfated Polysaccharides from Marine Algae. *Mar. Drugs* **2011,** *9* (2), 196-223.

65. Zhang, W. J.; Falcon, B.; Murzin, A. G.; Fan, J.; Crowther, R. A.; Goedert, M.; Scheres, S. H. W., Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **2019,** *8*.

66. Wang, T.; Hong, M., *Structure and Dynamics of Polysaccharidesin Plant Cell Walls from Solid-State NMR*. Royal Society of Chemistry: 2017.

67. Yao, X. L.; Schmidt-Rohr, K.; Hong, M., Medium- and long-distance ^{1}H-^{13}C heteronuclear correlation NMR in solids. *J Magn Reson* **2001,** *149* (1), 139-143.

68. Jain, S.; Bjerring, M.; Nielsen, N. C., Efficient and Robust Heteronuclear Cross-Polarization for High-Speed-Spinning Biological Solid-State NMR Spectroscopy. *The journal of physical chemistry letters* **2012,** *3* (6), 703-8.

69. Wei, D.; Akbey, U.; Paaske, B.; Oschkinat, H.; Reif, B.; Bjerring, M.; Nielsen, N. C., Optimal ^{2}H rf Pulses and ^{2}H-^{13}C Cross-Polarization Methods for Solid-State ^{2}H MAS NMR of Perdeuterated Proteins. *J Phys Chem Lett* **2011,** *2* (11), 1289-94.

70. Gelenter, M. D.; Wang, T.; Liao, S. Y.; O'Neill, H.; Hong, M., ^{2}H-^{13}C correlation solid-state NMR for investigating dynamics and water accessibilities of proteins and carbohydrates. *J. Biol. NMR* **2017,** *68* (4), 257-270.

71. Jaroniec, C. P., Structural studies of proteins by paramagnetic solid-state NMR spectroscopy. *J Magn Reson* **2015,** *253*, 50-9.

72. Sengupta, I.; Nadaud, P. S.; Helmus, J. J.; Schwieters, C. D.; Jaroniec, C. P., Protein fold determined by paramagnetic magic-angle spinning solid-state NMR spectroscopy. *Nat. Chem.* **2012,** *4* (5), 410-7.

73. Iwahara, J.; Clore, G. M., Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **2006,** *440* (7088), 1227-30.

74. Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R., NMR spectroscopy of paramagnetic metalloproteins. *Chembiochem* **2005,** *6* (9), 1536-1549.

75. Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J., Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **2004,** *382*, 769-781.

76. Marker, K.; Pingret, M.; Mouesca, J. M.; Gasparutto, D.; Hediger, S.; De Paepe, G., A New Tool for NMR Crystallography: Complete $^{13}C/^{15}N$ Assignment of Organic Molecules at Natural Isotopic Abundance Using DNP-Enhanced Solid-State NMR. .*J Am. Chem. Soc.* **2015,** *137* (43), 13796-13799.

77. Marker, K.; Paul, S.; Fernandez-de-Alba, C.; Lee, D.; Mouesca, J. M.; Hediger, S.; De Paepe, G., Welcoming natural isotopic abundance in solid-state NMR: probing pi-stacking and supramolecular structure of organic nanoassemblies using DNP. *Chem. Sci.* **2017,** *8* (2), 974-987.

78. Olsen, R. A.; Struppe, J.; Elliott, D. W.; Thomas, R. J.; Mueller, L. J., Through-bond ^{13}C-^{13}C correlation at the natural abundance level: Refining dynamic regions in the crystal structure of vitamin-D-3 with solid-state NMR. *J. Am. Chem. Soc.* **2003,** *125* (39), 11784-11785.

79. Smith, A. N.; Marker, K.; Piretra, T.; Boatz, J. C.; Matlahov, I.; Kodali, R.; Hediger, S.; van der Wel, P. C. A.; De Paepe, G., Structural Fingerprinting of Protein Aggregates by Dynamic Nuclear Polarization-Enhanced Solid-State NMR at Natural Isotopic Abundance. *J. Am. Chem. Soc.* **2018,** *140* (44), 14576-14580.

80. Singh, C.; Rai, R. K.; Aussenac, F.; Sinha, N., Direct Evidence of Imino Acid-Aromatic Interactions in Native Collagen Protein by DNP-Enhanced Solid-State NMR Spectroscopy. *The journal of physical chemistry letters* **2014,** *5* (22), 4044-4048.

81. Zhang, L.; Gao, C.; Mentink-Vigier, F.; Tang, L.; Zhang, D.; Wang, S.; Cao, S.; Xu, Z.; Liu, X.; Wang, T.; Zhou, Y.; Zhang, B., Arabinosyl Deacetylase Modulates the Arabinoxylan Acetylation Profile and Secondary Wall Formation. *Plant Cell* **2019**.

82. Mathies, G.; Caporini, M. A.; Michaelis, V. K.; Liu, Y. P.; Hu, K. N.; Mance, D.; Zweier, J. L.; Rosay, M.; Baldus, M.; Griffin, R. G., Efficient Dynamic Nuclear Polarization at 800 MHz/527 GHz with Trityl-Nitroxide Biradicals. *Angew. Chem. Int. Edit.* **2015,** *54* (40), 11770-11774.

83. Jaudzems, K.; Bertarello, A.; Chaudhari, S. R.; Pica, A.; Cala-De Paepe, D.; Barbet-Massin, E.; Pell, A. J.; Akopjana, I.; Kotelovica, S.; Gajan, D.; Ouari, O.; Tars, K.; Pintacuda, G.; Lesage, A., Dynamic Nuclear Polarization-Enhanced Biomolecular NMR Spectroscopy at High Magnetic Field with Fast Magic-Angle Spinning. *Angew. Chem. Int. Edit.* **2018,** *57* (25), 7458-7462.

84. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paepe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

85. Kang, X.; Zhao, W.; Dickwella Widanage, M. C.; Kirui, A.; Ozdenvar, U.; Wang, T., CCMRD: A Solid-State NMR Database for Complex Carbohydrates. *J. Biomol. NMR* **2019,** *accepted*.

CHAPTER 3: CURRENT LIMITATIONS OF SOLID-STATE NMR IN CARBOHYDRATE AND CELL WALL RESEARCH

Review paper reprinted with permission: *J. Magn. Reson.* 341, 107263 (2022)

Authors: Wancheng Zhao, Fabien Deligey, S. Chandra Shekar,

Frederic Mentink-Vigier, and Tuo Wang

3.1 Abstract

High-resolution investigation of cell wall materials has emerged as an important application of biomolecular solid-state NMR (ssNMR). Multidimensional correlation experiments have become a standard method for obtaining sufficient spectral resolution to determine the polymorphic structure of carbohydrates and address biochemical questions regarding the supramolecular organization of cell walls. Using plant cellulose and matrix polysaccharides as examples, we will review how the multifaceted complexity of polysaccharide structure is impeding the resonance assignment process and assess the available biochemical and spectroscopic approaches that could circumvent this barrier. We will emphasize the ineffectiveness of the current methods in reconciling the ever-growing dataset and deriving structural information. We will evaluate the protocols for achieving efficient and homogeneous hyperpolarization across the cell wall material using magic-angle spinning dynamic nuclear polarization (MAS-DNP). Critical questions regarding the line-broadening effects of cell wall molecules at cryogenic temperature and by paramagnetic biradicals will be considered. Finally, the MAS-DNP method will be placed into a broader context with other structural characterization techniques, such as cryo-electron microscopy, to advance ssNMR research in carbohydrate and cell wall biomaterials.

3.2 Introduction

Carbohydrates and glycoconjugates regulate cellular recognition and communication, serve as energy sources, and provide structural support to the cells of many organisms[1]. These biologically important functions are often driven by the structural characteristics of carbohydrate polymers, such as glycosidic linkages and conformations. However, complex carbohydrates are difficult to characterize in their native physical state and have remained under-investigated, unlike proteins and nucleic acids. Solid-state NMR (ssNMR) has a long history (approximately 50 years) of being employed to characterize cellulose-based materials, such as fibers, derivatives, and lignocellulosic biomass[2-4]. Early studies typically relied on the extensive use of 1D ^{13}C spectra to resolve different allomorphs, quantify their contents, and determine the crystallinity of cellulosic materials. Another important application, as summarized in a recent review[5], is to understand the supramolecular organization of cell walls and biofilms in many organisms across biological kingdoms. In these studies, multidimensional (2D and 3D) correlation approaches are often required to achieve sufficient resolution for investigating cellular samples[5-8]. Because ssNMR can address questions of polymer structure and packing on the angstrom-to-nanometer length scale, the results complement the information provided by many diffraction and imaging methods that focus on the nanoscale and mesoscale.

The high-resolution studies of cell wall materials were built on a toolbox that combines ssNMR methods commonly used in both structural biology and polymer research[9]. Unlike for purified biomolecules, the complexity and heterogeneity of native biomaterials prevent the construction of atomistic models or of an ensemble of interchanging structures[10]. As shown by multiple studies of fungal and plant cell walls[11-13], only the model-based physical and structural principles that regulate the assembly and function of biomolecules can be obtained, similar to the

outcomes expected for polymer ssNMR. However, the ever-growing dataset is evolving striking resemblance to that used in NMR structural biology. For example, a recent study has compared the lignin-carbohydrate packing interfaces in grasses, hardwoods, and softwoods, proposing structural models based on more than a thousand site-specific data of polymer interactions and dynamics[11].

The goal of this review is to highlight the underexplored directions where method development is urgently needed. We will critically review the intrinsic structural complexity underlying the difficulty in analyzing carbohydrate polymers and recognize the immediate need for calculation methods for converting under-used structural constraints into atomistic models. We will then selectively address questions regarding the applications of magic-angle spinning dynamic nuclear polarization (MAS-DNP) in cell wall research. Sections are also devoted to discussing the potential opportunities opening up for each of the current limitations.

3.3 Overview of Experimental Approach

Typically, uniformly $^{13}C/^{15}N$-labeled cell wall materials or whole cells are measured using 2D and 3D correlation experiments to examine biopolymer structure and cell wall assembly (**Figure 3.1**)[5]. The most useful experiments for resonance assignment are *J*- or dipolar-based refocused INADEQUATE experiments, which unambiguously display the carbon connectivity in each type of sugar units. Cellular carbohydrates usually come with structural polymorphism to a high extent; a phenomenon that can be closely traced via peak multiplicity in 2D/3D ^{13}C-$^{13}C/^{15}N$ spectra. From the peak intensities, which are represented as the peak volumes or areas, the composition of the cell wall molecules can be quantified. It is always beneficial to cross-compare these values with other analytical methods such as chromatography and mass spectrometry (MS). To examine the sub-nanometer physical packing, ssNMR offers a large collection of long-range

correlation experiments, including but not limited to PDSD, CORD, CHHC, and PAR[14-16]. Molecular dynamics are examined using relaxation, dipolar couplings, and exchange NMR, depending on the timescale of the dynamics of interest. Spectral-editing techniques are effective for selecting molecules with unique structural motifs (for example, aromatics and nitrogen sites), rigidity or mobility, and water association, against the bulk of the cell wall[17-19].

Depending on the growth condition, plants can be grown in solution media containing ^{13}C-glucose or chambers containing $^{13}CO_2$[20]. The former is usually limited to young seedlings grown in dark to avoid isotope dilution by ^{12}C, while the latter has a broader range of applications regarding the plant species and developmental stages. Well-established protocols are also available for labeling other organisms such as fungi, bacteria, and algae. Sample handling should be carefully designed to minimize structural perturbation, notably by avoiding chemical treatment and harsh physical processing techniques such as ball-milling. Typically, a razor blade is used to cut the tissues into millimetric pieces to allow for an even distribution of weight in the rotor during magic-angle spinning (MAS). The native hydration status is another crucial consideration in biomolecular ssNMR. Because the hydroxyl groups of carbohydrates efficiently associate with water molecules, it is not surprising that highly similar spectral features were found in both never-dried cell walls and in samples submitted a dehydration-rehydration cycle. This has been observed in the primary cell walls of *Arabidopsis*, the secondary cell walls of spruce and eucalyptus, and microalgae named *Parachlorella*[11, 21, 22]. However, moderate irreversible alternations induced by dehydration have been detected in sorghum and pine. The changes mainly happen to the hemicellulose (xylan and mannan) located on the packing interface with other polymers (lignin and cellulose)[23, 24]. Native and fresh cells or tissues are recommended whenever available.

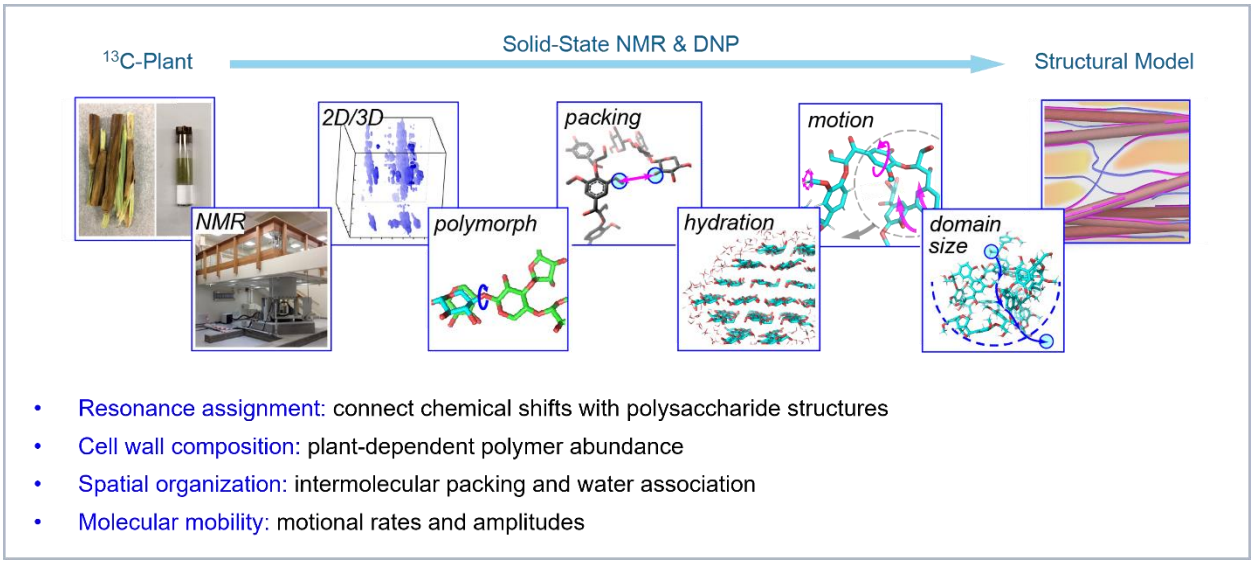

Figure 3.1 Schematic representation of ssNMR studies of cell wall materials. Plants are taken as an example, and only cell wall polysaccharides and lignin are shown here for illustration. Some panels of the figure are adapted from reference[25].

3.4 Resonance Assignment Impeded by Structural Complexity

Like protein structural determination by ssNMR, analysis of a ^{13}C-labeled cell wall sample begins with resonance assignment, which also turned out to be the most challenging step. The cell wall is typically the most rigid component and can be selected against intracellular molecules using experimental schemes based on cross-polarization (CP). However, the spectrum is still overcrowded due to the coexistence of a large variety of carbohydrates in the cell wall. The structural polymorphism of carbohydrate polymer further leads to peak multiplicity. Over the past decade, we have progressively uncovered the link between the structural and spectroscopic characteristics, which are summarized in **Figure 3.2** and detailed below.

First, non-cellulosic polysaccharides have highly diverse patterns of covalent linkages. When a linkage occurs to a carbon site, its ^{13}C isotropic chemical shift becomes larger by several ppm (also called downfield shift). In solids, this principle was demonstrated by systematically

tracking the signals of matrix polysaccharides in the model grass *Brachypodium distachyon*[26]. **Figure 3.2a** shows an example of arabinose, which is a monosaccharide unit commonly found in hemicellulose and pectin. Arabinose can accommodate linkages at carbons 1, 2, 3, and 5. The 1,5-linked arabinose (also named 5-Ara) is featured with a uniquely large C5 chemical shift of 68-70 ppm, instead of the regular 62-63 ppm value for the terminal arabinose residues (t-Ara). Similarly, an additional linkage at carbon 2 will change the isotropic ^{13}C chemical shifts of C2 from 80-82 ppm to 88-90 ppm.

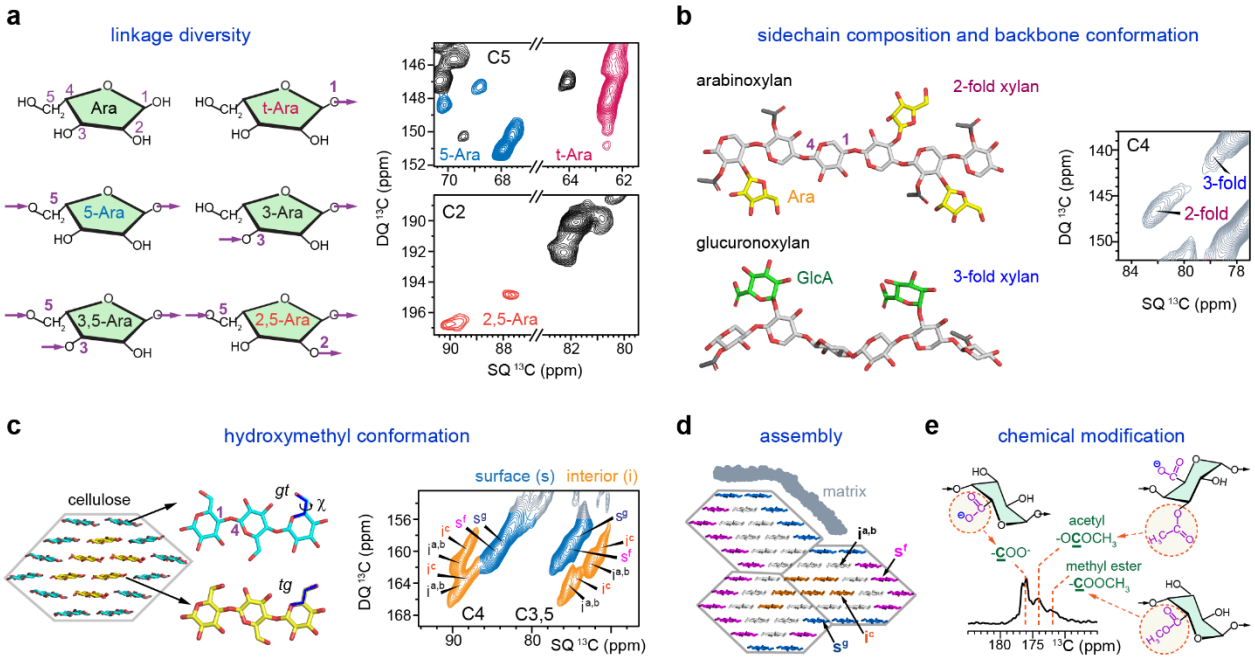

Figure 3.2 Structural complexity and peak multiplicity of plant carbohydrates. **a,** Linkage diversity demonstrated on the arabinose (Ara) unit. The C5 and C2 regions of Ara are shown using a 2D refocused J-INADEQUATE spectrum measured on *Brachypodium*. The key linkage sites are labeled in the structure, with the corresponding NMR signals color-coded in the spectra. **b,** Simplified examples of arabinoxylan and glucuronoxylan containing arabinose and glucuronic acid (GlcA) sidechains. Note that the two-fold and three-fold helical screw symmetries of the xylan backbone have been observed for both arabinoxylan and glucuronoxylan. The C4 region of xyloses

Figure 3.2 (cont'd)

differentiates signals from two-fold and three-fold xylan in spruce. **c,** The interior (yellow) and surface (cyan) glucan chains of cellulose have distinct hydroxymethyl conformation and NMR signals (from eucalyptus). The cross-section view of a hypothetical 18-chain model for an elementary cellulose fibril is shown. **d,** Scheme of a large bundle formed by three elementary cellulose microfibrils. The high-order assembly forms the basis for many different structural environments for the glucose residues. **e,** Acetylation and methyl-esterification change the ^{13}C chemical shifts of the carbonyl group (measured on *Arabidopsis*). The structures of three galacturonic acid residues, a component commonly found in pectin, are shown to match the NMR peaks. These panels are reconstituted from three previous publications[11, 26, 27].

Third, the next factor is the broad distribution of conformations in native polysaccharides. The hemicellulose xylan is typically found in two-fold (flat-ribbon) or three-fold (non-flat) helical screw conformations[11, 28-30], which can be resolved using the signals of C4 and C1, the two important carbon sites for the glycosidic bond along the xylan backbone (**Figure 3.2b**). Moreover, a continuous band of xylose C4 signals has also been observed in dried stems of *Arabidopsis* and the hydrated stems of hardwoods (eucalyptus and poplar)[11, 29], as a result of the coexistence of many intermediate conformations between the two-fold and three-fold conformers. For cellulose, the surface and interior glucan chains have distinct hydroxymethyl conformations and resolvable signals (**Figure 3.2c**). The dominant conformation is *gauche-trans* (*gt*) for surface chains and *trans-gauche* (*tg*) for interior chains, as recently determined using ^{1}H-^{1}H distance measurement[31].

Fourth, the high-order assembly of macromolecules also contributes to peak multiplicity. Since 2016, we have been consistently observing 7 major types of glucose units (namely types a-g) for the cellulose in the native cell walls of *Arabidopsis*, *Brachypodium*, maize, switchgrass, rice,

spruce, eucalyptus, and poplar (see the representative spectrum in **Figure 3.2c**). In the cross-section of a cellulose microfibril, the outer layer contains two types of surface chains (types f and g) with differentiable levels of water association, which has been examined using the water-edited ^{1}H polarization transfer method[31]. As a result, connections have been made between these two types of glucan chains with the concepts of hydrophobic and hydrophilic surfaces widely used in cellulose research, but further investigations are needed to assess this hypothesis. Types a and b are directly underneath the surface layer and exhibit strong cross peaks with surface chains[32, 33]. The origin of type-c has remained vague for long. Type-c chains failed to show strong cross peaks with surface chains; therefore, it is supposed to be deeply embedded in the core so that it is separated from the surface chains by an intermediate layer (types a and b). However, the prevalent 18-chain model of a cellulose microfibril (as shown in **Figure 3.2c**) could not accommodate such structural complexity. The presence of type-c chains can be better justified by the existence of larger bundles formed by multiple microfibrils as shown in **Figure 3.2d**. The other two forms (d and e) have noticeably weaker signals. In addition, there are ongoing efforts using density functional theory calculations to connect structural variables, such as hydrogen bonds and other torsional angles, to the observed ^{13}C chemical shifts of cellulose[34, 35].

Finally, chemical modifications, primarily acetylation and methyl esterification, are carried out by corresponding enzymes. The well-resolved carbonyl peaks (**Figure 3.2e**) and methyl signals can be used for distinguishing these functional groups and identifying their occurrence sites in cell wall polysaccharides[27, 29, 36, 37]. The connection between the NMR fingerprint and the structural complexity is exemplified using plant carbohydrates but the five general principles are appliable to polysaccharides in other organisms.

3.5 Biochemical and SsNMR Strategies to Facilitate Resonance Assignment

Nowadays, resonance assignment is no longer a barrier for well-studied materials (e.g., bacteria and plants), but additional caution should be taken when dealing with uncharted biosystems. In this scenario, the best strategy is to combine ssNMR with chemical and genomics approaches. Chemical assays, such as compositional and linkage analyses, should be conducted to guide and validate the NMR assignment. When ambiguity arises, the ssNMR assignment can be validated by either generating mutant lines or using chemical extracts that selectively deplete the carbohydrate component of interest, thus rendering the corresponding signals absent in the spectra. An alternative is to compare with other species that are known to lack such molecules. When *Arabidopsis thaliana* was first studied using 2D and 3D ssNMR, a xyloglucan-deficient mutant and a pectin-depleted cell wall were compared with the wild-type plant[38, 39]. A similar strategy was applied to the model fungus *Aspergillus fumigatus*, for which a collection of mutants depleting chitin, α-1,3-glucan, galactomannan, and galactosaminogalactan were generated[13].

A carbohydrate platform, complex carbohydrate magnetic resonance database (CCMRD), was created in 2019 to index the results from publications containing high-resolution ssNMR spectra and well-validated assignments[40]. Started from the 450 entries available in 2019[40], the number of entries has increased to 720 as of Spring 2022. The database was designed to accommodate the prolific growth of datasets in upcoming years, in preparation for implementing automated approaches to encourage more ssNMR studies of carbohydrate systems[41, 42].

Sequential assignment experiments (e.g., NCACX and NCOCX) have been extensively employed in protein ssNMR, but there are no equivalent methods for carbohydrates. This is partially due to the high degree of polymerization and the lack of nitrogen sites in most carbohydrates. However, novel carbohydrate sequencing strategies, analogous to those used in

solution NMR[44], might be promising in solids for small polysaccharides with highly variable monosaccharide structures or with nitrogenated sugar units. This possibility needs to be explored.

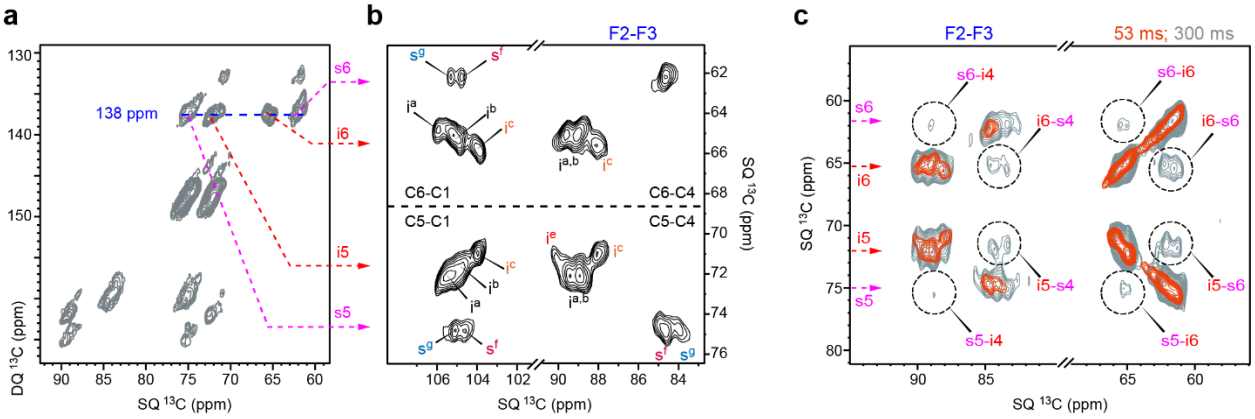

Figure 3.3 3D correlation experiments for resonance assignment and structural determination. **a,** Representative F1-F3 plane of a 3D refocused-INADEQUATE-CORD spectrum of spruce (with CORD recoupling turned off). **b,** 2D F2-F3 plane of spruce extracted from a 3D spectrum (with 53 ms CORD) at F1 = 138 ppm. **c,** Long-range inter-chain interactions in spruce cellulose highlighted using dash line circles. The F2-F3 planes (F1 = 138 ppm) of 3D spectra measured with short (53 ms) and long (300 ms) CORD mixing periods were compared to uncover low-intensity but highly relevant cross peaks reporting interactions. (Adapted from Shekar *et al.*)[43]

3D ^{13}C correlation experiments provide additional spectral resolution over the 2D spectra and have been applied to biomass characterization[29, 38]. Recently, a 3D ^{13}C DQ-SQ-SQ correlation experiment has been successfully employed to resolve the polysaccharide signals in the never-dried stems of spruce[43]. This experiment is free of the cube's body-diagonal, and the involvement of the DQ chemical shifts in the F1 dimension allows simultaneous observation of two parallel lines of signals in the F2-F3 planes for each coupled spin pair, e.g., s5-s6 and i5-i6 shown in **Figure 3.3a, b**. The experiment is also useful for identifying long-range correlation cross peaks when a

long-mixing ^{13}C-^{13}C mixing was implemented (**Figure 3.3c**). Novel 3D experiments, such as those involving underexplored nuclei (e.g. ^{1}H and ^{17}O for carbohydrates)[45-47], should be developed to counterbalance the lack of nitrogen sites in regard to proteins.

3.6 Underutilized Structural Data for Determining the Packing Interface

The ssNMR technique also provides information on the spatial proximities of biopolymers, which is its most appreciated capability, besides quantification, for investigating cell wall materials. A recent study of the plant secondary cell walls has proposed three comparative models for the secondary cell wall organization and the lignin-carbohydrate interface in the grass, hardwood, and softwood (**Figure 3.4**)[11]. The function of hemicellulose xylan is found to be conformation-dependent: coating the even surface of cellulose via its flat-ribbon structure and primarily binding lignin domains via its non-flat conformation using preferential surface contacts stabilized by numerous electrostatic interactions. However, this selectivity is partially compromised in the stems of woody plants, with some of the lignin particles closely packed with the flat-ribbon part of xylan as well as the surface of cellulose fibrils, which is likely forced by molecular crowding in the densely packed stems of trees. Finally, the softwood spruce has the best molecular mixing of biopolymers on the nanoscale, with lignin adapting a dispersive distribution instead of forming nanoparticles.

Although over-simplified, these cartoon illustrations were built on the structural concepts derived from more than 1,172 site-specific ssNMR data[11, 30]. The constantly improving resolution allows us to differentiate more carbon sites, and naturally a larger number of long-range cross peaks between these carbons. The complete dataset includes 508 intermolecular cross peaks (272 cross peaks in woody plants and 234 cross peaks in model plants), 475 ^{13}C-T_1 and ^{1}H-$T_{1\rho}$ relaxation time constants, and 189 water-edited intensities for various carbon sites. Nowadays, each pair of

2D ^{13}C-^{13}C correlation spectra (with long and short mixing times as contrast) allow the identification of 90-100 intermolecular cross peaks in a plant sample (**Figure 3.5a**).

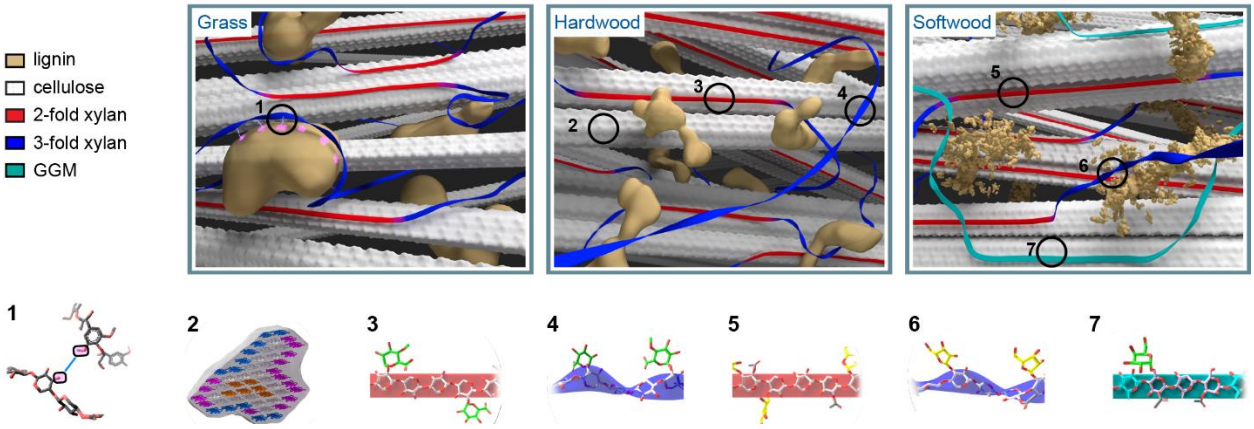

Figure 3.4 Illustrative cartoons of biopolymer packing in different plant species restrained by ssNMR. The scheme summarizes the structural concepts summarized from 1,172 site-specific data regarding lignin (yellow), cellulose (white), xylan in two-fold (red, flat-ribbon) and three-fold (blue, twisted bands) screw conformations, and mannan (GGM, green). The structures of polymers are shown for seven regions. Figure adapted from reference[11].

As shown in polymer contact maps (**Figure 3.5b)**, intermolecular cross peaks are categorized by their intensities following different chemical or structural motifs, such as the acetyl (Ac) and furanose ring of xylan (Xn) in two-fold or three-fold (2f/3f) screw conformations, the interior (i) and surface (s) chains in cellulose, as well as the ring carbons of guaiacyl (G) and syringyl (S) monolignol units, and the methoxyl (OMe) groups of lignin. In this way it is straightforward to recognize the structure selectivity of molecules in forming physical contacts: lignin primarily interacts with xylan instead of the other carbohydrate component (i.e., cellulose), and the S-residue contributes more to lignin-carbohydrate packing in poplar when compared with the G-residue. In particular, numerous strong cross peaks were identified in spruce, revealing the

unique architecture of softwood and the homogeneous mixing of biopolymers happening on the nanoscale[11]. This systemic treatment of data allowed us to efficiently summarize the principal interactions stabilizing the polymer interface, but we still lack a method of converting the overwhelmingly abundant spectroscopic data into structurally meaningful physical models.

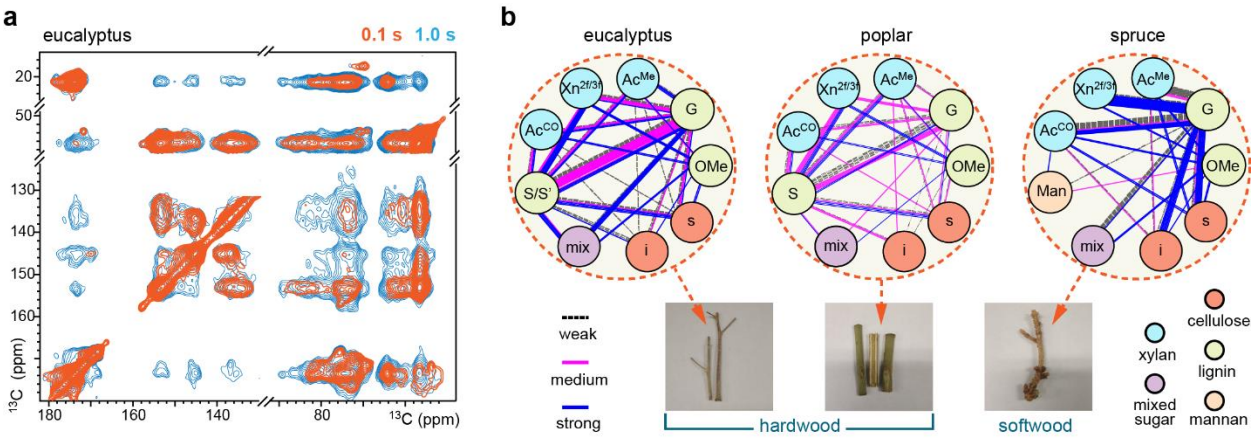

Figure 3.5 Structural restraints on polymer packing. **a,** Overlay of 2D dipolar-gated PDSD spectra with short (0.1 s; orange) and long (1.0 s; cyan) mixing times resolving 98 intermolecular cross peaks in eucalyptus. **b,** Polymer contact map summarizing 272 intermolecular cross peaks arising among different structural units of xylan (cyan), lignin (lime/yellow), cellulose (red), mannan (light orange), and mixed sugars (purple). All cross peaks are categorized as strong (blue), medium (magenta), and weak (black dash lines) following their intensities. Pictures of the samples used for NMR characterization are shown. Figures adapted from reference[11].

Unfortunately, the analyses of water-association and motional dynamics are even cruder. We rely on these data to distinguish the rigid and mobile domains of polymers, and to differentiate the hydrophilic matrix from the hydrophobic centers of the cell walls. Two polymers tightly packed to form larger aggregates are expected to show reduced mobility and water permeability, in addition to strong intermolecular cross peaks. These features were spotted for the chitin-α-glucan cores in the fungal cell wall of *Aspergillus fumigatus*[12, 13] and the cellulose-two-fold-xylan core in

plant secondary cell walls[11, 30]. In contrast, molecules well dispersed in the matrix should be better solvated and more dynamic, expecting fewer and weaker cross peaks with other components. This trend has been observed in the β-glucan matrix of fungi and the hemicellulose-pectin matrix of plants. A complication arises as a single polysaccharide may have multiple domains with distinct dynamics. For instance, the pectin in *Arabidopsis* has a mobile domain forming the matrix that is prone to alkali extraction as well as an alkali-resistant and stiff domain that is packed with cellulose microfibrils[21].

Data analysis in this polymer/residue-specific manner has made it straightforward to cross-inspect different samples (e.g., from different species or mutant lines), but it has compromised the advantage of high-resolution site-specific data. We are facing the challenge of converting structural restraints into an energy-minimized ensemble of macromolecular structures. Inputs from experts in modeling and structural calculation will be highly valued.

3.7 Critical Factors of MAS-DNP on Cell Wall Samples

Nowadays, MAS-DNP is regularly used for characterizing cell walls by providing the needed sensitivity for visualizing the interface between lowly-populated polymers (compared to the bulk of the cell wall) and by enabling high-resolution characterization of unlabeled materials using 2D ^{13}C correlation experiments at natural isotopic abundance (1.1%)[48-52]. These applications have been reviewed recently[53]. Here we will mainly address two frequently asked questions regarding 1) how to ensure homogeneous hyperpolarization throughout the biomaterial being studied, and 2) how to mitigate the line-broadening effect due to the low temperature of the sample (e.g., ~100 K for MAS-DNP) and the paramagnetic effect from the radicals.

It is expected that samples with homogenous hyperpolarization should show comparable patterns in 1D ^{13}C spectra measured with and without microwave (MW) irradiation (**Figure 3.6a**).

Occasionally, lipid polymers and membranes may be poorly polarized. Lipid polymers mostly form separated domains and their signals are distinguishable from carbohydrate peaks, thus alleviating the interference. Cell walls are porous layers on the surface of the cell that can be freely accessed and easily penetrated by the biradicals dissolved in solution[54], which are small molecules only 2-3 nm across[55, 56]. The polarization of electron spins will be first transferred to the ^{1}H in the biradical[57-59], and then to those in the solvent, and finally spread over a very long range via ^{1}H-^{1}H spin diffusion, which ensures homogeneous hyperpolarization, for example, across at least tens of nanometers as estimated using wood fibers[60]. This is the case for carbohydrates as the nuclear relaxation times are long enough at low temperature, unlike the lipids that still relax quickly.

Cellulose can largely retain its linewidth at the cryogenic temperature of DNP due to its high crystallinity (**Figure 3.6b**, **c**)[50, 61]. Indeed, we have observed impressively narrow full width at half maximum (FWHM) linewidths of 0.9-1.0 ppm for cotton cellulose on a 600 MHz/395 GHz MAS-DNP, which allowed us to track the glucose units in the Iα and Iβ model allomorphs frequently observed in model cellulosic materials[49]. The C-H dipolar order parameters are typically larger than 0.85 for cellulose; therefore, motion is not a major factor to consider for these fibrils. In contrast, hemicellulose and pectin are highly dynamic, and their signals are broadened out at a low temperature as shown in **Figure 3.6b**. The resulting structure contains conformers with different isotropic chemical shifts, which inhomogeneously broaden the NMR lines. Despite this limitation, some peaks of these mobile components have been recently resolved in 2D ^{13}C-^{13}C correlation MAS-DNP spectra[50]. Under most circumstances, biradicals mainly stay in the solvent instead of being packed with biomolecules. Therefore, the paramagnetic relaxation enhancement effect from the biradicals is neither pronounced nor an issue for biomaterial DNP (**Figure 3.6c**), especially for the solutions with a low concentration (5-10 mM) of radicals[62].

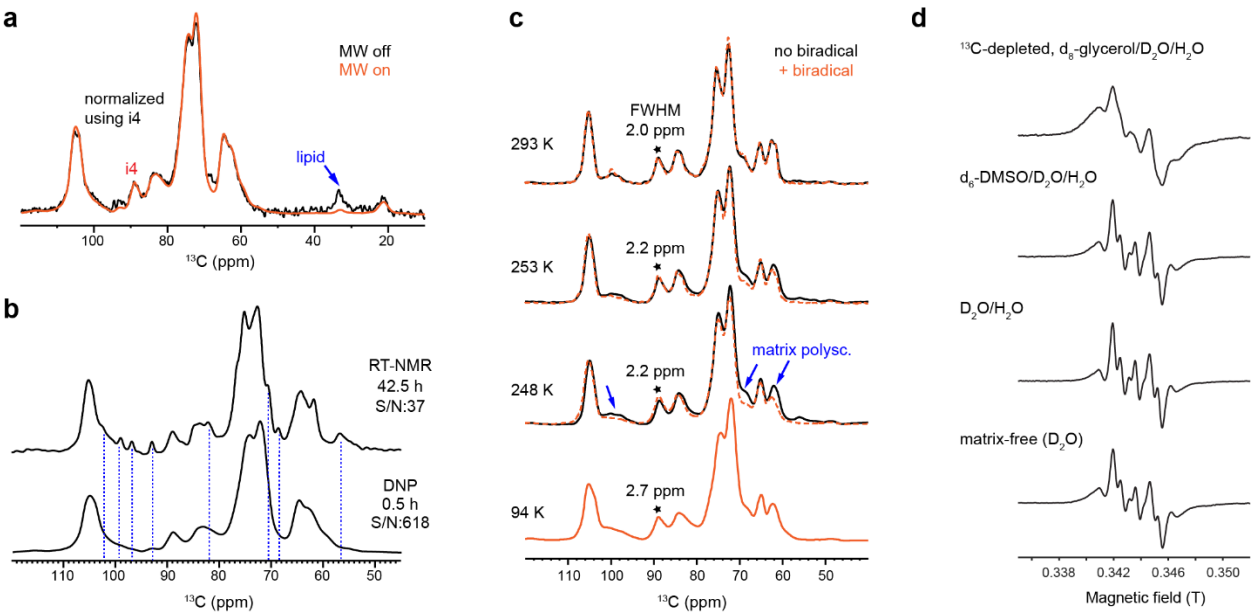

Figure 3.6 Evaluation of sample and biradical conditions for MAS-DNP. **a,** Uniform hyperpolarization across the cell wall confirmed by the consistent patterns of carbohydrate signals in the microwave-on (MW on) and microwave-off (MW off) spectra when normalized by the interior cellulose carbon 4 peaks (i4). Lipid components are not well polarized. **b,** Time saving and improved signal-to-noise ratios of DNP over room-temperature NMR (RT-NMR). Blue dash lines indicate the signals of mobile components of rice stems broadened by the cryogenic temperature of DNP. **c,** ^{13}C spectra of extracted *Arabidopsis* cell walls at different temperatures and biradical conditions. **d,** Room-temperature EPR spectra of AMUPol at 9.6 GHz in four rice stem samples (enhancement factor $\varepsilon_{on/off}$ of 38-57 on 600 MHz/395 GHz MAS-DNP) prepared using different solvents and procedures. The figures are remade from references[50, 61].

3.8 Expanding the Application of MAS-DNP in Carbohydrate Research

For most cell wall molecules in rice stems, it has been shown that the linewidths in 2D ^{13}C correlation spectra collected on a 600 MHz/395 GHz MAS-DNP were comparable to those collected on a 400 MHz conventional ssNMR[50]. This is observed not only for cellulose but also

for some rigid portions of matrix polymers, such as the backbone of xylan. However, the most dynamic components, such as the arabinose sidechains of xylan, still evades DNP detection. The current resolution allowed for resonance assignment and compositional analysis of unlabeled cell wall materials. In the same study, the MAS-DNP analysis was further combined with room-temperature measurements (by conventional ssNMR) of relaxation and dipolar order parameters, providing a more complete assessment of these unlabeled samples[50]. However, an unaccomplished task is to develop methods that can efficiently detect intermolecular cross peaks under natural isotopic abundance[63].

Further questions include whether we can better handle matrix polysaccharides and other mobile components such as intracellular molecules. Finding answers to these questions might introduce new applications of MAS-DNP to biomaterial characterization. It is also unclear though if MAS-DNP at higher fields would help resolve the resolution issues. If the line-broadening is inhomogeneous, higher fields will not help, otherwise higher magnetic fields are desirable. However, in most cases, MAS-DNP at 800 MHz/527 GHz or higher generates lower enhancement in general. Ongoing progress notably in biradical design might improve MAS-DNP at high field[56, 64, 65]. It should be noted that smaller rotors (e.g., 1.3 mm)[66, 67] are needed to improve the enhancement, but the use of a small rotor sacrifices the absolute sensitivity during ^{13}C detection as compared to the conventional 3.2 mm rotors. Overall, high-field DNP bears some promises but still presents major challenges and may not surpass the sensitivity provided by the lower-field DNP setups used today.

Unexpectedly, optimizing DNP samples, especially for the mixing protocol of biradical (in the solvent) and the cell wall material, is still fully empirical at this stage. Typically, the most limiting step of a MAS-DNP project is finding a sample preparation maximizing the sensitivity by

both a great enhancement factor and a relatively short DNP buildup time. This often requires re-processing the samples with solvents and radicals for multiple cycles, by trial and error. Recently, a benchtop EPR was used to examine a collection of rice samples before subjection to MAS-DNP measurements (**Figure 3.6d**). The EPR spectra suggests that the biradicals are favorably partitioning into the solvent in the samples with good sensitivity enhancement ($\varepsilon_{on/off}$ of 38-57). It will be beneficial if a protocol could be developed to rapidly screen the condition of radicals in the samples and pick up the most promising ones. This thematic is presently becoming even more relevant with the introduction of new radicals to the commercially available library.

The next question is how to place MAS-DNP into the context of other structural and biophysical techniques. Cryo-EM has established a leading role in protein structure determination and viral research but is not frequently used for carbohydrates or biomaterials. We have developed a strategy for analyzing the structures of unlabeled biomaterials by combining the nanoscale resolution of cryo-electron tomography (CET) with the atomic-level information provided by MAS-DNP (**Figure 3.7**)[68]. CET of a cellulose fibril (cross-sectional FWHM width of 5.2-5.7 nm) revealed periodic structure along the fibril axis formed by two wrapped, yet not twisted filaments, each of which generally agrees with an 18-chain model for elementary cellulose microfibril. MAS-DNP resolved the carbon sites of interior and surface glucan chains, and intensity quantification supports the 18-chain arrangement with a relatively low extent of fibrillar bundling. The sample used was cellulose fibrils synthesized *in vitro* by protein machinery isolated from moss, which provided a simplified target for this successful attempt. It remains to be seen if DNP has synergism with cryo-EM for investigating more sophisticated polysaccharides and biopolymer complexes such as cell walls.

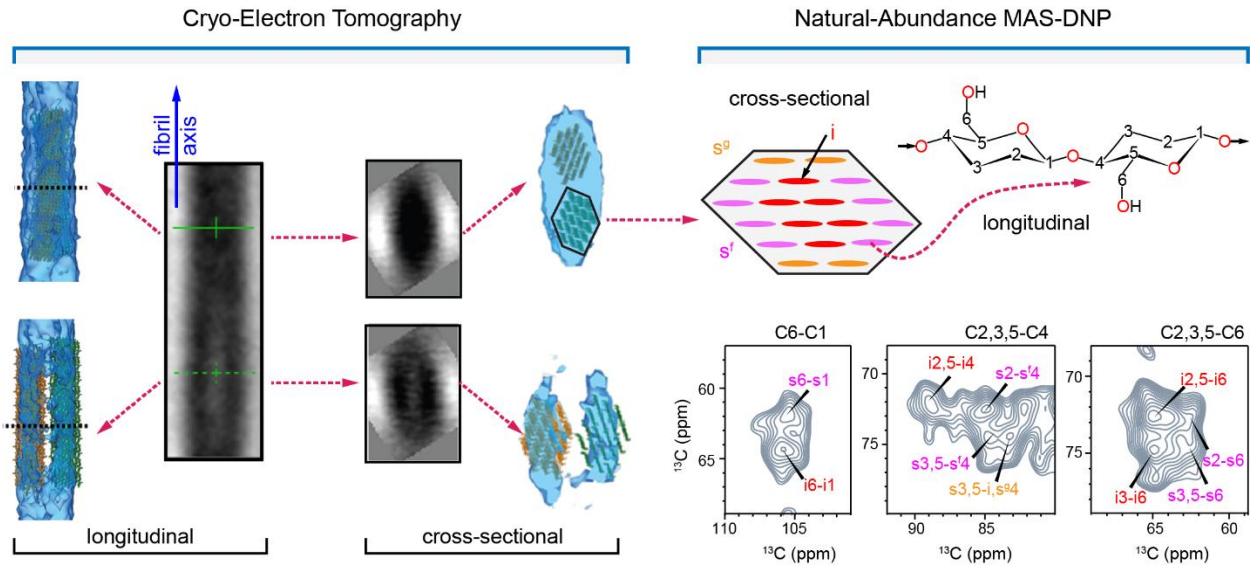

Figure 3.7 Synergism of Cryo-EM and DNP for carbohydrate characterization. Subtomogram averages of *in vitro* fibers and isosurface rendering of the density map (left) reveal a fibril structure formed by two wrapped filaments. Each filament fibril can be fit to an 18-chain cellulose model, where the conformational distribution of glucan chains and the surface-to-interior ratio are analyzed by MAS-DNP (right). A 2D ^{13}C–^{13}C CHHC spectrum is included here to show the spectral resolution. Adapted from Deligey et al.[68]

3.9 Conclusion

Over the last decade, significant advances have been made in the structural characterization of cell walls and carbohydrate-based biomaterials. However, due to the high heterogeneity of these cellular samples and the unique structural complexity of carbohydrates, the tools currently available for polymer research and protein structural biology are not fully applicable here. Development of semi-automatic tools for facilitating resonance assignment, novel pulse sequences and structural calculation methods for carbohydrate structural determination, as well as new radicals and optimized protocols for the application of high-field MAS-DNP may lead to breakthroughs in carbohydrate and biomaterial research.

3.10 Acknowledgments

This work was supported by the U.S. Department of Energy (grant no. DE-SC0021210). F.D. was supported as part of the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under award number DE-SC0001090. The National High Magnetic Field Laboratory is supported by National Science Foundation through NSF/DMR-1644779 and the State of Florida.

REFERENCES

1. Duus, J.; Gotfredsen, C. H.; Bock, K., Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000,** *100*, 4589-4614.

2. Atalla, R. H.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. *Science* **1984,** *223* (4633), 283-285.

3. El hariri El Nokab, M.; Habibi, M. H.; Alassmy, Y. A.; Abduljawad, M. M.; Alshamrani, K. M.; Sebakhy, K. O., Solid State NMR a Powerful Technique for Investigating Sustainable/Renewable Cellulose-Based Materials. *Polymers* **2022,** *14*, 1049.

4. Foston, M., Advances in solid-state NMR of cellulose. *Curr. Opin. Biotechnol.* **2014,** *27*, 176-184.

5. Ghassemi, N.; Poulhazan, A.; Deligey, F.; Mentink-Vigier, F.; Marcotte, I.; Wang, T., Solid-State NMR Investigations of Extracellular Matrices and Cell Walls of Algae, Bacteria, Fungi, and Plants. *Chem. Rev.* **2021**, in press.

6. Reichhardt, C.; Cegelski, L., Solid-state NMR for bacterial biofilms. *Mol. Phys.* **2014,** *112* (7), 887-894.

7. Kelly, J. E.; Chrissian, C.; Stark, R. E., Tailoring NMR experiments for structural characterization of amorphous biological solids: A practical guide. *Solid State Nucl. Magn. Reson.* **2020,** *109*, 101686.

8. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. *J. Biomol. NMR* **2018,** *70* (2), 123-131.

9. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-state NMR spectroscopy. *Nat. Rev. Methods Primers* **2021,** *1* (1), 2.

10. Markwick, P. R. L.; Malliavin, T.; Nilges, M., Structural Biology by NMR: Structure, Dynamics, and Interactions. *PLoS Comput. Biol.* **2008,** *4*, e1000168.

11. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

12. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

13. Chakraborty, A.; Fernando, L. D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latgé, J. P.; Wang, T., A molecular vision of fungal cell wall organization by functional genomics and solid-state NMR. *Nat. Commun.* **2021,** *12*, 6346.

14. Hou, G.; Yan, S.; Trébosc, J.; Amoureux, J.-P.; Polenova, T., Broadband homonuclear correlation spectroscopy driven by combined R2nv sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J. Magn. Reson.* **2013,** *232*, 18-30.

15. Paëpe, G. D.; Lewandowski, J. R.; Loquet, A.; Böckmann, A.; Griffin, R. G., Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **2008,** *129* (24), 245101.

16. Lange, A.; Luca, S.; Baldus, M., Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids. *J. Am. Chem. Soc.* **2002,** *124* (33), 9704-9705.

17. Ader, C.; Schneider, R.; Seidel, K.; Etzkorn, M.; Becker, S.; Baldus, M., Structural Rearrangements of Membrane Proteins Probed by Water-Edited Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2009,** *131* (1), 170-176.

18. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water–Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis thaliana from Polarization Transfer Solid-State NMR. *J. Am. Chem. Soc.* **2014,** *136* (29), 10399-10409.

19. Williams, J. K.; Schmidt-Rohr, K.; Hong, M., Aromatic spectral editing techniques for magic-angle-spinning solid-state NMR spectroscopy of uniformly ^{13}C-labeled proteins. *Solid State Nucl. Magn. Reson.* **2015,** *72*, 118-126.

20. Gao, Y.; Mortimer, J. C., Unlocking the architecture of native plant cell walls via solid-state nuclear magnetic resonance. *Methods Cell Biol.* **2020,** *160*, 121-143.

21. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis* primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiol.* **2015,** *168* (3), 871-884.

22. Poulhazan, A.; Dickwella Widanage, M. C.; Muszynski, A.; Arnold, A. A.; Warschawski, D. E.; Azadi, P.; Marcotte, I.; Wang, T., Identification and Quantification of Glycans in Whole Cells: Architecture of Microalgal Polysaccharides Described by Solid-State Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **2021,** *46*, 19374-19388.

23. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* **2020,** *11* (1), 6081.

24. Cresswell, R.; Dupree, R.; Brown, S. P.; Pereira, C. S.; Skaf, M. S.; Sorieul, M.; Dupree, P.; Hill, S., Importance of water in maintaining softwood secondary cell wall nanostructure. *Biomacromolecules* **2021,** *8*, 4669-4680.

25. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: a critical review. *Solid State Nucl. Magn. Reson.* **2020,** *107*, 101660.

26. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and Dynamics of Brachypodium Primary Cell Wall Polysaccharides from Two-Dimensional 13C Solid-State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2014,** *53* (17), 2840-2854.

27. Kirui, A.; Du, J.; Zhao, W.; Barnes, W.; Kang, X.; Anderson, C. T.; Xiao, C.; Wang, T., A pectin Methyltransferase modulates polysaccharide dynamics and interactions in *Arabidopsis* primary cell walls: Evidence from solid-state NMR. *Carbohydr. Polym.* **2021,** *270*, 118370.

28. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Pöppler, A.-C.; Brown, S. P.; Deazevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

29. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional 13C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015,** *54* (14), 2335-2345.

30. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 1-9.

31. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct determination of hydroxymethyl conformations of plant cell wall cellulose using ^{1}H polarization transfer solid-state NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

32. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

33. Wang, T.; Hong, M., Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **2016,** *67* (2), 503-514.

34. Yang, H.; Kubicki, J. D., A density functional theory study on the shape of the primary cellulose microfibril in plants: effects of C6 exocyclic group conformation and H-bonding. *Cellulose* **2020,** *27* (5), 2389-2402.

35. Yang, H.; Wang, T.; Oehme, D.; Petridis, L.; Hong, M.; Kubicki, J. D., Structural factors affecting 13C NMR chemical shifts of cellulose: a computational study. *Cellulose* **2018,** *25* (1), 23-36.

36. Wang, T.; Chen, Y.; Tabuchi, A.; Cosgrove, D. J.; Hong, M., The target of β-expansin EXPB1 in maize cell walls from binding and solid-state NMR studies. *Plant Physiol.* **2016,** *172* (4), 2107-2119.

37. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10*, 4978.

38. Dick-Pérez, M.; Zhang, Y.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell-wall polysaccharides by two-and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* **2011,** *50* (6), 989-1000.

39. Dick-Perez, M.; Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Multidimensional solid-state NMR studies of the structure and dynamics of pectic polysaccharides in uniformly ^{13}C-labeled Arabidopsis primary cell walls. *Magn. Reson. Chem.* **2012,** *50*, 539-550.

40. Kang, X.; Zhao, W.; Widanage, M. C. D.; Kirui, A.; Ozdenvar, U.; Wang, T., CCMRD: a solid-state NMR database for complex carbohydrates. *J. Biomol. NMR* **2020,** *74* (4), 239-245.

41. Li, D.-W.; Hansen, A. L.; Yuan, C.; Bruschweiler-Li, L.; Bruschweiler, R., DEEP picker is a deep neural network for accurate deconvolution of complex two-dimensional NMR spectra. *Nat. Commun.* **2021,** *12*, 5229.

42. Cordova, M.; Balodis, M.; Simoes de Almeida, B.; Ceriotti, M.; Emsley, L., Bayesian probabilistic assignment of chemical shifts in organic solids. *Sci. Adv.* **2021,** *7*, eabk2341.

43. Shekar, S. C.; Zhao, W.; Fernando, L. D.; Hung, I.; Wang, T., A ^{13}C three-dimensional DQ-SQ-SQ correlation experiment for high-resolution analysis of complex carbohydrates using solid-state NMR. *J. Magn. Reson.* **2022,** *336*, 107148.

44. Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L., Advancing Solutions to the Carbohydrate Sequencing Challenge. *J. Am. Chem. Soc.* **2019,** *141*, 14463-14479.

45. Phyo, P.; Hong, M., Fast MAS ^{1}H–^{13}C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73* (12), 661-674.

46. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206* (1), 66-72.

47. Shen, J.; Terskikh, V.; Struppe, J.; Hassan, A.; Monette, M.; Hung, I.; Gan, Z.; Brinkmann, A.; Wu, G., Solid-state 17O NMR study of α-d-glucose: exploring new frontiers in isotopic labeling, sensitivity enhancement, and NMR crystallography. *Chem. Sci.* **2022,** *13*, 2591-2603.

48. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paëpe, G., Rapid Natural-Abundance 2D ^{13}C–^{13}C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew. Chem. Int. Ed.* **2012,** *51* (47), 11766-11769.

49. Kirui, A.; Ling, Z.; Kang, X.; Widanage, M. C. D.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic resolution of cotton cellulose structure enabled by dynamic nuclear polarization solid-state NMR. *Cellulose* **2019,** *26* (1), 329-339.

50. Zhao, W.; Kirui, A.; Deligey, F.; Mentink-Vigier, F.; Zhou, Y.; Zhang, B.; Wang, T., Solid-state NMR of unlabeled plant cell walls: high-resolution structural analysis without isotopic enrichment. *Biotechnol. Biofuels* **2021,** *14* (1), 14.

51. Berruyer, P.; Gericke, M.; Moutzouri, P.; Jakobi, D.; Bardet, M.; Karlson, L.; Schantz, S.; Heinze, T.; Emsley, L., Advanced characterization of regioselectively substituted methylcellulose

model compounds by DNP enhanced solid-state NMR spectroscopy. *Carbohydr. Polym.* **2021,** *262*, 117944.

52. Perras, F. A.; Luo, H.; Zhang, X.; Mosier, N. S.; Pruski, M.; Abu-Omar, M. M., Atomic-level structure characterization of biomass pre-and post-lignin treatment by dynamic nuclear polarization-enhanced solid-state NMR. *J. Phys. Chem. A* **2017,** *121* (3), 623-630.

53. Chakraborty, A.; Deligey, F.; Quach, J.; Mentink-Vigier, F.; Wang, P.; Wang, T., Biomolecular complex viewed by dynamic nuclear polarization solid-state NMR spectroscopy. *Biochem. Soc. Trans.* **2020,** *48* (3), 1089-1099.

54. Takahashi, H.; Ayala, I.; Bardet, M.; De Paëpe, G.; Simorre, J. P.; Hediger, S., Solid-state NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. *J. Am. Chem. Soc.* **2013,** *135* (13), 5105-10.

55. Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem. Int. Ed.* **2013,** *52* (41), 10858-10861.

56. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

57. Mentink-Vigier, F.; Dubroca, T.; Van Tol, J.; Sigurdsson, S. T., The distance between g-tensors of nitroxide biradicals governs MAS-DNP performance: The case of the bTurea family. *J. Magn. Reson.* **2021,** *329*, 107026.

58. Mentink-Vigier, F.; Vega, S.; De Paëpe, G., Fast and accurate MAS–DNP simulations of large spin ensembles. *Phys. Chem. Chem. Phys.* **2017,** *19* (5), 3506-3522.

59. Perras, F. A.; Pruski, M., Large-scale ab initio simulations of MAS DNP enhancements using a Monte Carlo optimization strategy. *J. Chem. Phys.* **2018,** *149* (15), 154202.

60. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of pretreated wood fibers using dynamic nuclear polarization. *J. Phys. Chem. C* **2019,** *123* (50), 30407-30415.

61. Wang, T.; Park, Y. B.; Caporini, M. A.; Rosay, M.; Zhong, L.; Cosgrove, D. J.; Hong, M., Sensitivity-enhanced solid-state NMR detection of expansin's target in plant cell walls. *Proc. Natl. Acad. Soc. USA* **2013,** *110* (41), 16444-16449.

62. Takahashi, H.; Hediger, S.; De Paëpe, G., Matrix-free dynamic nuclear polarization enables solid-state NMR ^{13}C–^{13}C correlation spectroscopy of proteins at natural isotopic abundance. *Chem. Commun.* **2013,** *49* (82), 9479-9481.

63. Märker, K.; Paul, S.; Fernández-de-Alba, C.; Lee, D.; Mouesca, J.-M.; Hediger, S.; De Paëpe, G., Welcoming natural isotopic abundance in solid-state NMR: probing π-stacking and

supramolecular structure of organic nanoassemblies using DNP. *Chem. Sci.* **2017,** *8* (2), 974-987.

64. Zhai, W.; Lucini Paioni, A.; Cai, X.; Narasimhan, S.; Medeiros-Silva, J.; Zhang, W.; Rockenbauer, A.; Weingarth, M.; Song, Y.; Baldus, M.; Liu, Y., Postmodification via Thiol-Click Chemistry Yields Hydrophilic Trityl-Nitroxide Biradicals for Biomolecular High-Field Dynamic Nuclear Polarization. *J. Phys. Chem. B* **2020,** *124* (41), 9047-9060.

65. Lund, A.; Casano, G.; Menzildjian, G.; Kaushik, M.; Stevanato, G.; Yulikov, M.; Jabbour, R.; Wisser, D.; Renom-Carrasco, M.; Thieuleux, C.; Bernada, F.; Karoui, H.; Siri, D.; Rosay, M.; Sergeyev, I. V.; Gajan, D.; Lelli, M.; Emsley, L.; Ouari, O.; Lesage, A., TinyPols: a family of water-soluble binitroxides tailored for dynamic nuclear polarization enhanced NMR spectroscopy at 18.8 and 21.1 T. *Chem. Sci.* **2020,** *11* (10), 2810-2818.

66. Chaudhari, S. R.; Berruyer, P.; Gajan, D.; Reiter, C.; Engelke, F.; Silverio, D. L.; Copéret, C.; Lelli, M.; Lesage, A.; Emsley, L., Dynamic nuclear polarization at 40 kHz magic angle spinning. *Phys. Chem. Chem. Phys.* **2016,** *18* (15), 10616-10622.

67. Wisser, D.; Karthikeyan, G.; Lund, A.; Casano, G.; Karoui, H.; Yulikov, M.; Menzildjian, G.; Pinon, A. C.; Purea, A.; Engelke, F.; Chaudhari, S. R.; Kubicki, D.; Rossini, A. J.; Moroz, I. B.; Gajan, D.; Copéret, C.; Jeschke, G.; Lelli, M.; Emsley, L.; Lesage, A.; Ouari, O., BDPA-Nitroxide Biradicals Tailored for Efficient Dynamic Nuclear Polarization Enhanced Solid-State NMR at Magnetic Fields up to 21.1 T. *J. Am. Chem. Soc.* **2018,** *140* (41), 13340-13349.

68. Deligey, F.; Frank, M. A.; Cho, S. H.; Kirui, A.; Mentink-Vigier, F.; Swulius, M. T.; Nixon, B. T.; Wang, T., Structure of In Vitro-Synthesized Cellulose Fibrils Viewed by Cryo-Electron Tomography and 13C Natural-Abundance Dynamic Nuclear Polarization Solid-State NMR. *Biomacromolecules* **2022,** *in press.*

CHAPTER 4: POLYSACCHARIDE ASSEMBLIES IN FUNGAL AND PLANT CELL WALLS EXPLORED BY SOLID-STATE NMR

Review paper reprinted with permission: *Structure* 31, 1375-1385 (2023)

Authors: Liyanage D. Fernando[#], Wancheng Zhao[#], Isha Gautam, Ankur Ankur, and Tuo Wang

[#]Equally contributions

4.1 Summary

Structural analysis of macromolecular complexes within their natural cellular environment presents a significant challenge. Recent applications of solid-state NMR (ssNMR) techniques on living fungal cells and intact plant tissues have greatly enhanced our understanding of the structure of extracellular matrices. Here we selectively highlight the most recent progress in this field. Specifically, we discuss how ssNMR can provide detailed insights into the chemical composition and conformational structure of pectin, and the consequential impact on polysaccharide interactions and cell wall organization. We elaborate on the use of ssNMR data to uncover the arrangement of the lignin-polysaccharide interface and the macrofibrillar structure in native plant stems or during degradation processes. We also comprehend the dynamic structure of fungal cell walls under various morphotypes and stress conditions. Finally, we assess how the combination of NMR with other techniques can enhance our capacity to address unresolved structural questions concerning these complex macromolecular assemblies.

4.2 Introduction

The cell walls of plants and fungi play vital roles in cell shape, mechanics, integrity, adhesion, and extensibility[1-4]. These macromolecular assemblies are also crucial for energy and carbon storage, as well as antimicrobial resistance, and serve as potential targets for novel antifungal agents[5, 6]. From a structural perspective, cell walls represent highly sophisticated

biomaterials created by nature, inspiring extensive efforts to develop artificial materials that mimic the chemical and physical principles governing macromolecular assembly[7, 8]. However, the characterization of such nanocomposites in their native state presents significant challenges.

Recent advancements in solid-state NMR (ssNMR) techniques have enabled high-resolution investigation of cell wall materials by leveraging methods initially developed for NMR structural biology of proteins, nucleic acids, and polymers[9-12]. This approach has been adopted due to the sufficient spectroscopic resolution available, which allows for the acquisition of numerous structural constraints to visualize the complex composition of biomolecules within cell walls. Additionally, it addresses the limitations posed by the inherent heterogeneity and complexity of the cell wall by utilizing simplified representations[13]. Extensive studies have been conducted on numerous plant and fungal species, leading to significant revisions in our understanding of cell wall structure[14]. In this review, we aim to provide a concise overview of the key principles in ssNMR methodology and the most recent structural findings, with a focus on those published within the past three years.

4.3 Methodology Advances Enabling SsNMR of Cellular Samples

One of the major merits of ssNMR lies in its ability to directly characterize living cells or intact tissues[15, 16]. Since solubilization or extraction is not necessary, the samples fully preserve the inherent physical and chemical characteristics of the biomolecules. The samples analyzed are typically enriched with NMR-active isotopes by incorporating ^{13}C-enriched precursors like ^{13}C-glucose, ^{13}C-maltose, and $^{13}CO_2$ (for plants), as well as ^{15}N-labeled amino acids and salts[17, 18]. Occasionally, selective-labeled precursors are used to simplify the spectra and track the biosynthesis of macromolecules[19]. Deuterated fatty acids can also be incorporated for determining the phospholipid profile in cellular membranes[20]. The collected material can be directly placed

into a magic-angle spinning (MAS) rotor, which can hold approximately between 1 mg (for a 0.7 mm diameter) and 100 mg (for a 4 mm diameter) of material, depending on the requirements of the NMR experiment.

A versatile toolbox is available for studying various aspects of biomolecules in cell walls and cellular samples, including molecular composition, structural variations, dynamical distribution, water association, and sub-nanometer packing[13, 14, 21]. While 1D spectra are useful for quick screening, high-resolution structural characterization often mandates the use of 2D and 3D homonuclear (such as ^{13}C-^{13}C) and heteronuclear (such as $^{15}N/^{1}H$-^{13}C) correlation experiments. These techniques rely primarily on ^{13}C detection, which provides excellent spectral dispersion to differentiate many magnetically inequivalent sites[22-27]. However, recent advancements in ^{1}H detection methods have also yielded valuable insights into the structure and packing of polysaccharides and proteins in plants, fungi, and bacteria[28-31]. The success of these studies has been facilitated by fast-to-ultrafast magic-angle spinning (MAS) techniques, with MAS rotation frequencies ranging from 60 to 110 kHz, where a higher level of deuteration and the presence of more dynamic molecules allow for reduced reliance on high MAS frequencies.

Dynamic Nuclear Polarization (DNP) is another significant advancement that has greatly expanded the capabilities of ssNMR in the structural characterization of macromolecular complexes[32-34]. In MAS-DNP, polarization of unpaired electrons in stable biradicals are transferred to ^{1}H and subsequently to other nuclei of interest, resulting in enhanced sensitivity[35]. The prominent biradicals used extensively in structural studies include AMUPol and the recently developed AsympolPOK[36, 37]. The 600 MHz/395 GHz MAS-DNP instrument strikes a favorable balance between spectral resolution and sensitivity enhancement, and a 30-80-fold boost can be easily achieved for cellular samples. This breakthrough has effectively overcome the sensitivity

limitations of NMR, enabling the use of small sample quantities, which is particularly valuable for difficult-to-replicate samples, exploring lowly populated molecules or macromolecular complexes embedded in bulk structures, and even characterizing fully unlabeled biosamples without the need for isotopic enrichment[38-43].

4.4 Pectin Methylation Influencing Plant Primary Cell Wall Structure

The mechanical roles of cellulose microfibrils, hemicellulose, and pectin in plant primary cell walls have been a topic of ongoing debate for many decades[44]. One prevailing model suggests the existence of a tethered polymer network, where a single hemicellulose strand (e.g., a xyloglucan) could simultaneously bind to multiple cellulose microfibrils and hold them together to form a load-bearing network[45, 46]. Pectin is an acidic polymer regulating wall porosity, pH, and ionic balance, and is influencing cell expansion and differentiation[47]. However, pectin is often considered as a separate component that forms a gel-like matrix, providing reinforcement to the cellulose-hemicellulose network. This perspective was primarily based on biochemical findings, wherein the vast majority of pectin is extractable by strong alkali, and the acidic pectin backbones do not bind to cellulose *in vitro*[48]. A significant advancement brought in by ssNMR is the identification of the interaction between pectin and cellulose within intact primary plant cell walls.

Pectins are partially interconnected complexes of polysaccharides and proteoglycans[47, 49, 50]. The major domains include homogalacturonan (HG), rhamnogalacturonan-I (RG-I) with branched arabinan and galactan, as well as RG-II[51]. By employing ssNMR, even with a simple 1D ^{13}C spectrum, it is possible to discern various specific chemical sites within the galacturonic acid (GalA) residues that comprise HG, the rhamnose (Rha) residues that alternate with GalA along the RG-I backbone, as well as the arabinan sidechains (**Figure 4.1a**)[52]. Such spectral resolution forms the basis for tracking the intermolecular cross peaks between major pectin domains and cellulose

microfibrils using 2D ^{13}C-^{13}C correlation experiments[53, 54]. However, RG-II, which is less prevalent and structurally more intricate compared to HG and RG-I, has yet to be investigated using high-resolution ssNMR techniques.

Previously, it was believed that pectin backbone and cellulose were physically separated entities, but the numerous pectin-cellulose cross peaks identified by ssNMR have unambiguously revealed their close spatial proximity in the cell wall[54, 55]. The new insight has contributed to a paradigm shift in our understanding of the architecture of the primary cell wall[44, 56], and comprehensive overviews of these ssNMR studies can be found in multiple recent reviews[14, 57]. Despite these advancements, the precise function of the packing between pectin and cellulose in supporting the structure of the cell wall remains unclear. Recently, mesoscale coarse-grained molecular dynamics (CGMD) simulations conducted on onion epidermal cell walls have shed some light on this matter[58]. The noncovalent interactions between multiple cellulose fibrils were found to be the primary factor determining the mechanical properties of the cell wall. Notably, the sliding of two aligned cellulose microfibrils against each other was identified as a significant contributor to the plasticity of the cell wall. In contrast, hemicellulose and pectin were suggested to play more indirect roles, potentially by influencing the arrangement of cellulose within the cell wall[58].

Growing evidence emphasizes the significance of HG methylation in influencing the organization of the cell wall. At first glance, this may appear contradictory, as methylation only involves localized chemical modifications to the carboxyl groups of GalA residues[59, 60], while the rearrangement of cell wall polymers happens on the nanoscale. An emerging concept is that methylation alters the charge and physicochemical properties of HG, and may induce repulsion among HG chains, thereby providing the force to expand the interfibrillar space between cellulose

microfibrils[61]. This mechanism has been proposed as a potential regulator of cell expansion and shaping, alongside the more widely accepted turgor-driven mechanism.

Dupree, Hong, and coworkers have shown that pectin methylation can change the conformational structure, interactions, and dynamics of polysaccharides in the primary cell wall[62]. They compared the pectin structure in wild-type *Arabidopsis* with that of the *gosamt1 gosamt2* mutant, which had a reduced degree of methyl esterification in HG. The mutant exhibited an increased abundance of HG in a structure known as the 2-fold screw conformation (2_1), with its C6 chemical structure as $-COO^-$ in this sample (**Figure 4.1b**). These observations match the characteristics of the egg-box structure that refers to the cross-linking of two anionic GalA units that joined two adjacent HG chains by a Ca^{2+} ion. The reduced methyl esterification in the mutant promoted the formation of egg-box clusters and enhanced HG aggregation.

Conversely, the content of the methylated $-COOCH_3$ motif in the 3-fold screw conformation (3_1) was reduced in the *gosamt1 gosamt2* mutant[62]. The double mutant also exhibited strong signals indicating a reduction in the chain length of HG, allowing the identification of signals from the reducing ends of the GalA units in both α- and β-configurations. Importantly, this study represents the first successful application of ssNMR to elucidate the detailed structural characteristics of HG *in muro*, providing direct experimental evidence supporting the widely accepted model of the pectin egg-box structure within the plant cell wall.

NMR data showed that HG methylation has an impact on the polysaccharide organization in the cell wall. Pectin backbones showed larger dipolar order parameters and stronger interactions with cellulose in the *gosamt1 gosamt2* mutant[62]. These changes in cell wall structure were found to be associated with impaired polarized growth in the double mutant. Interestingly, these findings appear to contrast with recent observations from other studies using different mutants and

experimental conditions. An increase in pectin methylation is often accompanied by stronger pectin-cellulose interactions and impaired plant growth as reported in ssNMR analyses of *Arabidopsis* primary cell walls and focused on mutants that produced shorter HG chains, cell walls with pH-alteration, or segments along the inflorescences stem[63-65].

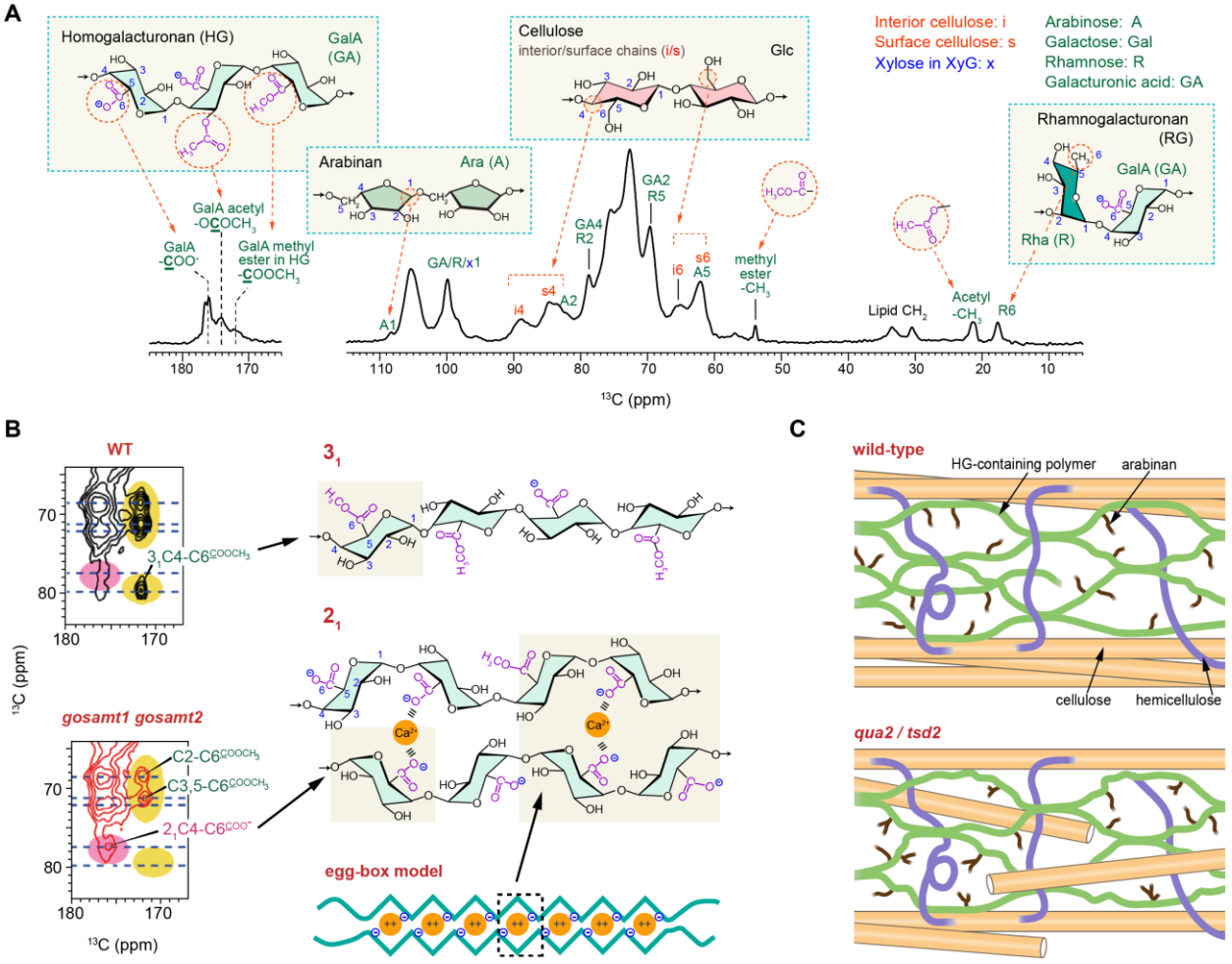

Figure 4.1 Pectin Structure and Its Effect on Primary Plant Cell Wall. **a,** Representative 1D ^{13}C ssNMR spectra with resolvable signals associated with pectin and cellulose structural motifs. Carbon numbers are shown for key sugar units. **b,** Selected regions of 2D ^{13}C-^{13}C spectra resolving the conformation and methylation state of GalA units in *Arabidopsis* pectin. Two-fold (2_1) and three-fold (3_1) screw conformations of HG displaying distinct chemical environments at the

Figure 4.1 (cont'd)

carbonyl site. An egg-box model is shown to depict two cross-linked HG chains. **c,** Structural rearrangement of the cell wall in *Arabidopsis* qua2/tsd2 mutant. Figures 1a, c adapted with permission from Kirui et al. *Carbohydr. Polym.* (2021). Figures 1b adapted with permission from Temple et al. *Nat. Plants* (2022).

The situation becomes more intricate due to another recent study involving two allelic mutants (*qua2* and *tsd2*) of a pectin methyltransferase[52]. These mutants exhibit similar cell wall composition and pectin methyl esterification levels as the wild-type sample, but, surprisingly, displayed stiffer pectin and stronger interactions with cellulose (**Figure 4.1c**). Atomic Force Microscopy (AFM) and Field Emission Scanning Electron Microscopy (FESEM) revealed a decrease in cellulose bundling in both mutants[66], and the more dispersed microfibril arrangement should have facilitated the extensive interactions as observed by ssNMR analysis.

These findings highlight how ssNMR can serve as a valuable tool to bridge the knowledge gaps between the chemical structure of carbohydrates, plant biology, and cell wall organization. However, pectin's role in cell wall structure remains enigmatic. It is now evident that pectin methylation can exert diverse effects on the cell wall structure, and there is no universal principle that applies to all scenarios. It is likely that the structural function of pectin is indirect, influencing the arrangement of cellulose microfibrils and hemicellulose within the cell wall, which requires further investigation.

4.5 Lignin-Carbohydrate Packing in Secondary Plant Cell Walls

Within the secondary cell wall, the coalescence of multiple cellulose microfibrils occurs frequently, e.g. an illustration of a 3-microfibril bundle is given in **Figure 4.2a**, which is accompanied by the association of hemicelluloses (xylan and galactoglucomannan) and lignin to

form larger macrofibrils[67, 68]. Inspired by the seminal work of Dupree and colleagues[69], many multidimensional ssNMR studies have been conducted in the past seven years to unravel the molecular architecture of secondary plant cell walls. These studies have led to two significant revisions in our understanding of lignocellulosic biomass. First, the chemical composition and conformational structure of xylan dictate its binding specificity within the cell wall[70, 71]. Second, the lignin-carbohydrate interface is stabilized by electrostatic interactions, with lignin primarily associated with the non-flat domains of xylan. However, in certain cases, lignin can also interact and co-localize with the xylan-coated cellulose microfibrils as the secondary target[72, 73].

In dehydrated environments, such as the dried stem of *Arabidopsis*, xylan exhibits a broad range of helical screw conformations that can be monitored through the continuous band observed in its C4 signals in ssNMR spectra[69]. In well-hydrated cell walls, including those of *Arabidopsis*, spruce, maize, switchgrass, rice, and *Brachypodium*, xylan preferentially adopts either 2-fold or 3-fold helical screw conformations resulting in two distinct C4 signals[70, 72-74]. This is illustrated by the flat-ribbon and non-flat structures depicted in **Figure 4.2a**. Notably, a continuous distribution of xylan helical screw conformation was even observed in hydrated hardwood materials, such as eucalyptus and poplar, which is likely induced by factors such as lower hydration levels and/or molecular crowding within hardwood stems[73]. In the case of sorghum, the secondary cell wall is primarily composed of 3-fold xylan conformations[75], while in *Brachypodium*, both the leaf and root showed two xylan conformers, but the stem shows a continuous range of conformations.

The presence of a two-fold flat-ribbon structure is relatively uncommon and was found to be facilitated by xylan's deposition onto the smooth surface of a cellulose microfibril[76]. Moreover, it has been discovered that an evenly distributed pattern of substitutions (e.g., by glucuronic acid or arabinose sidechains) along the xylan chain is crucial for maintaining this flat-ribbon structure

and promoting its interaction with cellulose[71]. This also explains why sorghum lacks this conformation as xylan is frequently and irregularly substituted by arabinosyl residues. On the other hand, the three-fold xylan plays a dominant role in carbohydrates' interactions with lignin, primarily through the polar functional groups, which highlights the significance of electrostatic interactions in stabilizing the interface between lignin nanodomains and carbohydrates (see the grass model in **Figure 4.2b**)[72].

The molecular architecture of lignin-carbohydrate interactions varies across different plant species (**Figure 4.2b**). Comparisons between grasses, hardwoods, and softwoods have revealed a gradual decrease in the domain size and polymer separation between lignin and polysaccharides. Lignin-cellulose packing interactions, which were found to be relatively limited in grass samples, also became abundant in woody plants[73]. During the analysis of the sparsely populated biopolymer interface, which exhibits structural deviations from the overall equilibrium state of the entire cell wall, ssNMR experiments employing spectral editing (e.g. the selected polarization of lignin aromatics or methoxy group or xylan acetyls) selection and DNP enhancement have proven to be instrumental[77, 78].

Dupree and coworkers successfully fit the ssNMR data on molecular fraction, polymer conformation, and polymer spatial proximity into a macrofibril observed under Cryo-SEM (**Figure 4.2c**)[74]. This large bundle that is tens of nanometers across to accommodate ~50 elementary microfibrils, each of which consists of 18 glucan chains. SsNMR analysis uncovered two distinct conformations for galactoglucomannan (GGM), reminiscent of the functional conformations in xylan. It was found that GGM undergoes conformational changes when bound to cellulose in a semi-crystalline manner. Some of the bound GGMs and two-fold xylan molecules can attach to the surface of the same microfibril, while the other GGMs contribute to the matrix alongside three-

fold xylan. The entire carbohydrate core is enveloped by a lignin layer, completing the molecular architecture of the microfibril assembly.

Water is another important structural component of the secondary cell wall, and high-temperature (105 °C) oven-drying can induce irreversible alteration to the structure of Monterey pine (Pinus radiata), even after subsequent rehydration[79]. SsNMR analysis of the rehydrated sample revealed a tighter packing between xylan and cellulose, along with the separation of some mannan into mobile phase, adopting a conformation similar to that observed in solution. These results support the hypothesis that water plays a crucial role in mediating the packing interactions between xylan and cellulose (**Figure 4.2d**). Importantly, this study provides direct evidence of the irreversible change in biomass ultrastructure caused by harsh drying procedures. It should be noted that another ssNMR study has demonstrated fully reversible changes of spruce and poplar through the process of lyophilization and rehydration[73].

Li, Kang, Yelle and colleagues have utilized ssNMR to investigate the perturbation of lignin linkages and lignin-polysaccharide packing by termite digestion [80]. This was acommplished by feeding ^{13}C-labeled sapwood sections of Canadian poplar (a natural hybrid *Populus* × *canadensis*), a hardwood, to a phylogenetically higher termite species (*Nasutitermes*) and Monterey pine, a softwood, to a lower termite species (*Cryptotermes*). The abundance of different lignin linkages and lignin-polysaccharide interactions in the termite diet and feces was assessed using 2D ^{13}C-^{13}C radio frequency-driven recoupling (RFDR) and long-range proton-driven spin diffusion (PDSD) spectra. Both termite species were found to effectively dissociate the electrostatic interface between lignin and polysaccharides but only the lower termite species can perturb the structure of the residual lignin. This study highlights the remarkable capability of ssNMR spectroscopy in examining the structural changes occurring in lignocellulosic materials,

and the same approach can be applied to investigate biomass degradation by various biological agents such as white/brown-rot fungi, as well as in industrial processes[81].

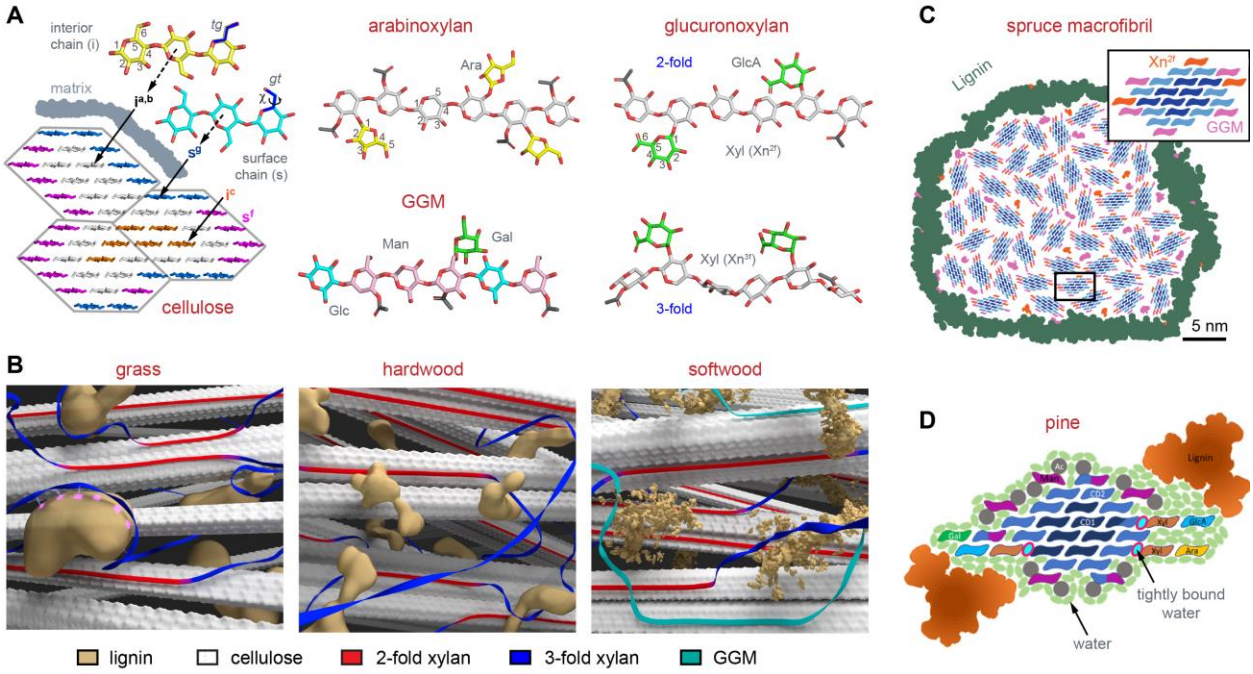

Figure 4.2 Macrofibril Structure and Lignin-Carbohydrate Interface. **a,**Structural features of polysaccharides in plant secondary cell walls resolvable by ssNMR. **b,** Structural view of the aromatic-carbohydrate packing interface in grass species, hardwood, and softwood. **c,** Best fit of NMR data collected on spruce secondary cell walls into a macrofibril. A local cellulose microfibril with two domains of glucan chains (light and dark blue) is shown with 2-fold xylan (Xn^{2f}) and galactoglucomannan (GGM) attached. **d,** Illustration of loosely associated water and tightly bound water in pine secondary cell wall. Figures 2a,b adapted from Kirui et al. *Nat. Commun.* (2022). Figure 2c modified from Terrett et al. *Nat. Commun.* (2019). Figure 2d adapted from Cresswell et al. *Biomacromolecules* (2021).

4.6 Dynamic Structure of Fungal Cell Wall: Challenge and Opportunity

Stark and colleagues have been at the forefront of utilizing ssNMR for nearly two decades

to study melanin deposition and its interactions with cell wall polysaccharides in *Cryptococcus* and *Saccharomyces* species[19, 82-85]. Recent ssNMR investigations on *Aspergillus* and *Schizophyllum*, with a focus on the cell wall structures, were largely inspired by these studies, and were built upon research strategies employed in plant cell wall characterization, albeit with methodological adaptations[57, 86]. However, an unexpected problem that significantly impeded progress was the dynamic nature of fungal cell walls, characterized by their remarkable ability to adapt to varying growth conditions and environmental contexts[2, 87]. These microorganisms exhibit an impressive capacity to compensate for the absence of specific polysaccharides through complex biosynthetic reactions, thereby generating a structurally intact cell wall[88]. Such adaptability, unlike their plant counterparts, challenges the notion that certain cell wall polysaccharides, such as cellulose and pectin in primary cell walls, are essential. The dynamic structural changes in fungal cell walls not only enable the survival and adaptation of these microbes to diverse environments but also present technical barriers to our fundamental understanding of these organelles and the development of antifungal medications to combat invasive infections[89]. As a result, it is crucial to carefully monitor the status of the samples and ensure the reproducibility of the culture conditions. In many cases, it becomes necessary to prepare fresh batches of samples specifically for lengthy experiments. Complications further arise from the distinct composition and assembly of biomacromolecules in different fungal species and the vast array of mutant strains found in nature. Therefore, it is crucial to first identify the conserved features of fungal cell walls and elucidate the structural functions of their key polysaccharides before investigating the variations across the wide spectrum of fungal species and strains.

4.7 Morphotype-Dependent Structure of *Aspergillus* Cell Walls

During the transition of morphotypes in the fungal life cycle, the cell wall undergoes

significant changes in its nanoscale morphology and molecular composition[2, 90]. These structural transformations play a vital role in many biological processes such as spore germination, hyphal growth, and sporulation. In the case of *A. fumigatus*, the fungus forms asexual hydrophobic spores called dormant conidia (dorC) on a specialized hyphal structure called conidiophore[1]. In immunocompromised patients, when the inhaled conidia reach the alveoli in the lungs within 4-6 hours, the germination process takes place. Germination involves the uptake of water by the conidia, resulting in their swelling and transformation into swollen conidia (swoC). Subsequently, the swollen conidia germinate (gerC) into short hyphae known as germ tubes[1]. Following germination, the growth of hyphae initiates, leading to the formation of a colony. This mycelial colony represents an aggregated form of hyphae that invade the pulmonary tissues, causing a life-threatening disease known as aspergillosis[90].

SsNMR analysis of the intact *Aspergillus* mycelium offers valuable insight into the structural organization of the cell wall during the vegetative growth of the fungus in the lungs. These living fungal cells exhibited high spectral resolution, enabling the differentiation of various conformers from six major types of fungal polysaccharides, including chitin, chitosan, α-1,3-glucan, three types of β-glucan, as well as galactomannan (GM), and galactosaminogalactan (GAG) (**Figure 4.3a**)[86, 91]. The partially crystalline chitin has a narrow linewidth of 0.5-0.7 ppm, whereas the plant counterpart, cellulose microfibrils, shows a broader linewidth of 0.7-1.0 ppm on high-field magnets[92]. The mobile molecules present in the fungus and the plant pectin have comparable linewidths of 0.2-0.5 ppm[93]. Similar to the polymorphic structure observed in plant cellulose [92], chitin also exhibits high polymorphism. The structural arrangement of chitin in fungi is predominantly aligned with the α-allomorph[94], characterized by an antiparallel packing of chains[95, 96]. Through ssNMR analysis, it was discovered that the rigid core of the cell wall

comprises α-1,3-glucan, β-glucan, and chitin.

Of particular surprise was the identification of α-1,3-glucan's structural function, as it was previously considered insignificant in the cell wall assembly, akin to the underappreciated role of pectin in primary plant cell walls[97]. Unlike the highly mobile nature of pectin in plants, α-1,3-glucan was found to be the most rigid polysaccharide in *Aspergillus fumigatus* mycelia. It forms a tightly packed, dehydrated core through extensive physical interactions with chitin. This finding was further strengthened by the observation that α-1,3-glucans are crucial to the stiffness of the cell wall in chitin-deficient mutant (**Figure 4.3a**). Moreover, α-1,3-glucan has diverse distribution, with its physical presence in both the mobile and rigid domains and chemical existence in both the alkali-soluble and insoluble fractions, as revealed by a combined chemical and NMR approach [91].

On the other hand, β-glucans in fungi, similar to xyloglucan in plant primary cell walls, have been proposed to play a crucial role in cross-linking. Chemical analysis has identified the presence of a small fraction of covalently linked mannan-β-1,3-glucan-chitin complex in *Aspergillus*[87]. The β-glucans encompass diverse linkages, existing as linear β-1,3-glucan, branched β-1,3/1,6-glucan, and terminal β-1,3/1,4-glucan[90], which has not been fully characterized by ssNMR so far. However, structural analysis of β-glucans has shown that these molecules are essential for retaining water in the cell wall and contribute to the formation of a hydrated matrix. Outside of this inner domain, there is a surface shell enriched with highly mobile glycoproteins, GM, and GAG[91].

Loquet and colleagues reported the evolution of the *Aspergillus fumigatus* conidial cell wall by examining the structure of dorC, which was the starting material, as well as swoC and gerC, which were prepared by incubating dorC by additional 5 h and 8 h, respectively[98]. Compared to dorC, the swoC exhibited isotropic growth to double its cellular size, based on which the gerC

cells further underwent polarized growth, leading to the formation of germ tubes. The polysaccharides present in the rigid fractions of *A. fumigatus* conidia cell walls were found to be consistent with those of mycelial cell walls, containing chitin, α-glucan, and β-glucan[98]. Another study published almost at the same time also reported highly conserved carbohydrate cores in both mycelia and conidia cell walls, but using unlabeled *A. fumigatus* as enabled by the sensitivity-enhancement of MAS-DNP[99].

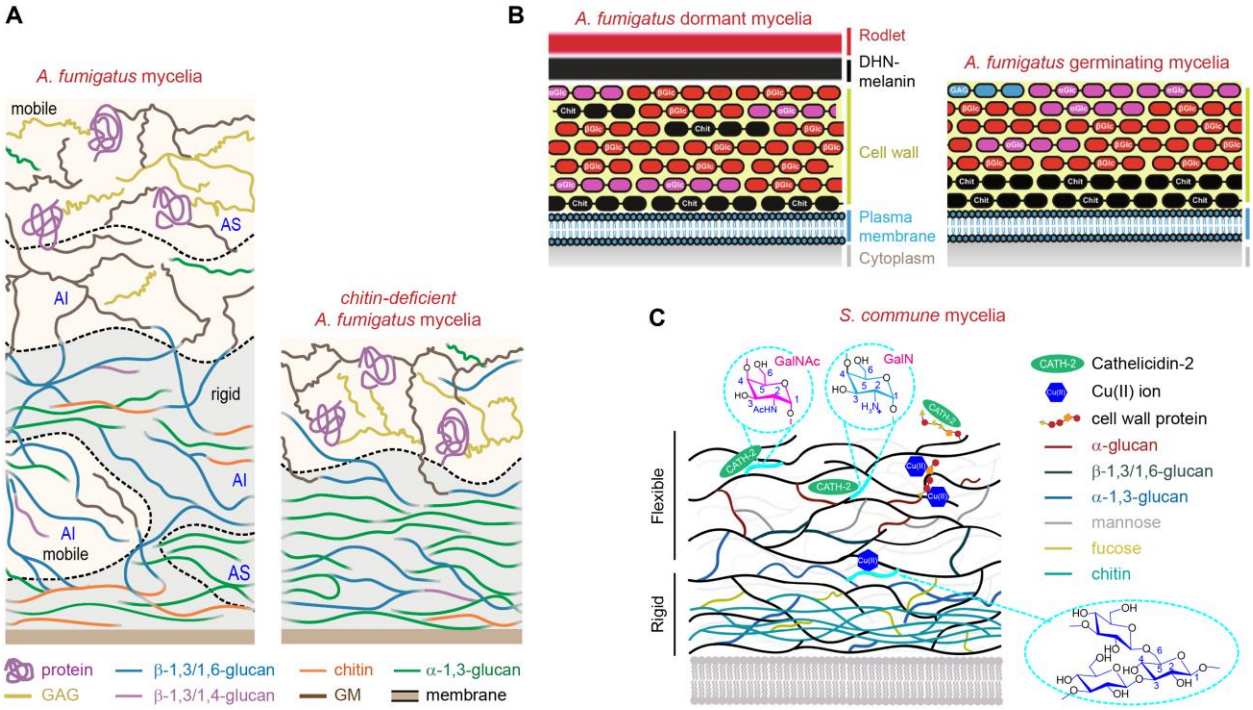

Figure 4.3 NMR-Restrained Structural Models of Fungal Cell Walls. **a,**Comparative views of mycelial cell walls in the parental strain (left) and chitin-deficient mutant (right) of *A. fumigatus*. **b,** Dormant and germinating conidial cell walls of *A. fumigatus*. **c,** Mycelial cell walls of *S. commune*, highlighting the binding polysaccharides (cyan) and proteins that bind Cu(II) ions and antimicrobial peptides. Figure 3a adapted from Chakraborty et al. *Nat. Commun.* (2021), Figure 3b adapted with permission from Lamon et al. *Proc. Natl. Acad. Sci. USA* (2023), and Figure 3c adapted with permission from Safeer et al. *Eur. J. Chem.* (2022) and Ehren et al. *Cell Surf.* (2020).

In the conidial cell wall, the ratio of α-glucan to β-glucan was approximately 1:3 in dorC and gerC, but became almost equal in swoC (**Figure 4.3b**), suggesting a less stable cell wall composition required for germination[98]. Both α-glucan and β-glucan exhibited increased hydration levels in swoC and gerC compared to dorC. Interestingly, α-glucan showed high hydration levels in all conidia cell walls[98], whereas it was poorly hydrated in mycelia cell walls[72], indicating differences in their structural organization. While the chitin content remained the same in the rigid portions of all three cell walls, chitin exhibited a noticeable decrease in both water-contact and structural polymorphism in gerC. This observation can be explained by a model where chitin becomes deeply embedded in the inner core and undergoes phase separation from glucans during the germinating stage (**Figure 4.3b**)[98]. In addition, the presence of GAG was detected by ssNMR in both swollen and germinating conidia, but not in dormant conidia, where its biosynthesis has yet to begin. This study demonstrates the capability of cellular ssNMR to effectively uncover the structural reorganization of the fungus during the transition between morphotypes.

Triglyceride signals were also identified in the mobile phase of *A. fumigatus* cells [98]. In other fungal species, such as the melanized *C. neoformans* and *S. cerevisiae* spore cell wall, triglyceride has been identified by ssNMR as a potential building molecule of stress-resistant cell walls[100]. The biological relevance of this molecule in *A. fumigatus* and its location (e.g., as a cell wall component or an intracellular molecule) requires careful evaluation.

4.8 Cell Wall Architecture of *Schizophyllum commune*

Baldus and coworkers conducted two ssNMR investigations of *Schizophyllum commune* mycelia[28,101]. The first study involved the analysis of sequentially digested cell walls, progressively removing molecules from the cell wall[101]. The rigid domain of the cell wall is composed of chitin and β-1,3/1,6-glucan, potentially linked together, as well as α-1,3-glucan and polymeric fucose

(**Figure 4.3c**). The mobile domain contains terminal hexoses included in α- and β-linked glucans. Mannose residues were also identified in the mobile phase, at least partially incorporated in mannosylated proteins. Using proton-detection methods and ultrahigh field magnets (1.2 GHz), the team further examined the binding sites of metals and antimicrobial peptides[28]. At a low ion concentration of 0.74 mM Cu(II), the paramagnetic relaxation enhancement (PRE) effect revealed signal quenching at the primary targets, which were mainly proteins. However, at a higher ion concentration of 18.5 mM, Cu(II) ions could penetrate into the cell wall region with β-1,3/1,6-glucans. Proton-detection methods were also employed to understand the binding of many other micronutrients, such as Ca^{2+}, Mg^{2+}, and anions, to the *S. commune* cell wall[102,103].

Additionally, the antimicrobial peptide Cathelicidin-2, which inhibits *S. commune* growth, was found to primarily bind to cell wall proteins[28]. This was demonstrated through chemical shift perturbation (CSP) observed in protein backbone Cα sites and the disappearance of signals from the sidechains of charged amino acids. Interestingly, the signals of GalN/GalNAc were also affected by the peptide[28]. This observation is intriguing as Cathelicidin-2 is typically positively charged with a theoretical isoelectric point (pI) in the range of 10-12 depending on the origin of the peptide[104]. At the same time, GalN can exist as cationic residues with an $-NH_3^+$ group under the pH of the fungal culture. The unidentified mechanisms underlying the association between the antimicrobial peptide and *S. commune* cell walls might inspire future studies.

4.9 Integrating NMR Restraints with Other Experimental Approaches

Chemical analysis, ssNMR, diffraction methods, imaging techniques can complement each other to complete the understanding of cell wall ultrastructure. Each of these methods rely on distinct physical and chemical principles, offering unique insights into various structural aspects at different length scales. Locally, ssNMR has been combined with chemical assays and sugar

analyses to confirm resonance assignment and provide detailed information on the composition and linkage patterns of carbohydrates[105]. However, the cross-linking between polymers and the extractability as probed by chemical analysis does not directly correlate with the physical packing as investigated by ssNMR[97]. For example, the majority of pectin and α-1,3-glucan are highly extractable by hot alkali, but they are both tightly associated with the crystalline core of the corresponding cell wall, cellulose in the plant and chitin in the fungus. And it has always been overlooked that a portion of pectin and α-1,3-glucan can never be fully isolated due to their physical entrapment by microfibrillar components[55, 91]. Similarly, the analysis of lignin also faces challenges in establishing consistency between ssNMR and the more widely used liquid NMR [106, 108]. Additionally, when it comes to polymer dynamics, it is not possible to directly combine NMR-reported parameters such as correlation time, relaxation time constants, or dipolar order parameters with values obtained from other techniques like persistence length[65].

Diffraction methods and imaging techniques are valuable for providing structural information at nanoscale or larger length scales. Consequently, ssNMR data can complement these techniques by offering atomic-to-nanometer structural information[109]. As we discussed earlier, this strategy has been applied to rationalize the organization of spruce microfibril[74]. Moreover, a combination of MAS-DNP and Cryo-electron tomography (CryoET) has been employed to examine the structure of in-vitro synthesized cellulose fibrils, revealing that they consist of two wrapped filaments fitting into an 18-chain microfibril[110]. Cryo-electron microscopy methods have also been employed to examine the fiber arrangement in the cell wall[111-113] and the structure of polysaccharide synthases[114-116]. Another example is the challenge of distinguishing proteins from different sources (e.g., cell walls, membrane proteins, and intracellular sources) using ssNMR alone. In such cases, confocal microscopy and atomic force microscopy (AFM) have been used to

investigate the assembly of RodA rodlets in *Aspergillus*, which were found to interact with melanin on the cell surface of DorC while swoC and gerC did not support rodlet formation[98]. However, caution is necessary when attempting to bridge the findings obtained from different approaches.

4.10 Conclusions and Perspectives

Significant progress has been made in applying ssNMR and other structural methods to unravel the complex structure of polysaccharides and associated biomacromolecules in plant and fungal cell walls. However, there are still numerous unanswered structural questions. For instance, the function of pectin in plant cell wall mechanics remains unknown, the resolution of lignin by ssNMR is insufficient, the role of GGM (galactoglucomannan) is not fully understood, and the identification of lignin-carbohydrate linkages[117] *in muro* remains challenging. Moreover, the fungal cell wall is even less explored, with many structural components, such as exopolysaccharides like glucuronoxylomannan (GXM) and galactoxylomannan (GalXM) in *Cryptococcus* capsules and β-1,6-glucan and phosphomannan in *Candida* sp., still requiring investigations. Understanding the mechanisms of cell wall remodeling in response to stress is also crucial. Recent research has unveiled that the halophilic fungus *Aspergillus sydowii* enhances the hydrophobicity and stiffness of its cell wall to resist osmotic pressure in hypersaline environments[118]. Similar changes have been observed in *A. fumigatus* during adaptation to internal stress caused by carbohydrate-deficiency and external stress such as antifungal treatment with caspofungin (unpublished results by Dickwella Widange et al.). It is essential to validate if these cell wall adaptations are universal features across various species during adaptation. Furthermore, high-resolution ssNMR techniques have been employed to address structural challenges in other organisms. An important application is the quantification and compositional analysis of carbohydrate components in algal cells, bacterial cell walls, and biofilms[31, 119-122]. As ssNMR

technology continues to evolve and our understanding of polysaccharide structure and cell wall assembly expands, it is now opportune to tackle important structural problems in the years to come.

4.11 Acknowledgments

This work was supported by the National Institutes of Health grant AI173270. The plant cell wall research was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under award number DE-SC0023702.

REFERENCES

1. van de Veerdonk, F. L.; Gresnigt, M. S.; Romani, M. G.; Latge, J. P., Aspergillus fumigatus morphology and dynamic host interactions. *Nat. Rev. Microbiol.* **2017,** *15*, 661-674.

2. Gow, N. A. R.; Lenardon, M. D., Architecture of the dynamic fungal cell wall. *Nat. Rev. Microbiol.* **2022,** *21*, 248-259.

3. Cosgrove, D. J., Building an extensible cell wall. *Plant Physiol.* **2022,** *189*, 1246-1277.

4. Cosgrove, D. J., Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* **2005,** *6*, 850-861.

5. Carpita, N. C.; McCann, M. C., Redesigning plant cell walls for the biomass-based bioeconomy. *J. Biol. Chem.* **2020,** *295*, 15144-15157.

6. Gow, N. A. R.; Latge, J. P.; Munro, C. A., The Fungal Cell Wall: Structure, Biosynthesis, and Function. *Microbiol. Spectr.* **2017,** *5*, FUNK-0035-2016.

7. Sarkar, P.; Bosneaga, E.; Auer, M., Plant cell walls throughout evolution: towards a molecular understanding of their design principle. *J. Exp. Bot.* **2009,** *60*, 3615-3635.

8. Paulraj, T.; Wennmalm, S.; Wieland, D. C. F.; Riazanova, A. V.; Dedinaite, A.; Pomorski, T. G.; Cardenas, M.; Svagan, A. J., Primary cell wall inspired micro containers as a step towards a synthetic plant cell. *Nat. Commun.* **2020,** *11*, 958.

9. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-State NMR Spectroscopy. *Nat. Rev. Methos Primers* **2021,** *1*, 2.

10. Kelly, J. E.; Chrissian, C.; Stark, R. E., Tailoring NMR experiments for structural characterization of amorphous biological solids: A practical guide. *Solid State Nucl. Magn. Reson.* **2020,** *109*, 101686.

11. Baldus, M., Biological solid-state NMR: Integrative across different scientific disciplines. *J. Struct. Biol. X* **2022,** *6*, 100075.

12. van der Wel, P. C. A., New applications of solid-state NMR in structural biology. *Emerg. Top. Life Sci.* **2018,** *2*, 57-67.

13. Zhao, W.; Deligey, F.; Shekar, S. C.; Mentink-Vigier, F.; Wang, T., Current limitations of solid-state NMR in carbohydrate and cell wall research. *J. Magn. Reson.* **2022,** *341*, 107263.

14. Ghassemi, N.; Poulhazan, A.; Deligey, F.; Mentink-Vigier, F.; Marcotte, I.; Wang, T., Solid-State NMR Investigations of Extracellular Matrixes and Cell Walls of Algae, Bacteria, Fungi, and Plants. *Chem. Rev.* **2022,** *122*, 10036-10086.

15. Warnet, X. L.; Arnold, A. A.; Marcotte, I.; Warschawski, D. E., In-Cell Solid-State NMR: An Emerging Technique for the Study of Biological Membranes. *Biophys. J.* **2015,** *109* (12), 2461-2466.

16. Renault, M.; Tommassen-van, B. R.; Bos, M. P.; Post, J. A.; Tommasseen, J.; Baldus, M., Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci. USA* **2012,** *109*, 4863-4868.

17. Gao, Y.; Mortimer, J. C., Unlocking the architecture of native plant cell walls via solid-state nuclear magnetic resonance *Methods Cell Biol.* **2020,** *160*, 121-143.

18. Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Wang, P.; Kang, X.; Wang, T., Preparation of Fungal and Plant Materials for Structural Elucidation Using Dynamic Nuclear Polarization Solid-State NMR. *J. Vis. Exp.* **2019,** *144*, e59152.

19. Chatterjee, S.; Prados-Rosales, R.; Tan, S.; Phan, V. C.; Chrissian, C.; Itin, B.; Wang, H.; Khajo, A.; Magliozzo, R. S.; Casadevall, A.; Stark, R. E., The melanization road more traveled by: Precursor substrate effects on melanin synthesis in cell-free and fungal cell systems. *J. Biol. Chem.* **2018,** *293* (52), 20157-20168.

20. Laydevant, F.; Mahabadi, M.; Llido, P.; Bourgouin, J. P.; Caron, L.; Arnold, A.; Marcotte, I.; Warschawski, D. E., Growth-phase dependence of bacterial membrane lipid profile and labeling for in-cell solid-state NMR applications. *BBA-Biomembranes* **2022,** *1864*, 183819.

21. Shcherbakov, A. A.; Medeiro-Silva, J.; Tran, N.; Gelenter, M. D.; Hong, M., From angstroms to nanometers: measuring interatomic distances by solid-state NMR. *Chem. Rev.* **2021,** *122*, 9848-9879.

22. Heise, H.; Seidel, K.; Etzkorn, M.; Becker, S.; Baldus, M., 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations. *J. Magn. Reson.* **2005,** *173*, 64-74.

23. Li, S.; Zhang, Y.; Hong, M., 3D 13C–13C–13C correlation NMR for de novo distance determination of solid proteins and application to a human α-defensin. *J. Magn. Reson.* **2010,** *202*, 203-210.

24. Shekar, S. C.; Zhao, W.; Fernando, L. D.; Hung, I.; Wang, T., A ^{13}C three-dimensional DQ-SQ-SQ correlation experiment for high-resolution analysis of complex carbohydrates using solid-state NMR. *J. Magn. Reson.* **2022,** *336*, 107148.

25. Pauli, J.; Baldus, M.; van Rossum, B. J.; de Groot, H.; Oschkinat, H., Backbone and Side-Chain 13C and 15N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *Chem. Biochem.* **2001,** *2*, 272-281.

26. Paëpe, G. D.; Lewandowski, J. R.; Loquet, A.; Böckmann, A.; Griffin, R. G., Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **2008,** *129* (24), 245101.

27. Cadars, S.; Lesage, A.; Emsley, L., Chemical Shift Correlations in Disordered Solids. *J. Am. Chem. Soc.* **2005,** *127* (12), 4466-4476.

28. Safeer, A.; Kleiburg, F.; Bahri, S.; Beriashvili, D.; Veldhuizen, E. J. A.; van Neer, J.; Tegelaar, M.; de Cock, H.; Wosten, H. A. B.; Baldus, M., Probing Cell-Surface Interactions in

Fungal Cell Walls by High-Resolution 1H-Detected Solid-State NMR Spectroscopy. *Chem. Euro. J.* **2022,** *29*, e202202616.

29. Phyo, P.; Hong, M., Fast MAS 1H–13C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73*, 661-674.

30. Marchand, T. L.; Schubeis, T.; Bonaccorsi, M.; Paluch, P.; Lalli, D.; Pell, A. J.; Andreas, L. B.; Jaudzems, K.; Stanek, J.; Pintacuda, G., 1H-Detected Biomolecular NMR under Fast Magic-Angle Spinning. *Chem. Rev.* **2022,** *122*, 9943-10018.

31. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206*, 66-72.

32. Ni, Q. Z.; Daviso, E.; Can, T. V.; Markhasin, E.; Jawla, S. K.; Swager, T. M.; Temkin, R. J.; Herzfeld, J.; Griffin, R. G., High Frequency Dynamic Nuclear Polarization. *Acc. Chem. Res.* **2013,** *46* (9), 1933-1941.

33. Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Copéret, C.; Emsley, L., Dynamic Nuclear Polarization Surface Enhanced NMR Spectroscopy. *Acc. Chem. Res.* **2013,** *46* (9), 1942-1951.

34. Takahashi, H.; Ayala, I.; Bardet, M.; De Paëpe, G.; Simorre, J. P.; Hediger, S., Solid-state NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. *J. Am. Chem. Soc.* **2013,** *135* (13), 5105-10.

35. Kundu, K.; Mentink-Vigier, F.; Feintuch, A.; Vega, S., DNP Mechanisms. *eMagRes* **2019,** *8* (3), 295-337.

36. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

37. Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem.* **2013,** *52* (41), 10858-10861.

38. Smith, A. N.; Märker, K.; Hediger, S.; De Paëpe, G., Natural Isotopic Abundance 13C and 15N Multidimensional Solid-State NMR Enabled by Dynamic Nuclear Polarization. *J. Phys. Chem. Lett.* **2019,** *10* (16), 4652-4662.

39. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paëpe, G., Rapid Natural-Abundance 2D 13C–13C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew. Chem.* **2012,** *51* (47), 11766-11769.

40. Zhao, W.; Kirui, A.; Deligey, F.; Mentink-Vigier, F.; Zhou, Y.; Zhang, B.; Wang, T., Solid-state NMR of unlabeled plant cell walls: high-resolution structural analysis without isotopic

enrichment. *Biotechnol. Biofuels* **2021,** *14* (1), 1-14.

41. Berruyer, P.; Gericke, M.; Moutzouri, P.; Jakobi, D.; Bardet, M.; Karlson, L.; Schantz, S.; Heinze, T.; Emsley, L., Advanced characterization of regioselectively substituted methylcellulose model compounds by DNP enhanced solid-state NMR spectroscopy. *Carbohydr. Polym.* **2021,** *262*, 117944.

42. Kirui, A.; Ling, Z.; Kang, X.; Widanage, M. C. D.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic resolution of cotton cellulose structure enabled by dynamic nuclear polarization solid-state NMR. *Cellulose* **2019,** *26* (1), 329-339.

43. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of pretreated wood fibers using dynamic nuclear polarization. *J. Phys. Chem. C* **2019,** *123* (50), 30407-30415.

44. Cosgrove, D. J., Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* **2014,** *22*, 122-131.

45. Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; Staehelin, A., *Plant Cell Walls, from Chemistry to Biology*. Garland Science: New York, 2011.

46. Pauly, M.; Albersheim, P.; Darvill, A.; York, W. S., Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J.* **1999,** *20*, 629-639.

47. Atmodjo, M. A.; Hao, Z.; Mohnen, D., Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* **2013,** *64*, 747-79.

48. Zykwinska, A.; Thibault, J. F.; Ralet, M. C., Organization of pectic arabinan and galactan side chains in association with cellulose microfibrils in primary cell walls and related models envisaged. . *J. Exp. Bot.* **2007,** *58*, 1795-1802.

49. Biswal, A. K.; Atmodjo, M. A.; Li, M.; Baxter, D.; Yoo, C. G.; Pu, Y.; Lee, Y. C.; Mazarei, M.; Black, I. M.; Zhang, J. Y.; Ramanna, H.; Bray, A. L.; King, Z. R.; LaFayette, P. R.; Pattathil, S.; Donohoe, B. S.; Mohanty, S. S.; Ryno, D.; Yee, K.; Thompson, O. A.; Rodriguez, M., Jr.; Dumitrache, A.; Natzke, J.; Winkeler, K.; Collins, C.; Yang, X.; Tan, L.; Sykes, R. W.; Gjersing, E.; Ziebell, A.; Turner, G. B.; Decker, S. R.; Hahn, M. G.; Davison, B. H.; Udvardi, M. K.; Mielenz, J. R.; Davis, M. F.; Nelson, R. S.; Parrott, W. A.; Ragauskas, A. J.; Stewart, N. J.; Mohnen, D., Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nat. Biotechnol.* **2018,** *36*, 249-257.

50. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water–Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis thaliana from Polarization Transfer Solid-State NMR. *J. Am. Chem. Soc.* **2014,** *136* (29), 10399-10409.

51. Ridley, B. L.; O'Neill, M. A.; Mohnen, D., Pectins: Structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **2001,** *57* (6), 929-967.

52. Kirui, A.; Du, J.; Zhao, W.; Barnes, W.; Kang, X.; Anderson, C. T.; Xiao, C.; Wang, T., A

pectin Methyltransferase modulates polysaccharide dynamics and interactions in *Arabidopsis* primary cell walls: Evidence from solid-state NMR. *Carbohydr. Polym.* **2021,** *270*, 118370.

53. Wang, T.; Zabotina, O.; Hong, M., Pectin-cellulose interactions in the *Arabidopsis* primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry-Us* **2012,** *51* (49), 9846-9856.

54. Dick-Perez, M.; Zhang, Y. A.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry-Us* **2011,** *50* (6), 989-1000.

55. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis* primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiol.* **2015,** *168* (3), 871-884.

56. Zhang, T.; Tang, H.; Vavylonis, D.; Cosgrove, D. J., Disentangling loosening from softening: insights into primary cell wall structure. *Plant J.* **2019,** *100* (6), 1101-1117.

57. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: a critical review. *Solid State Nucl. Magn. Reson.* **2020,** *107*, 101660.

58. Zhang, Y.; Yu, J.; Wang, X.; Durachko, D. M.; Zhang, S.; Cosgrove, D. J., Molecular insights into the complex mechanics of plant epidermal cell walls. *Science* **2021,** *372* (6543), 706-711.

59. Pelloux, J.; Rusterucci, C.; Mellerowicz, E. J., New insights into pectin methylesterase structure and function. *Trends Plant Sci.* **2007,** *12* (6), 267-277.

60. Du, J.; Anderson, C. T.; Xiao, C., Dynamics of pectic homogalacturonan in cellular morphogenesis and adhesion, wall integrity sensing and plant development. *Nat. Plants* **2022,** *8*, 332-340.

61. Haas, K. T.; Wightman, R.; Meyerowitz, E. M.; Peaucelle, A., Pectin homogalacturonan nanofilament expansion drives morphogenesis in plant epidermal cells. *Science* **2020,** *367*, 1003-1007.

62. Temple, H.; Phyo, P.; Yang, W.; Lyczakowski, J. J.; Echevarria-Poza, A.; Yakunin, I.; Parra-Rojas, J. P.; Terrett, O. M.; Saez-Aguayo, S.; Dupree, R.; Orellana, A.; Hong, M.; Dupree, P., Golgi-localized putative S-adenosyl methionine transporters required for plant cell wall polysaccharide methylation *Nat. Plants* **2022,** *8*, 656-669.

63. Phyo, P.; Gu, Y.; Hong, M., Impact of acidic pH on plant cell wall polysaccharide structure and dynamics: insights into the mechanism of acid growth in plants from solid-state NMR. *Cellulose* **2019,** *26* (1), 291-304.

64. Phyo, P.; Wang, T.; Xiao, C.; Anderson, C. T.; Hong, M., Effects of Pectin Molecular Weight Changes on the Structure, Dynamics, and Polysaccharide Interactions of Primary Cell Walls of Arabidopsis thaliana: Insights from Solid-State NMR. *Biomacromolecules* **2017,** *18* (9), 2937-

2950.

65. Phyo, P.; Wang, T.; Kiemle, S. N.; O'Neill, H.; Pingali, S. V.; Hong, M.; Cosgrove, D. J., Gradients in Wall Mechanics and Polysaccharides along Growing Inflorescence Stems. *Plant Physiol.* **2017,** *175* (4), 1593-1607.

66. Du, J.; Kirui, A.; Huang, S.; Wang, L.; Barnes, W. J.; Kiemle, S. N.; Zheng, Y.; Rui, Y.; Ruan, M.; Qi, S.; Kim, S. H.; Wang, T.; Cosgrove, D. J.; Anderson, C. T.; Xiao, C., Mutations in the Pectin Methyltransferase QUASIMODO2 Influence Cellulose Biosynthesis and Wall Integrity in Arabidopsis. *Plant Cell* **2020,** *32*, 3576-3597.

67. Cosgrove, D. J.; Jarvis, M. C., Comparative structure and biomechanics of plant primary and secondary cell walls. *Front Plant Sci* **2012,** *3*, 204.

68. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *Proc. Natl. Acad. Sci. USA* **2011,** *108* (47), E1195-E1203.

69. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry-Us* **2015,** *54* (14), 2335-2345.

70. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

71. Grantham, N. J.; Wurman-Rodrich, J.; Terrett, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **2017,** *3* (11), 859-865.

72. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 347.

73. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

74. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 4978.

75. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose-xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* **2020,** *11* (1), 6081.

76. Busse-Wicher, M.; Gomes, T. C. F.; Tryfona, T.; Nikolovski, N.; Stott, K.; Grantham, N. J.; Bolam, D. N.; Skaf, M. S.; Dupree, P., The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of Arabidopsis thaliana. *Plant J.* **2014,** *79* (3), 492-506.

77. Chakraborty, A.; Deligey, F.; Quach, J.; Mentink-Vigier, F.; Wang, P.; Wang, T., Biomolecular complex viewed by dynamic nuclear polarization solid-state NMR spectroscopy. *Biochem. Soc. Trans.* **2020,** *48*, 1089-1099.

78. Addison, B.; Stengel, D.; Bharadwaj, V. S.; Happs, R. M.; Doeppke, C.; Wang, T.; Bomble, Y. J.; Holland, G. P.; Harman-Ware, A. E., Selective One-Dimensional 13C–13C Spin-Diffusion Solid-State Nuclear Magnetic Resonance Methods to Probe Spatial Arrangements in Biopolymers Including Plant Cell Walls, Peptides, and Spider Silk. *J. Phys. Chem. B* **2020,** *124*, 9870-9883.

79. Cresswell, R.; Dupree, R.; Brown, S. P.; Pereira, C. S.; Skaf, M. S.; Sorieul, M.; Dupree, P.; Hill, S., Importance of Water in Maintaining Softwood Secondary Cell Wall Nanostructure. *Biomacromolecules* **2021,** *22*, 4669-4680.

80. Li, H.; Kang, X.; Yang, M.; Kasseney, B. D.; Zhou, X.; Liang, S.; Zhng, X.; J., W.; Yu, B.; Liu, N.; Zhao, Y.; Mo, J.; Currie, C. R.; Ralph, J.; Yelle, D., Molecular Insight Into the Evolution of Woody Plant Decay in the Gut of Termites. *Sci. Ad.* **2023,** *9*, eadg1258.

81. Rytioja, J.; Hilden, K.; Yuzon, J.; Hatakka, A.; de Vries, R. P.; Makela, M. R., Plant-Polysaccharide-Degrading Enzymes from Basidiomycetes. *Microbiol. Mol. Biol. Rev.* **2014,** *78*, 614-649.

82. Chatterjee, S.; Prados-Rosales, R.; Itin, B.; Casadevall, A.; Stark, R. E., Solid-state NMR Reveals the Carbon-based Molecular Architecture of *Cryptococcus neoformans* Fungal Eumelanins in the Cell Wall. *J. Biol. Chem.* **2015,** *290* (22), 13779-13790.

83. Chatterjee, S.; Prados-Rosales, R.; Frases, S.; Itin, B.; Casadevall, A.; Stark, R. E., Using Solid-State NMR To Monitor the Molecular Consequences of Cryptococcus neoformans Melanization with Different Catecholamine Precursors. *Biochemistry-Us* **2012,** *51* (31), 6080-6088.

84. Zhong, J. Y.; Frases, S.; Wang, H.; Casadevall, A.; Stark, R. E., Following fungal melanin biosynthesis with solid-state NMR: Biopolyrner molecular structures and possible connections to cell-wall polysaccharides. *Biochemistry* **2008,** *47* (16), 4701-4710.

85. Baker, R. P.; Chrissian, C.; Stark, R. E.; Casadevall, A., Cryptococcus neoformans melanization incorporates multiple catecholamines to produce polytypic melanin. *J. Biol. Chem.* **2022,** *298*, 101519.

86. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

87. Latgé, J.-P., The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.* **2007,** *66*

(2), 279-290.

88. Wagener, J.; Striegler, K.; Wagener, N., α- and β-1,3-Glucan Synthesis and Remodeling. In *The Fungal Cell Wall: An Armour and a Weapon for Human Fungal Pathogens*, Latgé, J. P., Ed. Springer, Cham.: 2020; Vol. 425, pp 53-82.

89. Wagener, J.; Loiko, V., Recent Insights into the Paradoxical Effect of Echinocandins. *J. Fungi* **2018,** *4*, 5.

90. Latgé, J. P.; Chamilos, G., Aspergillus fumigatus and Aspergillosis in 2019. *Clin. Microbiol. Rev.* **2019,** *33* (1), e00140-18.

91. Chakraborty, A.; Fernando, L. D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latgé, J. P.; Wang, T., A molecular vision of fungal cell wall organization by functional genomics and solid-state NMR. *Nat. Commun.* **2021,** *12*, 6346.

92. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

93. Wang, T.; Hong, M., Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **2016,** *67*, 503-514.

94. Fernando, L. D.; Dickwella Widanage, M. C.; Penfield, J.; Lipton, A. S.; Washton, N.; Latge, J.-P.; Wang, P.; Zhang, L.; Wang, T., Structural Polymorphism of Chitin and Chitosan in Fungal Cell Walls from Solid-State NMR and Principal Component Analysis. *Front. Mol. Biosci.* **2021,** *8*, 727053.

95. Heux, L.; Brugnerotto, J.; Desbrieres, J.; Versali, M. F.; Rinaudo, M., Solid state NMR for determination of degree of acetylation of chitin and chitosan. *Biomacromolecules* **2000,** *1* (4), 746-751.

96. Tanner, S. F.; Chanzy, H.; Vincendon, M.; Roux, J. C.; Gaill, F., High-Resolution Solid-State C-13 Nuclear-Magnetic-Resonance Study of Chitin. *Macromolecules* **1990,** *23* (15), 3576-3583.

97. Latgé, J. P.; Wang, T., Modern Biophysics Redefines Our Understanding of Fungal Cell Wall Structure, Complexity, and Dynamics. *mBio* **2022,** *13*, e01145-22.

98. Lamon, G.; Lends, A.; Valsecchi, I.; Wong, S. S. W.; Dupres, V.; Lafont, F.; Tolchard, J.; Schmitt, C.; Mallet, A.; Grelard, A.; Morvan, E.; Dufourc, E.; Habenstein, B.; Guijarro, J. I.; Aimanianda, V.; Loquet, A., Solid-state NMR molecular snapshots of Aspergillus fumigatus cell wall architecture during a conidial morphotype transition. *Proc. Natl. Acad. Sci. USA* **2023,** *120*, e2212003120.

99. Fernando, L. D.; Dickwella Widanage, M. C.; Shekar, C. S.; Mentink-Vigier, F.; Wang, P.; Wi, S.; Wang, T., Solid-state NMR analysis of unlabeled fungal cell walls from Aspergillus and Candida species. *J. Struct. Biol. X* **2022,** *6*, 100070.

100. Chrissian, C.; C.P.-C., L.; Camacho, E.; Casadevall, A.; Neiman, A. M.; Stark, R. E., Unconventional Constituents and Shared Molecular Architecture of the Melanized Cell Wall of C. neoformans and Spore Wall of S. cerevisiae. *J. Fungi* **2020,** *6*, 329.

101. Ehren, H. L.; Appels, F. V. W.; Houben, K.; Renault, M. A. M.; Wösten, H. A. B.; Baldus, M., Characterization of the cell wall of a mushroom forming fungus at atomic resolution using solid-state NMR spectroscopy. *Cell Surf.* **2020,** *6*, 100046.

102. Bahri, S.; Safeer, A.; Adler, A.; Smedes, H.; van Ingen, H.; Baldus, M., 1H-detected characterization of carbon–carbon networks in highly flexible protonated biomolecules using MAS NMR. *J. Biomol. NMR* **2023,** *77*, 111-119.

103. Kleijburg, F. E. L.; Safeer, A.; Baldus, M.; Wosten, H. A. B., Binding of micro-nutrients to the cell wall of the fungus Schizophyllum commune. *Cell Surf.* **2023,** *10*, 100108.

104. van Dijk, A.; Molhoek, E. M.; Veldhuizen, E. J. A.; Bokhoven, J. L. M. T.; Wagendrop, E.; Bikker, F.; Haagsman, H. P., Identification of chicken cathelicidin-2 core elements involved in antibacterial and immunomodulatory activities *Mol. Immunol.* **2009,** *46*, 2465-2473.

105. Ndukwe, I. E.; Black, I.; Heiss, C.; Azadi, P., Evaluating the Utility of Permethylated Polysaccharide Solution NMR Data for Characterization of Insoluble Plant Cell Wall Polysaccharides. *Anal. Chem.* **2020,** *92*, 13221-13228.

106. Mansfield, S. D.; Kim, H.; Lu, F. C.; Ralph, J., Whole plant cell wall characterization using solution-state 2D NMR. *Nat Protoc* **2012,** *7* (9), 1579-1589.

107. Kim, H.; Ralph, J., Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d6/pyridine-d5. *Org. Biomol. Chem.* **2010,** *8*, 576-591.

108. Wang, W.; Cai, J.; Xu, Z.; Zhang, Y.; Niu, F.; Gao, M.; Jun, Y., Structural characteristics of plant cell wall elucidated by solution-state 2D NMR spectroscopy with an optimized procedure. *Green Process. Synth.* **2020,** *9*, 650-653.

109. Reichhardt, C.; Joubert, L. M.; Clemons, K. V.; Stevens, D. A.; Cegelski, L., Integration of electron microscopy and solid-state NMR analysis for new views and compositional parameters of Aspergillus fumigatus biofilms. *Med. Mycol.* **2019,** *57*, S239-S244.

110. Deligey, F.; Frank, M. A.; Cho, S. H.; Kirui, A.; Mentink-Vigier, F.; Swulius, M. T.; Nixon, T.; Wang, T., Structure of In Vitro-Synthesized Cellulose Fibrils Viewed by Cryo-Electron Tomography and 13C Natural-Abundance Dynamic Nuclear Polarization Solid-State NMR. *Biomacromolecules* **2022,** *23*, 2290-2301.

111. Nicolas, W. J.; Fabler, F.; Dutka, P.; Schur, F. K. M.; Jensen, G.; Meyerowitz, E., Cryo-electron tomography of the onion cell wall shows bimodally oriented cellulose fibers and reticulated homogalacturonan networks. *Curr. Biol.* **2022,** *32*, 2375-2389.

112. Metskas, L. A.; Wilfong, R.; Jensen, G., Subtomogram averaging for biophysical analysis and supramolecular context. *J. Struct. Biol. X* **2022,** *6*, 100076.

113. Harapin, J.; Eibauer, M.; Medalia, O., Structural Analysis of Supramolecular Assemblies by Cryo-Electron Tomography. *Structure* **2013,** *21*, 1522-1530.

114. Jimenez-Ortigosa, C.; Jiang, J.; Chen, M.; Kuang, X.; Healey, K. R.; Castellano, P.; Boparai, N.; Ludtke, S. J.; Perlin, D. S.; Dai, W., Preliminary Structural Elucidation of β-(1,3)-glucan Synthase from Candida glabrata Using Cryo-Electron Tomography *J. Fungi* **2021,** *7*, 120.

115. Hu, X.; Yang, P.; Chai, C.; Liu, J.; Sun, H.; Wu, Y.; Zhang, M.; Zhang, M.; Liu, X.; Yu, H., Preliminary Structural Elucidation of β-(1,3)-glucan Synthase from Candida glabrata Using Cryo-Electron Tomography *Nature* **2023,** *616*, 190-198.

116. Purushotham, P.; Ho, R.; Zimmer, J., Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science* **2020,** *369*, 1089-1094.

117. Giummarella, N.; Pu, Y.; Ragauskas, A. J.; Lawoko, M., A critical review on the analysis of lignin carbohydrate bonds. *Green Chem.* **2019,** *21*, 1573-1595.

118. Fernando, L. D.; Perez-Llano, Y.; Dickwella Widanage, M. C.; Martinez-Avila, L.; Lipton, A. S.; Gunde-Cimerman, N.; Latge, J. P.; Bastista-Garcia, R. A.; Wang, T., Structural Organization of the Cell Wall of Halophilic Fungi. *Preprint at bioRxiv* **2023**, DOI: 10.1101/2023.04.15.537024.

119. Romaniuk, J. A. H.; Cegelski, L., Bacterial cell wall composition and the influence of antibiotics by cell-wall and whole-cell NMR. *Philos. Trans. R. Soc. B* **2015,** *370* (1679), 20150024.

120. Reichhardt, C.; Lim, J. Y.; Rice, D.; Fong, J. N.; Cegelski, L., Structure and Function of Bacterial Biofilms by Solid-State NMR. *Biophys. J.* **2014,** *106* (2), 192a.

121. Poulhazan, A.; Dickwella Widanage, M. C.; Muszynski, A.; Arnold, A.; Warschawski, D. E.; Azadi, P.; Marcotte, I.; Wang, T., Identification and Quantification of Glycans in Whole Cells: Architecture of Microalgal Polysaccharides Described by Solid-State Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **2021,** *143*, 19374-19388.

122. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. *J. Biomol. NMR* **2018,** *70* (2), 123-131.

CHAPTER 5: A THREE-DIMENSIONAL 13C-13C-13C CORRELATION EXPERIMENT FOR HIGH-RESOLUTION ANALYSIS OF COMPLEX CARBOHYDRATES USING SOLID-STATE NMR

Research paper reprinted with permission: *J. Magn. Reson.* 336, 107148 (2022)

Authors: S. Chandra Shekar, Wancheng Zhao, Liyanage D. Fernando, Ivan Hung, and Tuo Wang

5.1 Abstract

Complex carbohydrates are the key components of the protective cell walls of microbial pathogens and the bioenergy reservoir in plants and algae. Structural characterization of these polymorphic molecules requires assistance from multidimensional ^{13}C correlation approaches. To facilitate the analysis of carbohydrate structure using solid-state NMR, we present a three-dimensional (3D) ^{13}C-^{13}C-^{13}C experiment that includes a double-quantum (DQ) dimension and is thus free of the cube's body diagonal. The enhanced resolution supports the unambiguous resonance assignment of many polysaccharides in plant and fungal cell walls using uniformly ^{13}C-labeled cells of spruce and *Aspergillus fumigatus*. Long-range structural restraints were effectively obtained to revisit our understanding of the spatial organization of plant cellulose microfibrils. The method is widely applicable to the investigations of cellular carbohydrates and carbon-based biomaterials.

5.2 Introduction

Three-dimensional (3D) experiments are a standard approach used in biomolecular solid-state NMR studies but have rarely been applied to biomaterials[1-3]. Baldus, Rienstra, and Hong have established a collection of 3D ^{13}C-^{13}C-^{13}C (CCC) correlation experiments that flexibly combine the detection of single-quantum (SQ) and double-quantum (DQ) ^{13}C chemical shifts with various polarization transfer pathways and frequency filters to fulfill the needs for resonance assignment

and obtaining long-range structural restraints[4-7]. Such 3D correlation experiments also have the potential of assisting the identification of many coexisting polymers and understanding their spatial association in carbon-based biocomposites[8-10].

Indeed, recent NMR studies of plant cell walls have already benefited from the SQ-SQ-SQ CCC experiment (**Figure 5.1a**). In 2011, the first 3D CCC spectrum of plant materials was collected on uniformly ^{13}C-labeled primary cell walls of *Arabidopsis* seedlings, reporting numerous intermolecular cross peaks between cellulose and pectin[9]. This was the first set of consolidated molecular-level data that support the putative interactions between these two biopolymers. The results also led to a revised concept that the hemicellulose xyloglucan only interacts with cellulose at limited spots[9], which was later combined with biochemical evidence to build new conceptual models of primary plant cell walls[11-13]. In 2015, this experiment was extended to the investigation of secondary plant cell walls[10]. Demonstrated on the mature stems of the same model plant (*Arabidopsis thaliana*), the high-resolution data successfully revealed the conformational distribution of the hemicellulose xylan and the interconnectivity of different cellulose domains[10].

Currently, carbohydrate solid-state NMR has to rely on carbon-13 due to the highly restricted choice of NMR-active nuclei. Nitrogen-15 is widely used in protein studies but is only applicable to a small number of nitrogenated sugars present in nature. Proton detection and oxygen-17 studies have demonstrated their effectiveness in protein and material research[14-17], but both are still in the exploratory stage for carbohydrate studies[18-20]. Therefore, there is an urgent need for more versatile CCC experiments to assist the structural analysis of carbohydrate polymers.

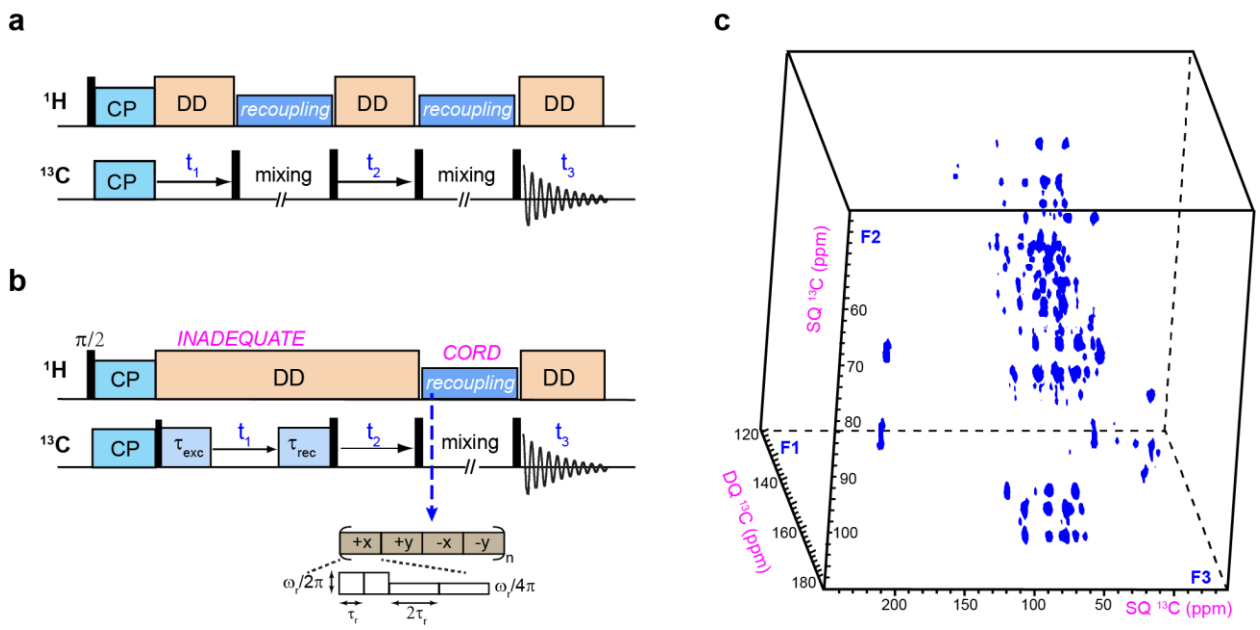

Figure 5.1 3D CCC experiments for characterizing carbon-rich biosolids. **a,** 3D ^{13}C-^{13}C-^{13}C correlation experiments relying on two ^{13}C-^{13}C mixing periods, for example, DARR/PDSD periods. **b,** 3D ^{13}C-^{13}C-^{13}C (DQ-SQ-SQ) experiment constructed as an INADEQUATE-CORD sequence. The INADEQUATE block is presented in a simplified format as it could be constructed using either J- or dipolar-based pulse sequence. The pulse program used for this study is provided in the **Supplementary Data**. **c,** Representative 3D INADEQUATE-CORD solid-state NMR spectrum collected on living cells of *A. fumigatus* at 800 MHz. The three dimensions (F1, F2, F3) report DQ, SQ, and SQ ^{13}C chemical shifts, respectively.

Here we adapt the 3D DQ-SQ-SQ CCC experiment, a scheme first demonstrated by Baldus and coworkers[4], to investigate the structure and packing of carbohydrate polymers in cellular environments. After the J-INADEQUATE module[21, 22], a COmbined $R2_{n\upsilon}$-Driven (CORD)[23, 24] block was implemented (**Figure 5.1b**). The consideration is that CORD is a broadband homonuclear correlation sequence that even functions well under fast MAS conditions; therefore, it should remain effective for future applications on high-field magnets under fast spinning. Also,

CORD needs very low recoupling power as the radiofrequency (rf) field strength is only needed on the ^{1}H channel and is matching only half of the spinning frequency during the majority (two-thirds) of the mixing period. This can effectively reduce the detrimental effect of rf power on sensitive biological samples such as the living cells of various organisms.

The unprecedented resolution provided by this 3D CCC (DQ-SQ-SQ) experiment allows us to systematically document the ^{13}C chemical shift of various glucans and chitin molecules in the cell walls of a pathogenic fungus *Aspergillus fumigatus*[25, 26], using its living cells. Signal multiplicity has been observed for chitin and α-1,3-glucan; the former is attributed to the structural complexity of sugar units when placed in the hydrogen-bonded chitin microfibrils and the latter is correlated with the functional diversity of α-1,3-glucans. The most rewarding improvement is on plant biomass, the research of which has long been hampered by insufficient resolution. We have resolved the signals of previously ambiguous carbon sites and further evaluated how different glucose conformers are packed in the bundled cellulose microfibrils. This experiment has shown its usefulness in the investigations of carbohydrate polymers and should facilitate the ongoing efforts in understanding the carbohydrate structure and assembly in algae, bacteria, fungi, plants, and human cells, as well as biomimetic materials[27-32]. The pulse program has been provided in the **Supplementary Data** to expedite future investigations.

5.3 Results and Discussion

The 3D INADEQUATE-CORD spectrum of the 3-day-old *A. fumigatus* mycelium showed remarkable dispersion of the carbon peaks (**Figure 5.1c**). The sample contains the living cells of this fungus, which should be a highly complex and challenging biosystem for NMR investigation. However, the outstanding spectral resolution (typically 0.3-0.7 ppm for ^{13}C linewidths as reported recently)[33, 34] allowed us to use it as a model system for cellular carbohydrates. The conventional

INADEQUATE spectrum reports the signals from directly bonded carbons (**Figure 5.2a**), and all the spin pairs of *A. fumigatus* carbohydrates were marked using their DQ chemical shifts. As ^{1}H-^{13}C cross polarization was used for the creation of initial magnetization, the signals detected here mainly originated from four carbohydrate polymers found in the rigid portion of fungal cell walls (**Figure 5.2b**). These molecules include chitin, α-1,3-glucan, β-1,3-glucan (occasionally with β-1,6-linkage for branching), and the mixed linked β-1,3/1,4-glucan, which is a terminal segment of the diversely linked β-glucan network[35-37],

The implementation of a 53-ms CORD mixing period after the INADEQUATE block allowed the ^{13}C magnetization to be transferred to all carbon sites in a molecule that showed up at a given double-quantum chemical shift (**Figure 5.3a**). For example, the β-1,3-glucan carbon 1 and carbon 2 (B1 and B2) have single-quantum (SQ) chemical shifts of 103.6 ppm and 74.4 ppm, respectively. Therefore, in the INADEQUATE spectrum, their signals showed up at the double-quantum (DQ) chemical shift of 178 ppm, which is the sum of the two SQ chemical shifts. When CORD is implemented, the other carbons of β-1,3-glucan (B3, B4, B5, and B6) also appeared in the same DQ line, with well-resolved signals for resonance assignment.

Such spectral feature presents a unique advantage over the conventional 2D correlation spectrum. Both the DQ-SQ correlation (such as INADEQUATE) and SQ-SQ correlation (such as DARR and CORD) spectra often have highly crowded regions, where the carbon connectivity cannot be explicitly tracked. For example, the carbon 1 signals of chitin and β-1,3-glucan are almost identical; therefore, their C1-CX (CX denotes other carbons) cross peaks are severely overlapped signals in the CORD spectrum (**Figure 5.3b**). However, their C1-C2 spin pairs have distinct DQ chemical shifts: 159 ppm for chitin C1-C2 and 178 ppm for β-1,3-glucan C1-C2 as shown in **Figure 5.3a**. Therefore, we can unambiguously resolve and assign all their carbon signals

in the collapsed INADEQUATE-CORD spectrum, which is free of diagonal and has better dispersion of signals as benefited from the implementation of a DQ dimension.

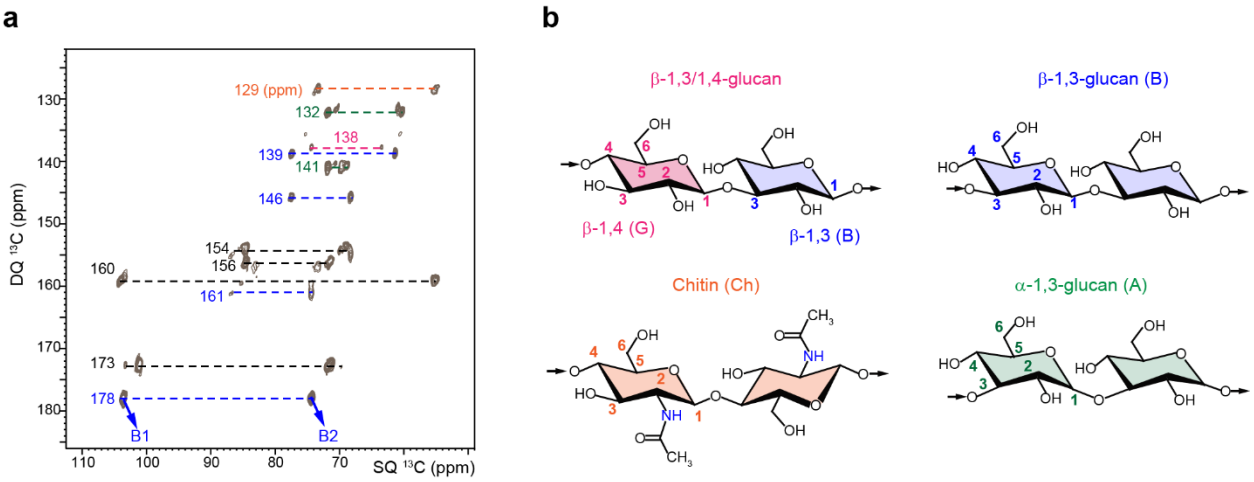

Figure 5.2 2D spectra of fungal cell wall carbohydrates. **a,** 2D ^{13}C-^{13}C INADEQUATE (DQ-SQ) spectrum of *A. fumigatus* cells. Key cross sections are marked using the DQ ^{13}C chemical shifts. **b,** Representative *A. fumigatus* carbohydrates with the NMR abbreviations labeled. All carbohydrates are color-coded, and the color scheme is used for all fungal figures.

When a DQ chemical shift contains only a single pair of carbons, all the carbon sites in the same molecule can be well resolved in the third dimension (**Figure 5.4a** and **Figure S5.1**). For example, in the F2-F3 plane of F1 = 178 ppm, both B1 and B2 will have cross peaks with the other carbons in β-1,3-glucan. Similar spectral patterns were also observed in other F2-F3 panels of this molecule, regardless of the F1 DQ chemical shifts, which provided confidence to the resonance assignment. The β-1,4-linked glucose units only account for a minor portion of the β-glucan matrix because the β-1,3/1,4-glucan constitute approximately 10% of all β-glucans[38]. It is not surprising that β-1,4-linked glucose residues have been particularly difficult to track using NMR. However, its ^{13}C signals can be assigned using the F2-F3 plane extracted at F1 = 138 ppm.

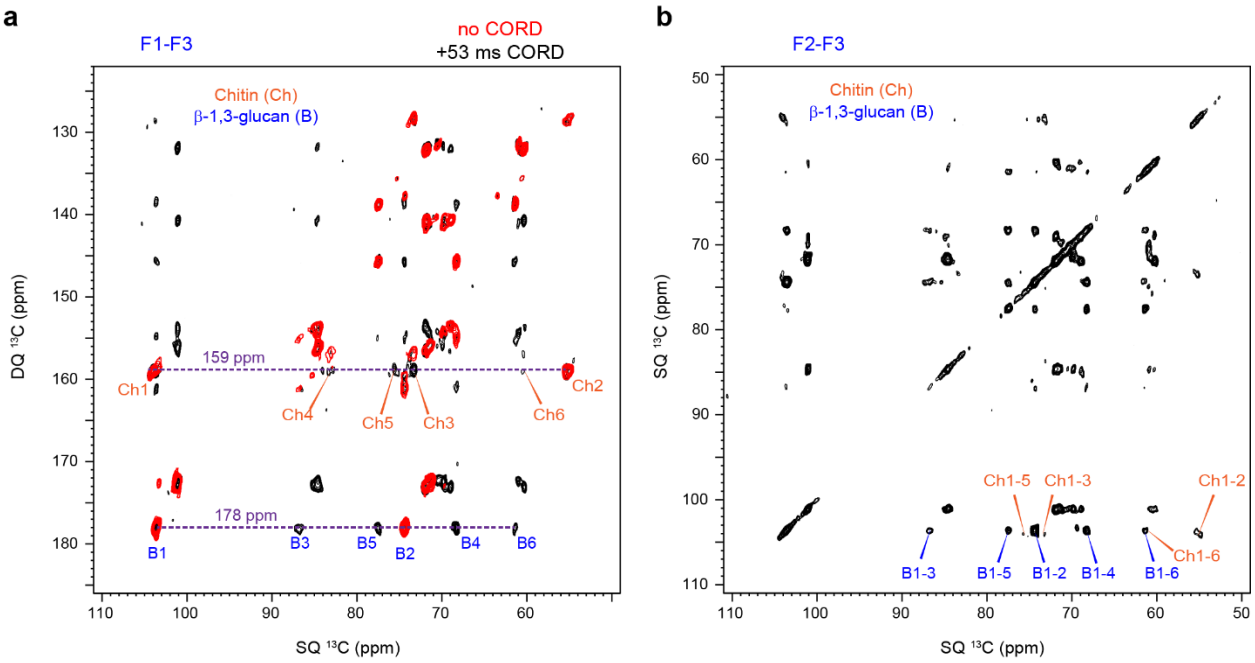

Figure 5.3 Collapsed 2D spectra of the 3D INADEQUATE-CORD experiment. The 2D spectra plotted in black represent all the signals of the 3D INADEQUATE-CORD when being projected to **a,** the F1-F3 dimensions and **b,** the F2-F3 dimensions. For comparison, a conventional INADEQUATE spectrum (red) is overlaid with the F1-F3 plane, where the extra signals are from multi-bond intramolecular correlations within each molecule. The representative signals of chitin (Ch) and β-1,3-glucan (B) are labeled. For example, Ch1 denotes chitin carbon 1.

Both α-1,3-glucan and chitin showed peak multiplicity, with multiple sets of signals for each carbon. This observation further correlates with the functional diversity and structural complexity of these biomolecules. Recent studies have revealed that α-1,3-glucan is widely distributed in all functional domains of fungal cell walls, including the rigid and mobile portions (from the physical perspective) of both the alkali-soluble and alkali-insoluble fractions (from the chemical view)[33, 34]. Differently, chitin has complex structural environments in the hydrogen-bonded microfibrils, which has been recently revealed by solid-state NMR and principal component analysis[39].

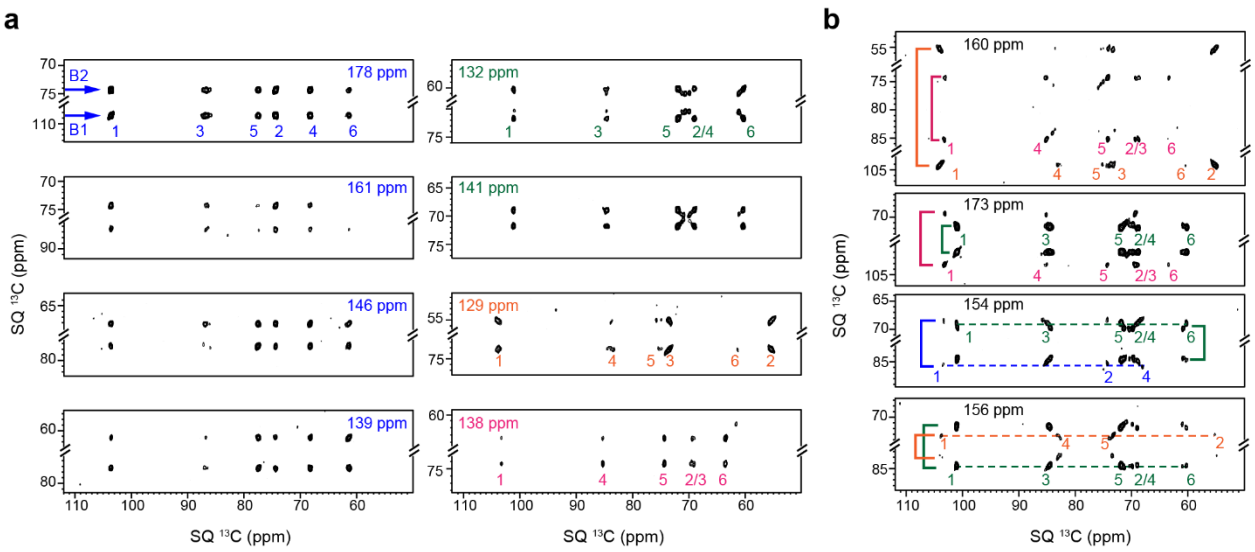

Figure 5.4 2D planes of the 3D spectra resolving signals of fungal cell wall carbohydrates. The double quantum ^{13}C chemical shifts (in ppm) are marked in panels (a) and (b) to match the signals in Figure 5.2a. **a,** Well-resolved F2-F3 planes extracted from 3D INADEQUATE-CORD (DQ-SQ-SQ) spectrum of the *A. fumigatus* sample. The left panels show signals of β-1,3-linked glucans or glucose units (blue) while the right panel shows α-1,3-glucan (green), chitin (orange), and β-1,4-linked glucose residues. The numbers (1-6) indicate the carbon numbers in each carbohydrate molecule as presented in Figure 5.2b. Two arrows are used to annotate the lines of peaks for β-1,3-glucan carbon 1 (B1) and carbon 2 (B2), with correlations to each of the six carbons in this molecule. **b,** F2-F3 planes with multiple molecules observed in each 2D plane. The full views of these panels are presented in **Figure S5.1**.

It is expected that a single DQ chemical shift might represent multiple pairs of carbon. For example, the 160 ppm DQ chemical shift is the sum of 104 ppm and 56 ppm from chitin C1 and C2, as well as 85 ppm and 75 ppm from the C4 and C5 of β-1,4-linked glucose units (**Figure 5.4b**). All the ^{13}C signals are well resolved even in the mixed planes, for example, those with concurrent contributions from chitin and β-1,3/1,4-glucan (F1 = 160 ppm), β-1,3/1,4-glucan and α-1,3-glucan

(F1 = 173 ppm), β-1,3-glucan and α-1,3-glucan (F1 = 154 ppm), as well as chitin and α-1,3-glucan (F1 = 156 ppm). It is obvious that the unprecedented resolution offered by the 3D should facilitate the NMR studies of carbohydrate-rich biomaterials.

This 3D CCC experiment is particularly useful for the analysis of highly complex biosystems, such as plant secondary cell walls, whose spectra are much more congested than those of the fungal cells. Even the collapsed 2D INADEQUATE-CORD spectrum measured on a spruce stem sample failed to provide sufficient resolution for resolving a large number of carbohydrate units and conformers present in the cellulose microfibrils and matrix polysaccharides (**Figure S5.2**). However, the F2-F3 planes extracted at F1 = 138 ppm (cellulose carbon 5 and 6) and F1 = 178 ppm (cellulose carbon 1 and carbon 2) efficiently separated the many carbon peaks in the glucan chains residing on the surface or the internal domains of cellulose microfibrils (**Figure 5.5a**). Matrix polysaccharides exhibited even better resolution (F1 = 174 ppm) as benefited from their partial mobility.

We further spotted the signals from the hydrophobic (s^g) and hydrophilic (s^f) chains of cellulose surface, the deeply embedded core chains (i^c), and an intermediate layer (i^a and i^b) bridging the core and the surface (**Figure 5.5b**). These glucose units have been recently identified using high-resolution 2D ^{13}C-^{13}C correlation spectra[40], with their hydroxymethyl conformations restrained by ^{1}H-^{1}H distance measurements[41]. However, the peak congestion of these glucose units caused ambiguity in the previous efforts to complete the resonance assignment[40]. The resolution presented by the well-isolated signals, such as these C6-C1 cross peaks, has never been achieved before. We can even resolve the carbon 2 signals of internal and surface chains that are apart by less than 0.5 ppm, which is below the ^{13}C FWHM of individual cellulose peaks (0.7-1.0 ppm).

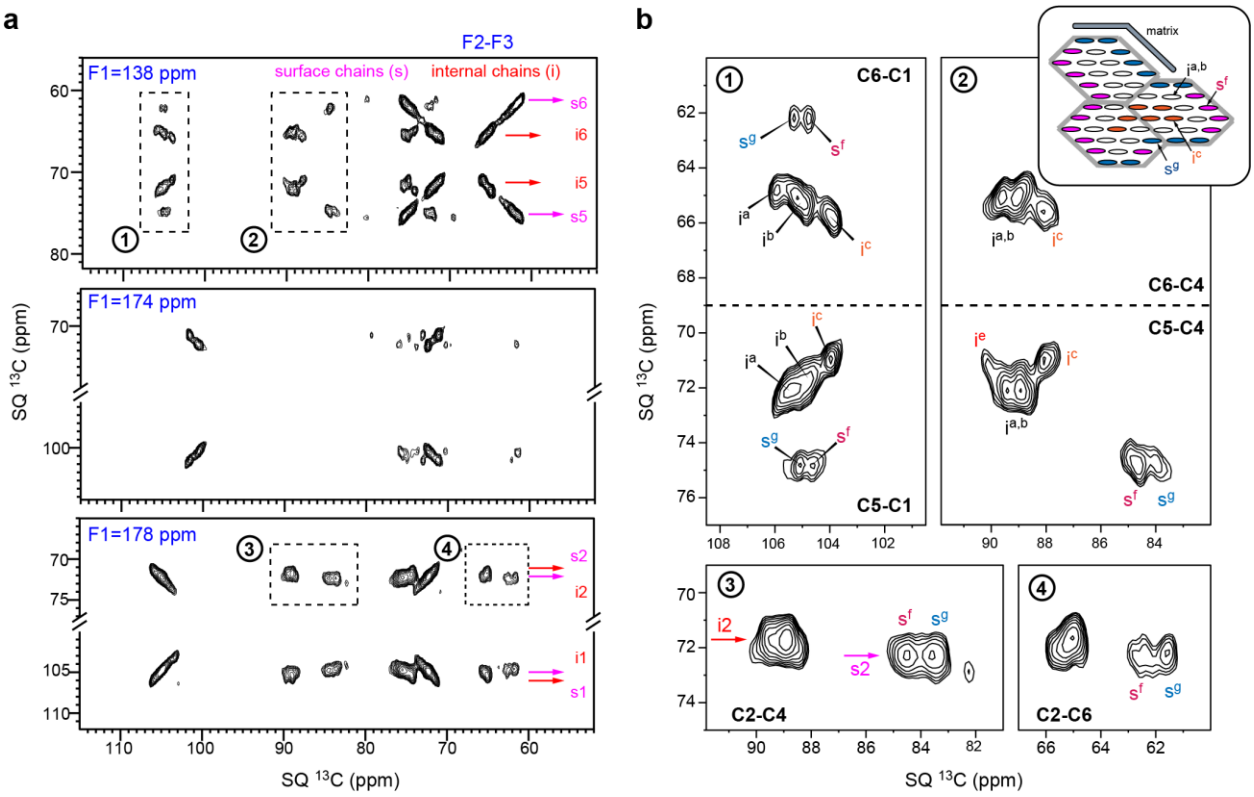

Figure 5.5 2D F2-F3 planes of spruce extracted from the 3D spectrum. **a,** Representative planes of spruce cellulose (F1 = 138 ppm and 178 ppm) and matrix polysaccharides (F1 = 174 ppm). The rows of internal (i) and surface (s) chains are shown using arrows. **b,** Zoom-in views of the numbered regions boxed by dashlines in panel (a). A bundle formed by three elementary cellulose microfibrils is shown to indicate the different packing environment for internal cellulose (types a, b, c) and surface chains (types f and g).

Although the spectrum is diagonal-free considering the 3D cube's body as shown in **Figure 5.1c**, the diagonals and anti-diagonals of the 2D F2-F3 planes are clearly visible (**Figure 5.6**), arising only from the correlated spins in F3-F2 plane at a given F1 (DQ) frequency, as has been noted by Baldus and co-workers[4]. For example, **Figure 5.6** shows the resonance assignment of all signals observed in the 2D plane extracted at F1 = 138 ppm, which is the DQ chemical shifts of interior cellulose C5-C6 and surface cellulose C5-C6 carbon pairs. Therefore, diagonal signals can

only be observed for C6-C6 and C5-C5 signals in the type-f and type-g surface chains as well as different types of interior residues. However, these diagonal signals differ from those of the conventional 2D SQ-SQ correlation spectra where all carbon sites present in the biosystem have their corresponding diagonal peaks. The diagonal signals observed in a 2D F2-F3 plane are more discrete as they only originate from the carbons selected by the DQ chemical shift (for example, only cellulose C5 and C6 here). Many other carbon sites in matrix polysaccharides have similar SQ chemical shifts but they are filtered out by their distinct DQ chemical shifts. Also, the diagonal does not span across the whole spectral width (for example, 10-180 ppm for the ^{13}C spectrum of plant cell walls) as typically observed in 2D correlation spectra. The DQ selection has clearly alleviated the spectral crowding issue.

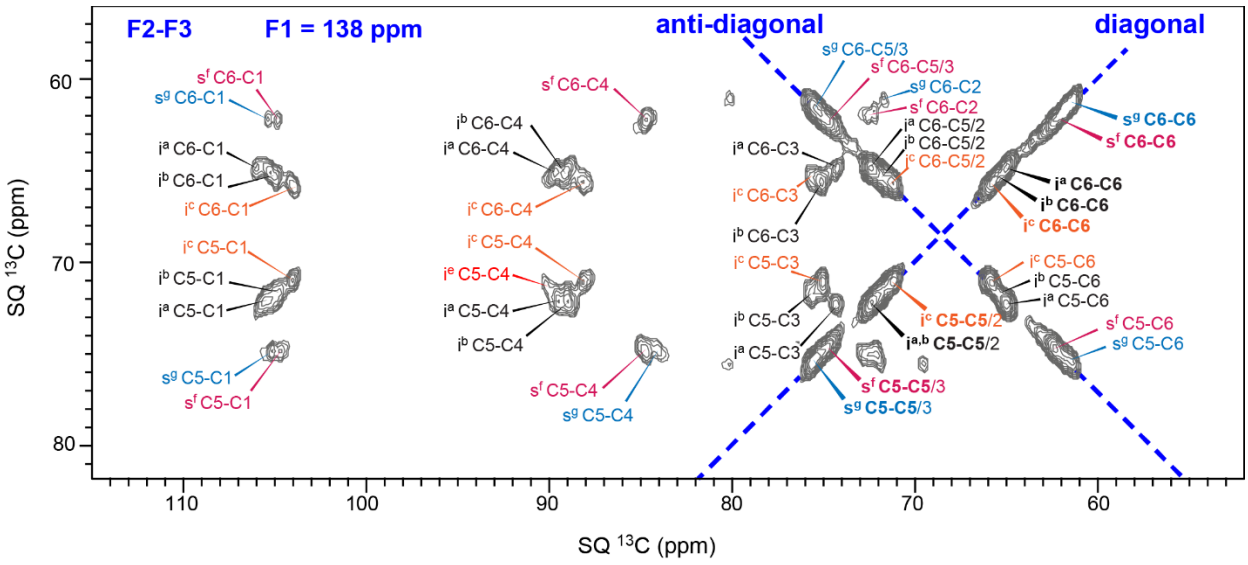

Figure 5.6 Diagonal and antidiagonal signals of F2-F3 planes. Resonance assignment is shown for a F2-F3 plane (F1 = 138 ppm; cellulose C5-C6) extracted from the 3D spectrum of the spruce sample. Dashlines show the positions of the diagonal and anti-diagonal cross peaks. The assignments of the diagonal peaks are highlighted in bold.

A more promising application of this 3D experiment is to identify long-range correlations for structural determination. This was evidenced by the numerous intermolecular cross peaks reported by the F2-F3 plane (F1 = 138 ppm) of the 3D spectrum measured with a 300-ms long CORD mixing period. Site-specific and conformer-specific information of polymer packing can be obtained. For example, instead of simply reporting a cross peak between surface cellulose carbon 6 and interior cellulose carbon 4 (s6-i4), we identified that the type-a and type-b glucose units ($i^{a,b}$) were responsible for contacting the surface chains (**Figure 5.7**). This observation perfectly aligns with the structural scheme presented in **Figure 5.5b**, as $i^{a,b}$ represents the layer of glucan chains that are directly underneath the surface chains. Always, two types of surface chains, s^f and s^g, were simultaneously involved in the cross peaks with internal chains (more specifically, the $i^{a,b}$ chains), with comparable intensities. This is understandable as these two types of surface chains, regardless of their capability of retaining water molecules, are both tightly associated with the internal chains. While only a single F2-F3 plane is shown here to demonstrate the technical feasibility, the full long-range correlation 3D spectrum carries a rich pool of data that could possibly revise our understanding of the structure and bundling of cellulose, as well as its interactions with matrix polymers.

Both the fungal and plant samples studied here are never-dried materials. The native hydration level is also key to the spectral resolution because the hydration state is particularly important for the mobile and non-crystalline components that rely on motional dynamics to narrow the lines. Fortunately, most carbohydrates could efficiently retain their associations with water molecules. Three recent studies conducted on the primary and secondary plant cell walls as well as the microalgal cells have demonstrated that the ^{13}C resolution of cell wall carbohydrates is largely independent of the hydration history of the sample and can be fully restored upon

rehydration[27, 42, 43]. However, irreversible changes might occur to the nanostructure of water-stabilized biopolymer interface upon dehydration[44, 45]. Therefore, the sample preparation procedures, in addition to the spectroscopic methods, should be carefully designed and tested to ensure sufficient resolution for investigating carbohydrate-rich biomaterials.

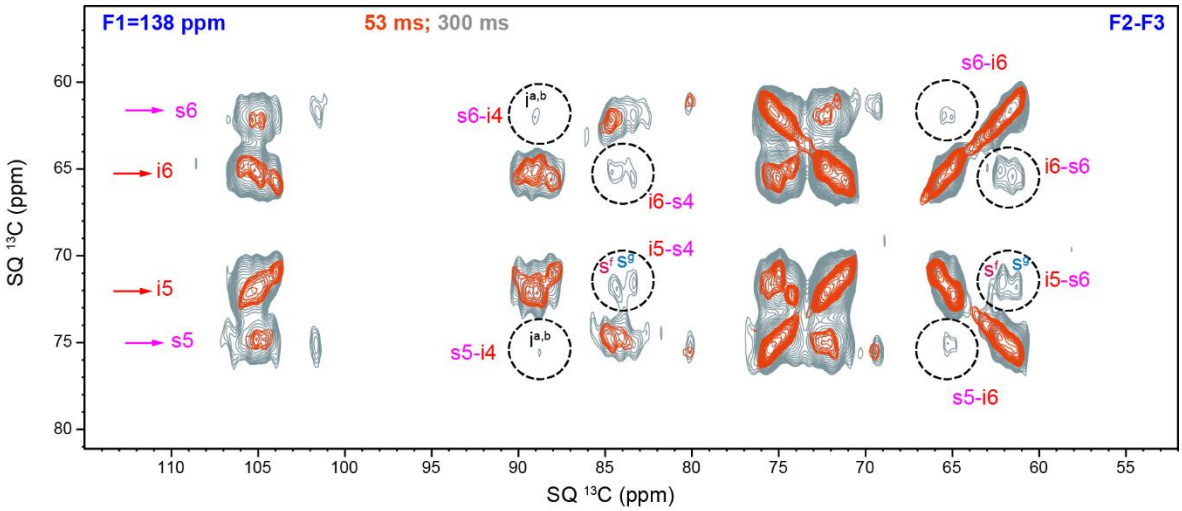

Figure 5.7 Long-range intermolecular interactions in spruce cellulose. The comparison of F2-F3 planes (F1 = 138 ppm) showed numerous long-range intermolecular cross peaks. The key interactions between internal and surface chains of cellulose are selectively highlighted using dash line circles. Within each circle, the signals of different types of glucose units can be unambiguously resolved, providing site- and conformer-specific information on the chain packing of cellulose.

5.4 Materials and Methods

5.4.1 Sample Preparation

Three uniformly ^{13}C-labeled samples were used in this study, including the intact cells of *A. fumigatus* (also with uniform ^{15}N-labeling), the never-dried stems of spruce, and the model tri-peptide formyl-Met-Leu-Phe-OH (MLF). For both *A. fumigatus* and spruce, around 35-40 mg of materials were packed into 3.2 mm pencil rotors. In addition, 7 mg of MLF was center packed into

a 4 mm rotor. The fungal sample was freshly prepared following previously established protocols[33, 34, 39]. Briefly, the mycelium of *A. fumigatus* was obtained by growing the wild-type strain (RL 578) in a minimum liquid medium containing ^{13}C-glucose (10.0 gram/L) and ^{15}N-sodium nitrate (6.0 gram/L). The culture was incubated at 30 °C for 3 days under 210 rpm shaking condition. The material was then harvested by centrifugation (7,000 × g; 20 mins), and the resulting pellet was washed using 10 mM phosphate buffer (pH 7.4). The ^{13}C-spruce sample was acquired from Isolife (the Netherlands), with the stems debarked manually. The wood material was hand-grinded using a pestle and a mortar to achieve relatively uniform size for the pieces in the range of a few millimeters across.

5.4.2 Experiments

The 3D experiment was first verified using MLF on a 400 MHz (9.4 Tesla) using a 4.0 mm probe under 10 kHz MAS. The CORD mixing was 53 ms for the 3D experiment. The recycle delay was 1.6 s, and 8 scans were collected. The acquisition time was 15 ms, 7.0 ms, and 5.2 ms for the F3, F2, and F1, respectively (with respective spectral widths of 497, 99, and 80 ppm). The F2 and F1 dimensions have 140 and 84 points (70 and 42 complex points), respectively. The total experimental time was 43 h. The F2-F3 planes showed good resolution of all carbon sites as demonstrated in **Figure S5.3**.

Both the *A. fumigatus* and spruce samples were measured on an 800 MHz NMR (18.8 Tesla) Bruker Avance III HD spectrometer at the National High Magnetic Field Laboratory (NHMFL) using a 3.2 mm Low-E HCN probe designed and constructed in-house. The experiments were conducted under 13.5 kHz MAS at 293 K. The ^{13}C chemical shifts were externally referenced to the tetramethylsilane (TMS) scale. The typical radiofrequency field strengths were 83 kHz for ^{1}H decoupling and pulses, and 62.5 kHz for ^{13}C pulses.

A 53-ms CORD recoupling time was used for the 3D experiment of *A. fumigatus*. The spectral widths were set to 497 ppm, 68 ppm, and 66 ppm for the F3, F2, and F1, respectively. The acquisition time was chosen to be 15 ms (F3), 6.0 ms (F2), and 6.0 ms (F1). This corresponds to 164 and 160 slices for the F2 and F1 dimensions. The recycle delay was 1.1 s, and 16 scans were collected. The total experimental time was 134 h.

Two INADEQUATE-CORD spectra were collected using the spruce sample, with short (53 ms) and long (300 ms) CORD mixing. Both 3D experiments used spectral widths of 497 ppm, 62 ppm, and 70 ppm for the F3, F2, and F1, respectively. The short-mixing 3D spectrum has acquisition times of 15 ms, 4.1 ms, and 4.1 ms for the F3, F2, and F1, respectively. These values correspond to 102 and 116 slices for the F2 and F1 dimensions. The short-mixing 3D spectrum was collected using 32 scans and 1 s recycle delays, with a total experimental time of 110 h. The number of scans was doubled for the long-mixing 3D spectrum for better signal-to-noise ratios, which is the barrier for observing the weak long-range cross peaks. The recycle delay was further shortened to 0.93 s. The F2 and F1 dimensions have 74 and 84 slices, respectively. The total experimental time for the long-mixing 3D was 136 h. For each sample, two collapsed 2D spectra of the F1-F3 and F2-F3 dimensions were measured separately using identical parameters as used in the 3D experiments.

5.5 Conclusions

Because the current solid-state NMR investigations of complex carbohydrates heavily rely on ^{13}C detection, the resolution improvement provided by this 3D DQ-SQ-SQ correlation experiment has significantly expanded our technical capability. This study has demonstrated the unambiguous assignment of all carbon signals in the chitin and glucans of fungal cell walls. The spectra have also shown sufficient resolution for investigating cellulose and matrix

polysaccharides in plant biomass. This 3D CCC approach could become a standard experiment for future investigations of the polymorphic structure and spatial association of carbohydrates in cellular systems and functional biomaterials[46]. Inclusion of novel homonuclear recoupling and third-spin-assisted recoupling methods[47-49], as well as the introduction of underexplored nuclei (such as ^{1}H, ^{2}H ^{15}N, ^{17}O, and ^{19}F)[18-20, 50-53] might further extend the use of 3D experiments in carbohydrate solid-state NMR research.

5.6 Acknowledgments

This work was supported by the National Science Foundation MCB-1942665. The National High Magnetic Field Laboratory is supported by National Science Foundation through NSF/DMR-1644779 and the State of Florida.

REFERENCES

1. Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H., Backbone and Side-Chain 13C and 15N Signal Assignments of the a-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem* **2001,** *2*, 272-281.

2. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-state NMR spectroscopy. *Nat. Rev. Methods Primers* **2021,** *1* (1), 2.

3. Hoffmann, J.; Ruta, J.; Shi, C.; Hendriks, K.; Chevelkov, V.; Franks, W. T.; Oschkinat, H.; Giller, K.; Becker, S.; Lange, A., Protein resonance assignment by BSH-CP-based 3D solid-state NMR experiments: A practical guide *Magn. Reson. Chem.* **2020,** *58*, 445-465.

4. Heise, H.; Seidel, K.; Etzkorn, M.; Becker, S.; Baldus, M., 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations. *J. Magn. Reson.* **2005,** *173*, 64-74.

5. Zhou, D.; Kloepper, K. D.; Winter, K. A.; Rienstra, C. M., Band-selective ^{13}C homonuclear 3D spectroscopy for solid proteins at high field with rotor-synchronized soft pulses. *J. Biomol. NMR* **2006,** *34*, 245-257.

6. Li, S.; Zhang, Y.; Hong, M., 3D ^{13}C-^{13}C-^{13}C Correlation NMR for De Novo Distance Determination of Solid Proteins and Application to a Human Alpha Defensin. *J. Magn. Reson.* **2010,** *202*, 203-210.

7. Dregni, A. J.; Mandala, V. S.; Wu, H.; Elkins, M. R.; Wang, H. K.; Hung, I.; DeGrado, W. F.; Hong, M., In vitro 0N4R tau fibrils contain a monomorphic β-sheet core enclosed by dynamically heterogeneous fuzzy coat segments. *Proc. Natl. Acad. Sci. USA* **2019,** *116*, 16357-16366.

8. Duer, M. J., The contribution of solid-state NMR spectroscopy to understanding biomineralization: atomic and molecular structure of bone *J. Magn. Reson.* **2015,** *253*, 98-110.

9. Dick-Pérez, M.; Zhang, Y.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell-wall polysaccharides by two-and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* **2011,** *50* (6), 989-1000.

10. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional 13C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015,** *54* (14), 2335-2345.

11. Cosgrove, D. J., Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* **2014,** *22*, 122-131.

12. Zhang, Y.; Yu, J.; Wang, X.; Durachko, D. M.; Zhang, S.; Cosgrove, D. J., Molecular insights into the complex mechanics of plant epidermal cell walls. *Science* **2021,** *372* (6543), 706-711.

13. Park, Y. B.; Cosgrove, D. J., A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases *Plant Physiol.* **2012,** *158*, 1933-1943.

14. Struppe, J.; Quinn, C. M.; Lu, M.; Wang, M.; Hou, G.; Lu, X.; Kraus, J.; Andreas, L. B.; Stanek, J.; Lalli, D.; Lesage, A.; Pintacuda, G.; Maas, W.; Gronenborn, A. M.; Polenova, T., Expanding the horizons for structural analysis of fully protonated protein assemblies by NMR spectroscopy at MAS frequencies above 100 kHz. *Solid State Nucl. Magn. Reson.* **2017,** *87*, 117-125.

15. Andreas, L. B.; Jaudzems, K.; Stanek, J.; Lalli, D.; Bertarello, A.; Le Marchand, T.; Cala-De Paepe, D.; Kotelovica, S.; Akopjana, I.; Knott, B.; Wegner, S.; Engelke, F.; Lesage, A.; Emsley, L.; Tars, K.; Herrmann, T.; Pintacuda, G., Structure of fully protonated proteins by proton-detected magic-angle spinning NMR. *Proc. Natl. Acad. Sci. USA* **2016,** *113* (33), 9187-9192.

16. Su, Y.; Andreas, L.; Griffin, R. G., Magic Angle Spinning NMR of Proteins: High-Frequency Dynamic Nuclear Polarization and ^{1}H Detection. *Annu. Rev. Biochem.* **2015,** *84* (1), 465-497.

17. Keeler, E. G.; Michaelis, V. K.; Wison, C. B.; Hung, I.; Wang, X.; Gan, Z.; Griffin, R. G., High-Resolution ^{17}O NMR Spectroscopy of Structural Water. *J. Phys. Chem. B* **2019,** *123*, 3061-3067.

18. Phyo, P.; Hong, M., Fast MAS ^{1}H–^{13}C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73* (12), 661-674.

19. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206* (1), 66-72.

20. Sefzik, T. H.; Houseknecht, J. B.; Clark, T. M.; Prasad, S.; Lowary, T. L.; Gan, Z.; Grandinetti, P. J., Solid-state 17O NMR in carbohydrates. *Chem. Phys. Lett.* **2007,** *434* (4), 312-315.

21. Cadars, S.; Sein, J.; Duma, L.; Lesage, A.; Pham, T. N.; Baltisberger, J. H.; Brown, S. P.; Emsley, L., The refocused INADEQUATE MAS NMR experiment in multiple spin-systems: Interpreting observed correlation peaks and optimising lineshapes. *J. Magn. Reson.* **2007,** *188* (1), 24-34.

22. Cadars, S.; Lesage, A.; Emsley, L., Chemical Shift Correlations in Disordered Solids. *J. Am. Chem. Soc.* **2005,** *127* (12), 4466-4476.

23. Hou, G.; Yan, S.; Sun, S.; Han, Y.; Byeon, I.-J. L.; Ahn, J.; Concel, J.; Samoson, A.; Gronenborn, A. M.; Polenova, T., Spin Diffusion Driven by R-Symmetry Sequences: Applications to Homonuclear Correlation Spectroscopy in MAS NMR of Biological and Organic Solids. *J. Am. Chem. Soc.* **2011,** *133* (11), 3943-3953.

24. Hou, G.; Yan, S.; Trébosc, J.; Amoureux, J.-P.; Polenova, T., Broadband homonuclear

correlation spectroscopy driven by combined R2nv sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids. *J. Magn. Reson.* **2013,** *232*, 18-30.

25. Brown, G. D.; Denning, D. W.; Gow, N. A. R.; Levitz, S. M.; Netea, M. G.; White, T. C., Hidden Killers: Human Fungal Infections. *Sci. Transl. Med.* **2012,** *4* (165), 165rv13.

26. Latgé, J.-P.; Chamilos, G., Aspergillus fumigatus and Aspergillosis in 2019. *Clin. Microbiol. Rev.* **2019,** *33* (1), e00140-18.

27. Poulhazan, A.; Dickwella Widanage, M. C.; Muszynski, A.; Arnold, A. A.; Warschawski, D. E.; Azadi, P.; Marcotte, I.; Wang, T., Identification and Quantification of Glycans in Whole Cells: Architecture of Microalgal Polysaccharides Described by Solid-State Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **2021,** *46*, 19374-19388.

28. Kelly, J. E.; Chrissian, C.; Stark, R. E., Tailoring NMR experiments for structural characterization of amorphous biological solids: A practical guide. *Solid State Nucl. Magn. Reson.* **2020,** *109*, 101686.

29. Ghassemi, N.; Poulhazan, A.; Deligey, F.; Mentink-Vigier, F.; Marcotte, I.; Wang, T., Solid-State NMR Investigations of Extracellular Matrices and Cell Walls of Algae, Bacteria, Fungi, and Plants. *Chem. Rev.* **2021**, in press.

30. Thongsomboon, W.; Serra, D. O.; Possling, A.; Hadjineophytou, C.; Hengge, R.; Cegelski, L., Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. *Science* **2018,** *359* (6373), 334-338.

31. Overall, S. A.; Barnes, A. B., Biomolecular Perturbations in In-Cell Dynamic Nuclear Polarization Experiments. *Front. Mol. Biosci.* **2021,** *8*, 743829.

32. Ghosh, R.; Xiao, Y.; Kragelj, J.; Frederick, K. K., In-Cell Sensitivity-Enhanced NMR of Intact Living Mammalian Cells. *J. Am. Chem. Soc.* **2021,** *143* (44), 18454-18466.

33. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

34. Chakraborty, A.; Fernando, L. D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latgé, J. P.; Wang, T., A molecular vision of fungal cell wall organization by functional genomics and solid-state NMR. *Nat. Commun.* **2021,** *12*, 6346.

35. Latgé, J.-P.; Beauvais, A.; Chamilos, G., The Cell Wall of the Human Fungal Pathogen Aspergillus fumigatus: Biosynthesis, Organization, Immune Response, and Virulence. *Annu. Rev. Microbiol.* **2017,** *71* (1), 99-116.

36. Latgé, J.-P., The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.* **2007,** *66* (2), 279-290.

37. Lee, M. J.; Sheppard, D. C., Recent advances in the understanding of the Aspergillus fumigatus

cell wall. . *J. Microbiol.* **2016,** *54*, 232-242.

38. Gastebois, A.; Clavaud, C.; Aimanianda, V.; Latgé, J.-P., *Aspergillus fumigatus*: cell wall polysaccharides, their biosynthesis and organization. *Future Microbiol.* **2009,** *4* (5), 583-595.

39. Fernando, L. D.; Dickwella Widanage, M. C.; Penfield, J.; Lipton, A. S.; Washton, N.; Latge, J.-P.; Wang, P.; Zhang, L.; Wang, T., Structural Polymorphism of Chitin and Chitosan in Fungal Cell Walls from Solid-State NMR and Principal Component Analysis. *Front. Mol. Biosci.* **2021,** *8*, 727053.

40. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

41. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct determination of hydroxymethyl conformations of plant cell wall cellulose using ^{1}H polarization transfer solid-state NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

42. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis* primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiol.* **2015,** *168* (3), 871-884.

43. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

44. Cresswell, R.; Dupree, R.; Brown, S. P.; Pereira, C. S.; Skaf, M. S.; Mathias, S.; Dupree, P.; Hill, S., Importance of Water in Maintaining Softwood Secondary Cell Wall Nanostructure. *Biomacromolecules* **2021,** *22*, 4669-4680.

45. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* **2020,** *11* (1), 6081.

46. Hariri El Nokab, M. E.; van der Wel, P. C. A., Use of Solid-state NMR spectroscopy for Investigating Polysaccharide-based Hydrogels: A Review. *Carbohydr. Polym.* **2020,** *240*, 116276.

47. Gelenter, M. D.; Dregni, A. J.; Hong, M., Pulsed Third-Spin-Assisted Recoupling NMR for Obtaining Long-Range 13C–13C and 15N–13C Distance Restraints. *J. Phys. Chem. B* **2020,** *124* (33), 7138-7151.

48. Paul, S.; Takahashi, H.; Hediger, S.; De Paepe, G., Third Spin-Assisted Recoupling in SSNMR: Theoretical Insights and Practicable Application to Biomolecular Structure Determination. *Annual Reports on NMR Spectroscopy* **2015,** *85*, 93-142.

49. Paëpe, G. D.; Lewandowski, J. R.; Loquet, A.; Böckmann, A.; Griffin, R. G., Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **2008,** *129* (24), 245101.

50. Wu, G., ^{17}O NMR studies of organic and biological molecules in aqueous solution and in the solid state. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2019,** *114*, 1335-191.

51. Fritz, J. H.; Kraus, J.; Quinn, C. M.; Yap, G. P. A.; Struppe, J.; Sergeyev, I. V.; Gronenborn, A. M.; Polenova, T., Measurement of accurate interfluorine distances in crystalline organic solids: a high-frequency magic angle spinning NMR approach. *J. Phys. Chem. B* **2019,** *123*, 10680-10690.

52. Shcherbakov, A. A.; Medeiros-Silva, J.; Tran, N.; Gelenter, M. D.; Hong, M., From Angstroms to Nanometers: Measuring Interatomic Distances by Solid-State NMR. *Chem. Rev.* **2021**, in press. .

53. Gelenter, M. D.; Wang, T.; Liao, S.-Y.; O'Neill, H.; Hong, M., ^{2}H–^{13}C correlation solid-state NMR for investigating dynamics and water accessibilities of proteins and carbohydrates. *J. Biomol. NMR* **2017,** *68* (4), 257-270.

APPENDIX

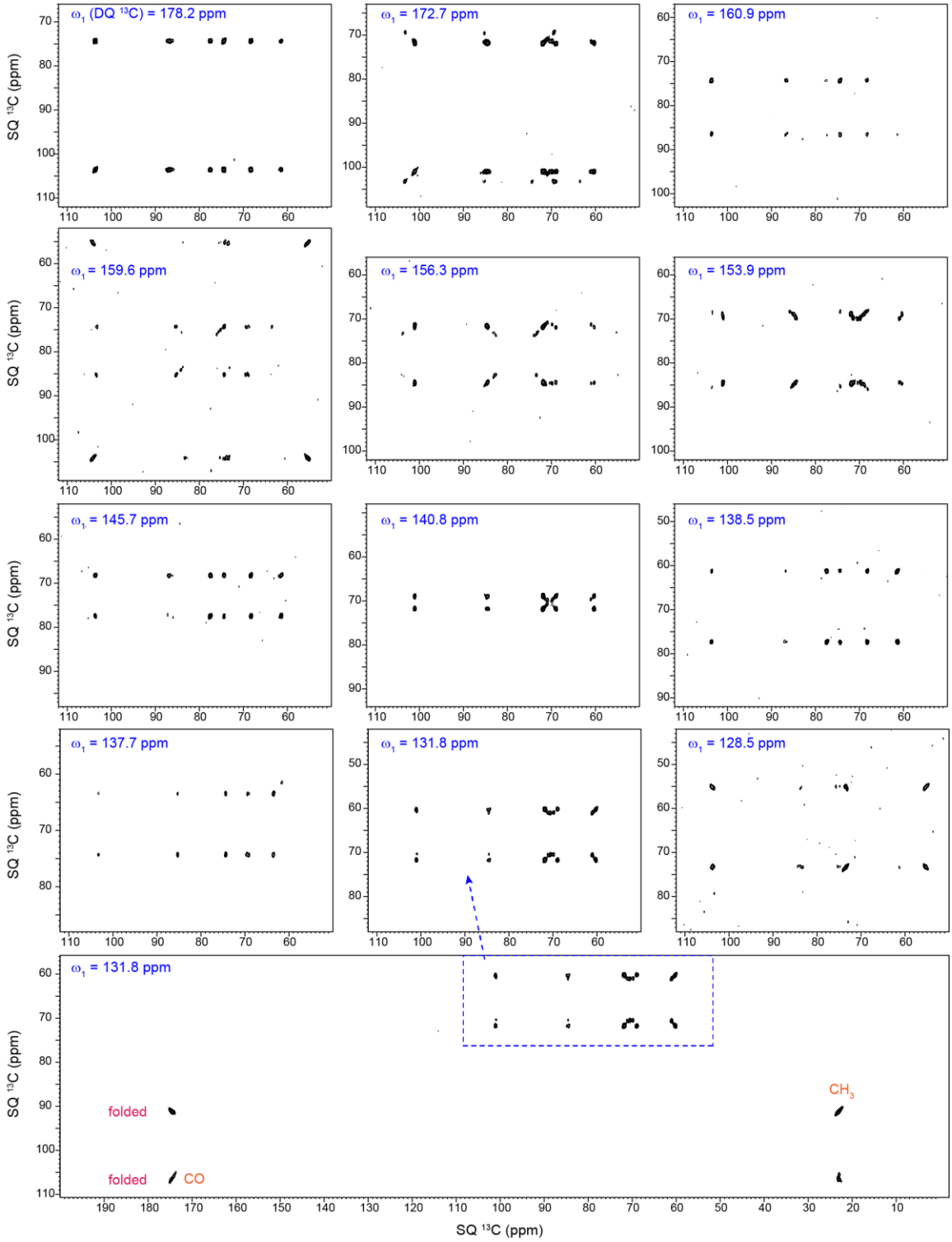

Figure S5.1 F2-F3 planes of *A. fumigatus* 3D spectra. The ω_1 double-quantum (DQ) ^{13}C chemical shifts were labeled for each 2D plane. Bottom panel contains folded peaks for chitin CO and CH_3.

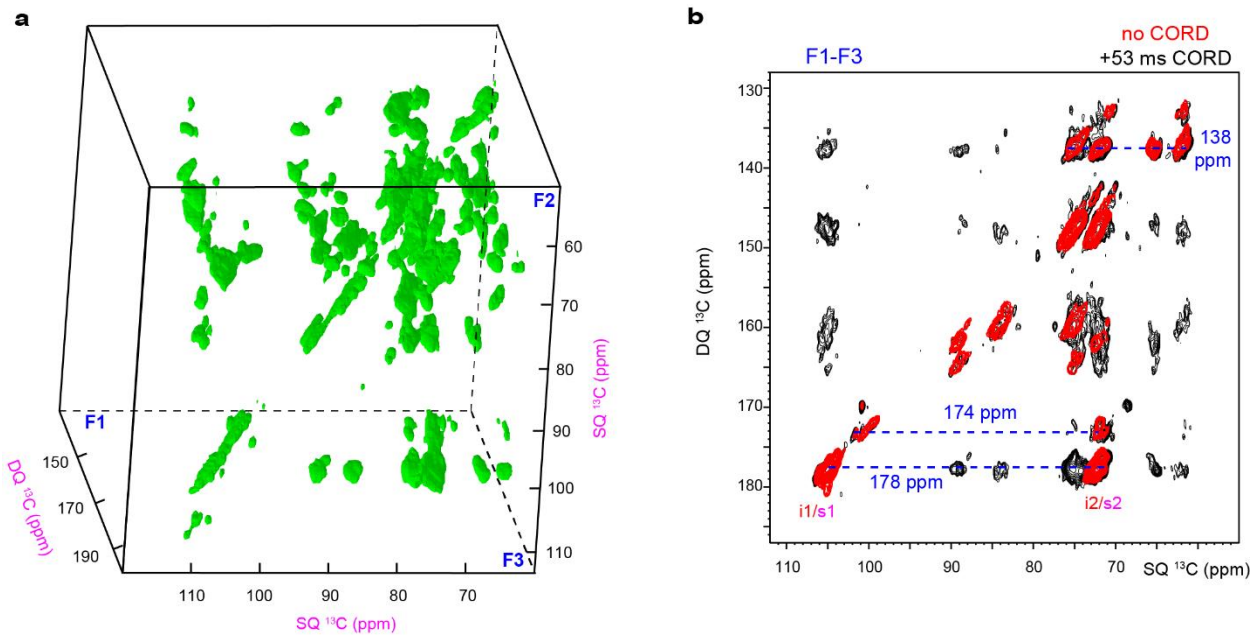

Figure S5.2 3D spectra of spruce stems with highly crowded spectra. **a,** 3D INADEQUATE-CORD of never-dried spruce measured on an 800 MHz NMR. **b,** The 2D F1-F3 plane of 3D INADEQUATE-CORD spectrum (black). A conventional INADEQUATE spectrum (red) is compared. Blue dashlines indicate the F1 positions at which 2D F2-F3 planes are extracted in Figure 5.4.

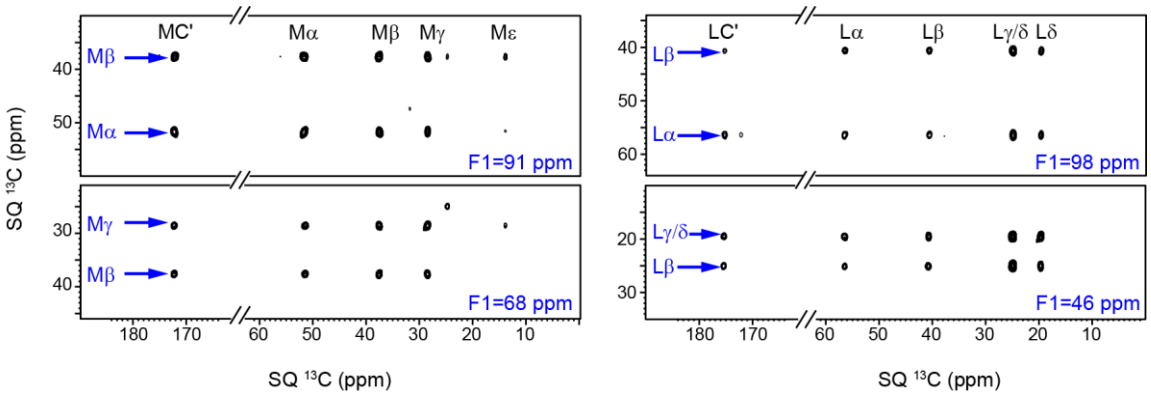

Figure S5.3 F2-F3 planes extracted from the 3D spectrum of MLF. Representative 2D planes were shown for Met and Leu. The signals of these two residues are well resolved from each other. The rows of Met and Leu carbons are indicated using blue. The data were collected on a 400 MHz NMR.

Supplementary Data: Pulse Program of DQ-SQ-SQ 3D CCC Experiment

---

```
;p1 : 13C  90
;p2 : 13C 180
;p3 : 1H   90
;p15: CP Contact Time
;p4 :CORD basic block
;p31:P3*170/90=TPPM dec.pulse
;pl1 : 13C CP Contact
;pl2 : 1H excitation
;pl3 : 1H CP Contact
;pl11: C-C J Mixing
;pl12: 1H Dec for acq
;pl22: 1H dec for evoln and J-mixing
;cpdprg2: 1H Dec4Inadq+t1: tppm15_p26_pl22
;cpdprg3: 1H Dec Acq, use tppm15_p31_pl12
;cnst31 : spinning speed (Hz)
;l6=cnst31*(d2+p1/2+p2/2+1u)
;l1=td1/2,   for States cosine/sine
;l2=td2/2,   for States cosine/sine
;l4:CORDloop# mixing=l4*24*tauR<500m
;d0      : t1=ntr
;d2=l6*1s/cnst31-p1/2-p2/2-1u  ;1/2 of J evolution, up to 5 ms
;d10     : t2=ntr
;d11=30m

;#include <Avancesolids.incl>
;#include <tppm.incl>

"l1=td1/2"                        ; for States (t1)
"in0=inf1"

"l2=td2/2"                       ; for States (t1 & t2)
"in10=inf2"
aqseq 321

"p4=(1s/cnst31)/2"
"p30=p31-0.4u"

"d2=l6*1s/cnst31-p1/2-p2/2-1u"        ; 1/2 of J evolution, up to 5 ms
"d11=30m"

dccorr
ze
```

```
1 d11                       ; t1-loop
;  d11                      ; t1 States loop
;   d11                      ; t2-loop
;    d11                      ; t2 States-loop
;     d11                      ; go-loop

      d1 do:f2
      1m rpp17
      10u reset:f1 reset:f2
      6u setnmr3|28 \n 4u  setnmr3^28

      10u pl1:f1
      10u pl2:f2

      p3:f2 ph19
      (p15:spf0 ph1):f1 (p15 pl3 ph20):f2
      1u cpds2:f2 pl22:f2

      d2 pl11:f1                ; C-C J mixing
      p2:f1 ph2                 ; X 180 pulse
      d2
      1u

      p1:f1 ph3                 ; X 90 mixing
      d0                        ; t1=ntr
      p1:f1 ph4                 ; X 90 mixing

      d2                        ; refocus
      p2:f1 ph5
      d2
      d10

      p1:f1 ph6                 ; X 90 mixing
      1u do:f2

;     cord
2     (p4 pl13 ph17^):f2
      (p4 pl13 ph17^):f2
      (p4 pl13 ph17^):f2
      (p4 pl13 ph17^):f2
      (p4*2 pl14 ph17^):f2
      (p4*2 pl14 ph17^):f2
      (p4*2 pl14 ph17^):f2
      (p4*2 pl14 ph17^):f2

      (p4 pl13 ph17^):f2
```

```
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2

   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2

   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4 pl13 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   (p4*2 pl14 ph17^):f2
   lo to 2 times l4

   1u do:f2
   p1:f1 ph7                    ; X 90 mixing
1u cpds3:f2 pl12:f2
go=1 ph31
30u do:f2                     ; decoupler off
d11 wr #0 if #0 zd                ; write2disk,incrmntFilePntr,0data

1m ip4
1m ip5
lo to 1 times 2                 ; States detection
1m rp4
1m rp5
9m id10                       ; increment d0 by in0 = tr
lo to 1 times l2
10m rd10

1m ip1
1m ip2
1m ip3
```

```
   lo to 1 times 2                ; States detection
   1m rp1
   1m rp2
   1m rp3
   9m id0                       ; increment d0 by in0 = tr
   lo to 1 times l1
   10m rd0

HaltAcqu, 1m
exit

ph1 = (8) 2 4 6 0               ;CP pulse on X
ph2 = (8) 0 2 4 6               ;first  180
ph3 = (8) 0 2 4 6               ;DQ generation
ph4 =  0                       ;DQ reconversion
ph5 =    {0}*8   {2}*8            ;second 180
ph6 =    2
ph7 =    0
ph17=   1 3 2 2  1 3 2 2
         2 0 3 3  2 0 3 3
         3 1 0 0  3 1 0 0
         0 2 1 1  0 2 1 1
ph19=   {1}*4   {3}*4            ;1H excitation
ph20=   0                        ;1H CP
ph31=   0 2 0 2  2 0 2 0

; NS=8*n
```

CHAPTER 6: EFFECT OF CROSS POLARIZATION RADIOFREQUENCY PHASES ON SIGNAL PHASE

Research paper reprinted with permission: *Solid State Nucl. Magn. Reson.* 117, 101771 (2022)

Authors: S. Chandra Shekar[#], Wancheng Zhao[#], Thomas K. Weldeghiorghis, and Tuo Wang

[#]Equal contribution

6.1 Abstract

Utilizing phases of radio frequency (RF) pulses to manipulate spin dynamics is routine in NMR and MRI, leading to spectacular techniques like phase cycling. In a very different area, cross polarization (CP) also has a long history as part of a vast number of solid-state NMR pulse sequences. However, a detailed study devoted to the effect of CP RF phases on NMR signal, seems not to be readily available. From the first principles, we arrive at a simple dependence of NMR signal on arbitrary CP RF phases, for static and MAS conditions, accompanied by experimental verification. In the process, the CP propagator emerges as a product of RF "pulses" and a period of "free precession", conforming to coherence transfer pathway theory. The theoretical expressions may lend confidence for dealing with CP blocks with tunable phases in pulse sequences.

6.2 Introduction

Effect of phases of radio frequency (RF) pulses on spin dynamics is a deeply studied topic in NMR and MRI[1]. It has led to spectacularly successful techniques such as phase cycling[2, 3] to select certain coherence transfer pathways (CTP) while blocking the undesirable ones, for spectral editing and suppression of image artifacts in countless number of NMR and MRI pulse sequences. This in turn has revolutionized NMR via pulse sequences such as COSY, NOESY, TOCSY, TROSY etc., as well as innumerable 3D pulse sequences built upon them.

Cross polarization (CP)[4-7] technique is employed, both as a standalone experiment, as well

as a building block in an innumerable number of solid-state NMR experiments[8]. Often as a first step, CP is an integral part of solid-state NMR pulse sequences[9] to transfer polarization from NMR sensitive protons (^{1}H) to insensitive X-nuclei such as ^{13}C and ^{15}N[10]. It is a regular feature of numerous 1D, 2D, and 3D solid-state NMR pulse sequences. CP is used in static as well as magic-angle spinning (MAS) conditions, including (by no means limited to) widely differing applications such as reduction of experimental time driven by the much shorter ^{1}H T_1, selective detection of rigid components in the sample[11], extraction of distance information by estimating the dipolar interaction based on the CP intensity build up curves[12], etc. Also, there has been a steady effort to extend applicability by expanding its capability: combining CP with MAS (resulting in CPMAS)[13], CP at fast MAS[14, 15] and ultra-fast MAS[16, 17], tilted angle CP (such as DOTA-CP)[18], Lee-Goldburg CP[19], broad band CP[20] in an attempt to broaden the Hartmann-Hahn (HH) match, CP for 'wide-line' NMR (BRAIN-CP)[21], exploiting chemical shift offset to reduce the power requirements for CP[10, 22], selective CP (SPECIFIC-CP)[23] for X-nuclei at specific chemical shifts as an assignment strategy, and repetitive CP blocks for quantitative detection (Multi-CP)[24]. Interestingly CPMAS itself becomes part of the strategy to transfer from even richer source of polarization (electrons) in indirect DNP experiments[25, 26].

Yet, there does not seem to be an in-depth study examining the effect of CP RF phases on the NMR signal (notwithstanding simple and standard phases $\pm x$ and $\pm y$[27-29]), to the best of our knowledge. The lack of related studies is further evidenced by the results of literature search. For example, what would be the effect of CP *S*-channel RF phase of 30^{o}, under static and MAS conditions? We explore this elementary question below in a general manner.

Both thermodynamic[4-6, 27, 30] and quantum mechanical descriptions[1, 14, 15, 20-23, 27, 30-40] are available to understand the CP phenomenon, which has been essential because the technique has undergone

many major developments to broaden its applicability since its inception. Drawing upon this established literature, we arrive at a surprisingly simple relationship between the phases of CP RF fields and the NMR signal for a heteronuclear spin-$\frac{1}{2}$ pair, for the pulse sequence shown in **Figure 6.1,** and followed by extensive experimental verification, carried out on several compounds by employing ^{1}H-X (X= ^{13}C, ^{15}N, ^{31}P) CP. Such expressions may come in handy to design pulse schemes combining transients resulting from different sets of RF phases during CP.

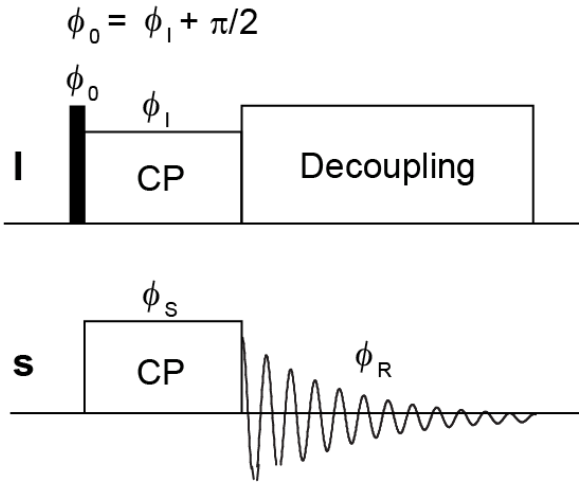

Figure 6.1 Effect of RF phases during cross polarization on NMR signal. In the pulse sequence, ϕ_0 and ϕ_I are respectively RF phases during $\frac{\pi}{2}$- pulse and CP on *I*-spins. Under the constraint $\phi_0 = \phi_I + \frac{\pi}{2}$, *S*-spin signal acuires same phase as that of RF phase ϕ_S during CP, experimentally confirming Eq. (20). ϕ_I *and* ϕ_S are independently cycled according to: $\phi_{I_k} = k2\pi/N_I;\ \ k = 0,1,\ldots.,N_I - 1$. Each ϕ_{I_k} was accompanied by $\phi_S(l) = \phi_S^0 + \phi_S = \phi_S^0 + \frac{l2\pi}{N_S}; l = 0,1,\ldots.,N_S - 1$. ϕ_S^0 is the constant overall phase shift. Receiver phase $\phi_R \rightarrow \phi_{Rl} = \phi_{Sl.}$, thus, selecting the observable -1QC for the *S*-spins (see text).

6.3 Methods

To test the theoretical predictions ensuing from the derivations in this work, CP experiments were performed on four samples, packed into 4 Bruker 4-mm zirconium rotors: Rotors (1) and (2): 10 mg and 75 mg of U(^{13}C,^{15}N) glycine (from Cambridge Isotope Laboratories), respectively; Rotor (3): DMPC multilamellar vesicles, prepared by dispersing 25mg of the lipid (from Avanti) in 40 μL of water; Rotor (4): around 5 mg of U(^{13}C,^{15}N) labeled tripeptide formyl-Methionine-Leucine-Phenylalanine-OH (MLF). All data were acquired on a Bruker Avance 400 MHz (9.4 Tesla) spectrometer, equipped with a multinuclear 4-mm MAS probe as detailed previously, along with experimental details, unless noted otherwise in this work[41-43]. Typical RF field strengths were, 83 kHz for ^{1}H $\frac{\pi}{2}$-pulse, 71 kHz for ^{1}H TPPM decoupling (with associated pulse widths of 6.7 and 6.3 μs), 50-63 kHz for ^{1}H CP, typically either nominally matched to the CP RF field of the X (^{13}C, ^{15}N, ^{31}P) nucleus ± spinning speed (corresponding respectively to DQ or ZQ HH match), or experimentally optimized. No ramp was applied on either RF channel during CP. A recycle delay of 3s was employed between free induction decay transients, and 4 to 256 transients were accumulated for signal averaging.

6.4 Theory

In **Figure 6.1**, the phase ϕ_0 of the excitation $\frac{\pi}{2}$-pulse and ϕ_I of the CP 'pulse' (on *I*-spins) are constrained by[44]

$$\phi_0 = \phi_I + \frac{\pi}{2} \qquad \textbf{Eq 6.1}$$

Starting with the density matrix corresponding to *I*-spin equilibrium magnetization at time t = 0,

$$\boldsymbol{\sigma}(0) = \mathbf{I_z} \qquad \textbf{Eq 6.2}$$

P, a $\frac{\pi}{2}$-pulse with phase $\phi_I + \frac{\pi}{2}$ on *I*-spins (see **Figure 6.1**), is given by[1, 45-47]

$$\mathbf{P}\left(\frac{\pi}{2}, \phi_I + \frac{\pi}{2}\right) = \mathbf{R}_z(\phi_I)\, \mathbf{R}_y(\tfrac{\pi}{2})\, \mathbf{R}_z^{\dagger} \qquad \textbf{Eq 6.3}$$

$$\mathbf{R}_k(\theta) = e^{-i\theta \mathbf{I_k}} \qquad \textbf{Eq 6.4}$$

is applied. († symbolizes *adjoint* operation defined by the relationship between matrix elements as follows: $A_{kl}^{\dagger} = A_{lk}^{*}$, with * signaling complex conjugation). Equation (Eq.) (4) represents a rotation in spin space about a direction specified by the unit vector $\hat{\mathbf{k}}$. The pulse rotates Eq. (2) into transverse magnetization at an angle ϕ_I relative to the x-axis in the rotating frame, given by the density matrix, $\boldsymbol{\sigma}(\mathbf{1})$[1, 45-47],

$$\boldsymbol{\sigma}(\mathbf{1}) = \mathbf{R}_z(\phi_I)\, \mathbf{I_x}\, \mathbf{R}_z^{\dagger} \qquad \textbf{Eq 6.5}$$

Evolution of the density matrix during CP, ignoring relaxation effects for simplicity, is governed by[1, 22, 24, 38, 45-48]

$$\boldsymbol{\sigma}(t) = \mathbf{U}(\phi_I \phi_S; t)\, \boldsymbol{\sigma}(\mathbf{1})\, \mathbf{U}^{\dagger} \qquad \textbf{Eq 6.6}$$

$$\mathbf{U}(\phi_I \phi_S; t) = \mathbf{R}\, \mathbf{u}(t)\, \mathbf{R}^{\dagger} \qquad \textbf{Eq 6.7}$$

Its elements are determined by the RF field strengths on both I and S channels and the heteronuclear orientation dependent dipolar coupling[38]. Above,

$$\mathbf{R} = \mathbf{R}_z(\phi_I \phi_S)\, \mathbf{R}_y\left(\frac{\pi^I}{2}\, \frac{\pi^S}{2}\right) \qquad \textbf{Eq 6.8}$$

$$\mathbf{R}_z(\phi_I \phi_S) = \mathbf{R}_z(\phi_I)\, \mathbf{R}_z(\phi_S) = e^{-i(\phi_I \mathbf{I_z} + \phi_S \mathbf{S_z})} \qquad \textbf{Eq 6.9}$$

$$\mathbf{R}_y\left(\frac{\pi^I}{2}\, \frac{\pi^S}{2}\right) = \mathbf{R}_y\left(\frac{\pi^I}{2}\right) \mathbf{R}_y\left(\frac{\pi^S}{2}\right) = e^{-i\frac{\pi}{2}(\mathbf{I_y} + \mathbf{S_y})} \qquad \textbf{Eq 6.1}$$

signaling[1, 14, 38, 45-47], (i) a transformation from the laboratory frame (with $B_0 \parallel z$) to a doubly rotating frame (at respective Larmor frequencies for the I and S spins), (ii) followed by a transformation to a doubly rotated frame along z, (iii) and a subsequent transformation to a doubly tilted frame along y;

Using Eqs. (5) and (7) in (6), $\boldsymbol{\sigma}(t) = \mathbf{R}\,\mathbf{u}\,\mathbf{R}^{\dagger}\,\mathbf{R_z}(\phi_I)\,\mathbf{I_x}\,\mathbf{R_z^{\dagger}}(\phi_I)\,\mathbf{R}\,\mathbf{u}^{\dagger}\,\mathbf{R}^{\dagger}$ since $(AB)^{\dagger} = B^{\dagger} A^{\dagger}$. But from Eqs. (8), (9), and (10),

$$\mathbf{R}^{\dagger}\,\mathbf{R_z}(\phi_I)\mathbf{I_x}\,\mathbf{R_z^{\dagger}}(\phi_I)\mathbf{R} = \mathbf{R_y^{\dagger}}\left(\frac{\pi^I}{2}\,\frac{\pi^S}{2}\right)\mathbf{R_z^{\dagger}}(\phi_I\phi_S)\mathbf{R_z}(\phi_I)\mathbf{I_x}\,\mathbf{R_z^{\dagger}}(\phi_I)\mathbf{R_z}(\phi_I\phi_S)\mathbf{R_y} =$$

$$\mathbf{R_y^{\dagger}}\left(\frac{\pi^I}{2}\,\frac{\pi^S}{2}\right)\mathbf{R_z^{\dagger}}(\phi_S)\mathbf{I_x}\mathbf{R_z}(\phi_S)\mathbf{R_y} = \mathbf{R_y^{\dagger}}\left(\frac{\pi^I}{2}\right)\mathbf{I_x}\mathbf{R_y} = \mathbf{I_z},$$

yielding,

$$\boldsymbol{\sigma}(t) = \mathbf{R}\,\mathbf{D}(t)\,\mathbf{R}^{\dagger} \qquad \textbf{Eq 6.11}$$

$$\mathbf{D}(t) = \mathbf{u}(t)\,\mathbf{I_z}\,\mathbf{u}^{\dagger} \qquad \textbf{Eq 6.12}$$

The amount of an operator **A** present in the density matrix at a time t can be extracted by the following trace operation[1, 45-48]

$$\langle \mathbf{A}(t) \rangle = \mathbf{Tr}\{\mathbf{A}\,\boldsymbol{\sigma}(t)\} = \mathbf{Tr}\{\mathbf{R}^{\dagger}\,\mathbf{A}\,\mathbf{R}\,\mathbf{D}(t)\} \qquad \textbf{Eq 6.13}$$

Using[46]

$$\mathbf{D}\mathbf{R}^{\dagger}\mathbf{S_x}\mathbf{R} = \mathbf{S_z}\cos\phi_S - \mathbf{S_y}\sin\phi_S$$

$$\mathbf{R}^{\dagger}\mathbf{S_y}\mathbf{R} = \mathbf{S_z}\sin\phi_S + \mathbf{S_y}\cos\phi_S$$

$$\mathbf{R}^{\dagger}\mathbf{S_z}\mathbf{R} = -\mathbf{S_x} \qquad \textbf{Eq 6.14}$$

in Eq. (13) yields

$$\langle \mathbf{S_x}(t) \rangle = \cos\phi_S\,P(t) - \sin\phi_S\,Q(t)$$

$$\langle \mathbf{S_y}(t) \rangle = \sin\phi_S\,P(t) + \cos\phi_S\,Q(t)$$

$$\langle \mathbf{S_z}(t) \rangle = -\mathbf{Tr}\{\mathbf{S_x}\mathbf{D}(t)\} \qquad \textbf{Eq 6.15}$$

with

$$P(t) \equiv \mathbf{Tr}\{\mathbf{S_z}\mathbf{D}(t)\}$$

$$Q(t) \equiv \mathbf{Tr}\{\mathbf{S_y}\mathbf{D}(t)\}(t)\} \qquad \textbf{Eq 6.16}$$

Distinction should be drawn between the scalars $\langle \mathbf{S_x}(t) \rangle$, $\langle \mathbf{S_y}(t) \rangle$ and the 4×4 matrices $\mathbf{S_x}$, $\mathbf{S_y}$, each with just 4 non-zero elements = $\frac{1}{2}$ or $\pm i\frac{1}{2}$. Moreover in Eq. (13), when **A** is one of the 16 possible product operators (PO's)[1, 47, 49-51], $\mathbf{R}^{\dagger}\mathbf{AR}$ is mapped back into one of the 16 PO's. Tabulates non-vanishing traces of the PO's with **D**. This enables evaluation of expectation values for all 16 PO's by Eq. (13).

Remarkably,

$$Q(t) \equiv \mathbf{Tr}\{\mathbf{S_y D}(t)\} = 0 \qquad \textbf{Eq 6.17}$$

as shown by Eq. (S31), leaving us with

$$\langle \mathbf{S_x}(t) \rangle = \cos\phi_S \, P(t)$$

$$\langle \mathbf{S_y}(t) \rangle = \sin\phi_S \, P(t) \qquad \textbf{Eq 6.18}$$

Using the ladder operators[1, 45-48]

$$\mathbf{S}_{\pm} = \mathbf{S_x} \pm i\mathbf{S_y} \qquad \textbf{Eq 6.19}$$

on Eq. (18), signal observed in quadrature is given by[1-3, 45-47, 52, 53]

$$f(t) \equiv \langle \mathbf{S}_{+}(t) \rangle = e^{i\phi_S} P(t) \qquad \textbf{Eq 6.20}$$

It can be seen that $P(t)$ is real. It depends only on the CP RF amplitudes and the orientation dependent $I - S$ dipolar interaction, contained in the matrix elements of **u**(t).

But the phase ϕ_S, imparted to the signal is impervious to such details, and Eq. (20) is valid for any regime (Hartmann-Hahn (HH) match, HH mismatch, HH anti match) of CP RF fields, and any distribution of orientations (crystal, partial order, powder). Some of these features have emerged in theoretical simulations previously[38], extending additional support for Eq. (20).

Though the effect of canonical phases of RF fields during CP have been treated before in excellent treatises[27-29], those pioneering works did not deal with arbitrary RF phases. For example, the outcome of phase cycling, employing RF and receiver phases $\pm x, \pm y$ only, can be found in a

very helpful table, but without actual derivation[27-29]; the underlying mechanism is not elucidated. It may be possible to arrive at such a dependence purely by trial-and-error experiments, and/or the intuition based on the effect of resonant RF pulses on NMR signal phase (see below).

Phase acquired[1, 47] by a CTP of coherence jump $\Delta m = -1$, under an RF "pulse" of phase f_S , $e^{-i\,Dm\,fS} = e^{i\,fS}$, is the same as in Eq. (20). At first glance, this seems surprising, as CTP theory requires[1] that the propagator (the 'block') comprise of a sequence of RF pulses, including free precession intervals. Since the CP Hamiltonian is a sum of the non-commuting RF and dipolar Hamiltonians, the ensuing evolution cannot be factored into a simple *product* of exponential operators, apparently failing to meet CTP theory requirements. But, upon closer inspection, Eqs. (7) & (8) reveal that the CP evolution, is indeed a product of actual and *equivalent* RF 'pulses' sandwiching a *free precession* interval. The propagator's phases ϕ_I, and ϕ_S stem from the phases of the actual RF irradiation on *I* and *S*-spins during CP. The diagonalizing y-rotations in Eq. (S20) play the role of RF 'pulses' along y-axis, and the z-rotation that of a free precession interval. Thus, the CTP theory indeed applies to CP evolution. Notice that, under MAS there is an additional effective free precession interval associated with the sample rotation, as well as additional orientation (ϒ) dependent phase factors, in addition to those imparted due to CP RF phases.

6.5 Results and Discussion

First experimental verification of Eq. (20) is furnished in **Figure 6.2**, which shows static ^{13}C NMR CP spectrum of U (^{13}C,^{15}N) glycine. (Please see the section on methods for additional experimental details.) The spectra were acquired at ^{13}C carrier frequency of 100.673 MHz with a spectral width of 50 kHz and 294 complex sampling points in quadrature for an acquisition time of 29.4 ms. ϕ_I and ϕ_S were independently phase cycled, with $\phi_S{}^0$ as the constant overall phase shift, as described in Figure 6.1. Explicitly in units of $\frac{\pi}{2}$, $\phi_S - \phi_S^0 = \phi_R = 0,1,2,3$; $\phi_0 = \phi_I +$

$1; \phi_I = 0,0,0,0, 1,1,1,1, 2,2,2,2, 3,3,3,3$. The time domain data was subject to Gaussian multiplication. The time domain data was subsequently zero filled to 16384 complex points prior to Fourier Transformation (FT). Phase correction[1, 45-47] on spectra was carried out with the same 1st order phase correction ϕ_1^c, regardless of the value of ϕ_S^0. On the other hand, the 0th order phase correction ϕ_0^c, depended on ϕ_S^0. In fact, for spectra acquired with $\phi_S^0 \neq 0$, corrections of $\phi_0^c(\phi_S^0) = \phi_0^c(0) + \phi_S^0$ did the trick, completing the experimental confirmation of Eq. (20). What is more, ϕ_S^0 dependence survives phase cycling of ϕ_I and ϕ_S, further consolidating the validity of Eq. (20). Even though Eq. (20) was derived for the static regime, we found that it held for MAS at 10, 5, and 2.5 kHz, for the same sample as above.

We also found that, for a powder sample of the tripeptide U(^{13}C,^{15}N) formyl-Methionine-Leucine-Phenylalanine-OH (MLF), the signal phase obeyed Eq. (20) for ^{13}C CP MAS at 10 kHz. In addition, we found that Eq. (20) held not only for the phase cycling employed previously, but also under various other schemes (maintaining $\phi_0 = \phi_I + \frac{\pi}{2}$), such as no phase cycling of ϕ_I and/or ϕ_S, N_I and/or N_S =2 or 4 steps.

For the same sample, the ϕ_S dependence was also tested for ^{15}N CP MAS at 2.5 kHz (**Figure 6.3**) and 10 kHz, adding to the growing list of experiments validating Eq. (20). Further affirmation of Eq. (20) followed, by ^{15}N CP MAS at 5 kHz, 1 kHz, and 0 kHz, once again on U(^{13}C,^{15}N) glycine. We extended the investigation of ϕ_S dependence to ^{31}P CP MAS of DMPC multilamellar vesicles that contain 25 mg of lipids and 40 mg of water. The MAS at 1 kHz and 5 kHz demonstrate the phase dependence described by Eq. (20), yet again.

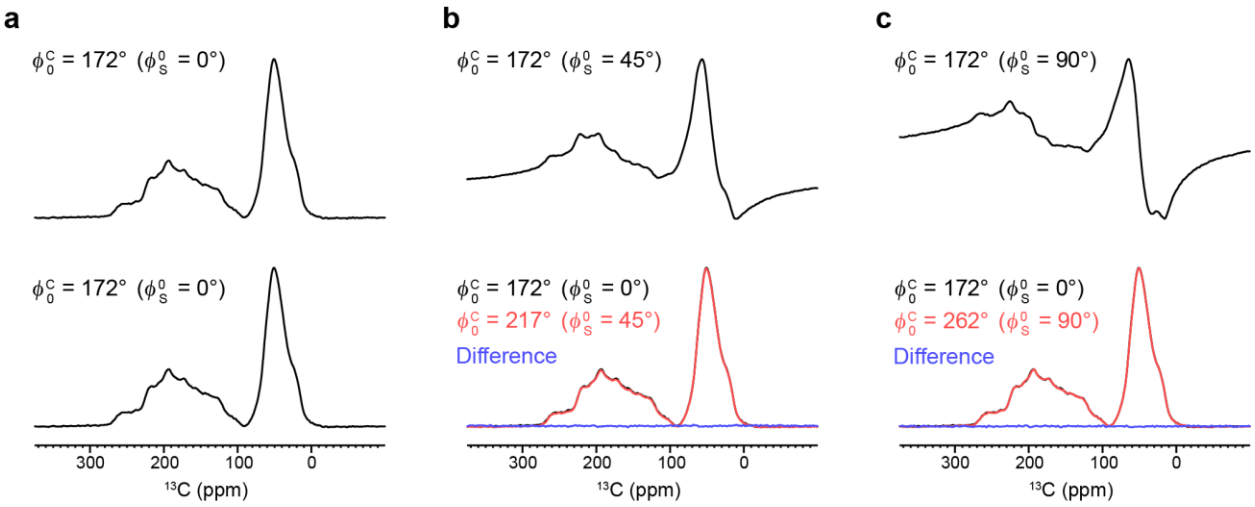

Figure 6.2 Effect of ϕ_S (Figure 6.1) on ^{13}C CP static NMR of U(^{13}C,^{15}N) glycine. B_0=9.4 T and Larmor Frequency ν_0(^{13}C)=100 MHz. ϕ_I and ϕ_S were independently cycled with $N_I = N_S = 4$. Phase correction[1, 45-47] on spectra was carried out with 1st order phase correction $\phi_1^c = 107°$, regardless of the value of ϕ_S^0! On the other hand, the 0th order phase correction ϕ_0^c, is a function of ϕ_S^0 (Figure 6.1). All spectra were acquired by accumulating N=128 shots. Top row: All with same $\phi_0^c = 172°$. Left to Right: ϕ_S^0=0°; 45° and 90° respectively. Bottom row: Same as top row, but with respective corrections ϕ_0^c = 172° + ϕ_S^0 = 172°; 217° and 262°. ϕ_S^0=0 spectrum was superposed in the middle and right panels. In each case, as the displayed difference spectrum shows, the overlays are virtually identical, validating Eq. (20). The ϕ_S^0 dependence persists despite cycling ϕ_I and ϕ_S, further supporting Eq. (20). See text for further details.

6.5.1 Validation on Phospholipid Membranes

Unlike the MAS spectra, the static CP of the same sample in **Figure 6.4**, displays a broad dip, that is deep in the vicinity of isotropic CS, which has been observed previously.[54-58] **Figure 6.4** displays a superposition of static ^{31}P NMR obtained without employing CP — dubbed as direct polarization (DP), under CP, and under CP MAS at 1 kHz.

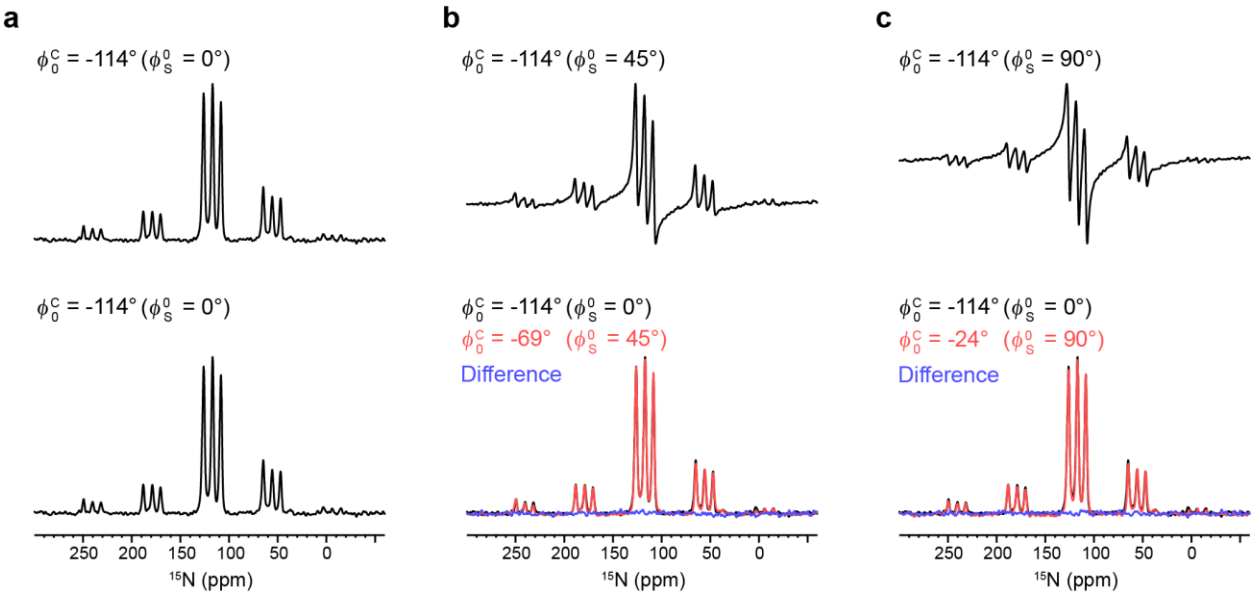

Figure 6.3 Effect of ϕ_S (Figure 6.1) on ^{15}N CP MAS (at 2.5 kHz) spectrum of U(^{13}C,^{15}N) tripeptide MLF. Unless otherwise noted, parameters are the same as in Figure 6.2 Larmor frequency ν_0 (^{15}N)=40 MHz. N=256 and $\phi_1^c = 95°$. Top row: $\phi_0^c = -114°$. Left to Right: ϕ_S^0=0°; 45° and 90° respectively. Bottom row: same as top row but with respective $\phi_0^c = -114° + \phi_S^0 = -114°, -69°,$ and $-24°$. ϕ_S^0 = 0 spectrum was superposed in the middle and right panels, and the difference displayed. In each case, the overlays are virtually identical, Like in Figure 6.2, ϕ_S dependence is indifferent to phase cycling. See text for further details.

When molecular rotation about an axis is rapid on the NMR time scale, the 2nd-rank CS and dipolar (D) tensors are averaged down in magnitude and projected onto the rotation axis as axially symmetric tensors[27, 49-51, 59]. This is borne out by the DP spectrum, which is characteristic of an effectively axially symmetric PAS for CSA, even if the tensor were asymmetric prior to averaging. The position of the underlying resonances that make up the powder pattern, accompanied by a nominal relative amplitude ($1/cos\beta_{ML}$).

For the static CP spectrum, these still determine the resonance position, but the amplitude depends on the residual dipolar interaction available for CP, by

$$\omega_D \equiv d = d_{00}^2(\beta_{PM}) d_{00}^2(\beta_{ML}) d_P \qquad \textbf{Eq 6.21}$$

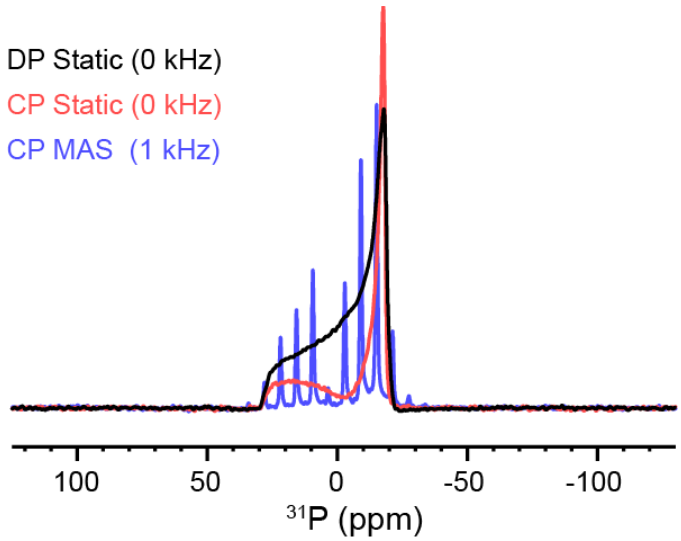

Figure 6.4 Static ^{31}P NMR (black), under CP (red), and under CP-MAS at 1 kHz (blue) of the same sample. The static powder pattern obtained without CP (dubbed as DP spectrum), is characteristic of an isotropic distribution of an effectively axially symmetric CSA tensors around B_0, brought about by fast (on the NMR time scale) rotation about the lipid axis, as shown by Eqs. (S62) and (S76). A nominal relative intensity within the spectrum $= 1/cos\beta_{ML}$. On the other hand, for the static CP spectrum, for the *same* resonance position, the actual amplitude is determined by Eq. (21). At *magic angle*, the resonance is at isotropic shift. For DP, the nominal intensity at this shift is $\sqrt{3}$x that at $\beta_{ML} = 0$. However, at magic angle, from Eq. (21), the residual dipolar interaction =0, resulting in a null signal from Table S2 and Table S3. The overall effect is that of a broad pit that is deep at this shift, in the CP spectrum. On the other hand, in the CP-MAS (at 1 kHz) spectrum, the averaged CSA tensor of size of ~ 8 kHz, breaks up the powder pattern into sidebands. See text for further details.

Which determines the CP amplitude. Above, β_{PM} (see Eq. (S68)) specifies the fixed orientation of the internuclear vector in M-frame, while β_{ML} specifies the orientation of the rotation axis relative to $\mathbf{B_0}$. At the magic angle, $\beta_{ML} = \theta_M$, the resonance is at the isotropic shift, and $d = 0$. This results in $\mathbf{Tr}(\mathbf{S_z}\,\mathbf{D}) = 0$, and in turn, leads to a null signal, corroborated by current experiments (**Figure 6.4**) and consistent with earlier observations[54-58].

For one last time, we interrogated the ϕ_S dependence, on the static ^{31}P CP spectrum of this sample, and confirmed its good agreement with Eq. (20). The totality of experimental findings, thus far, can be summarized as follows:

First and foremost the phase of the *S*-spin NMR signal obeys Eq. (20). Always (not just for 0^o & 180^o and 90^o & 270^o, corresponding to +/-x and +/- y).

No dependence on ϕ_I in the pulse sequence of **Figure 6.1**, as predicted by Eq. (20).

Hence, independent of ϕ_I cycling details.

When ϕ_S is cycled, overall shift ϕ_S^0 replaces ϕ_S in Eq. (20).

Hence, independent of ϕ_S cycling details.

It can be shown rather easily theoretically that cycling f_0 between f_I +/- p/2 (instead of imposing Eq. (1), for the purposes of eliminating thermal equilibrium magnetization of *S*-spins), still preserves Eq. (20), and was confirmed experimentally as well.

6.5.2 Consideration of CP-MAS

Attention was drawn earlier to the fact that Eq. (20) was derived under the assumption of *static* condition. Yet, it was found to be valid under MAS, from 1 to 10 kHz, across different samples and for different observed nuclei (^{13}C, ^{15}N, and ^{31}P). Initially, we entertained the idea that this is due to slow spinning relative to the CP RF amplitudes (~50kHz), and that Eq. (20) may break down as the spinning speed catches up to the CP RF amplitudes. However, we reexamined

the problem under the light of pioneering analytical treatments of CP MAS[14, 30, 39]. Closely following Zilm and coworkers[14], we rederived the MAS equivalents of the evolution operator **u**, and the operator **D**, occurring crucially in Eqs. (11), (12), and (13). It can be seen that $\mathbf{u}^{MAS}$ and $\mathbf{D}^{MAS}$ have identical structures as their static counterparts. Thus, remain valid under MAS, with elements of **D** replaced by the elements of $\mathbf{D}^{MAS}$. In particular, all remain valid under MAS, along with Eqs. (17) and (20), experimentally validated time and again throughout this manuscript. Also, since seem to account for changes in spinning speed, at least the phase dependence given by Eq.(20), may survive even ultra-fast MAS rates that are available currently (> 100 kHz), but remains to be experimentally confirmed.

6.6 Conclusions

We have examined the effect of RF phases during CP on the evolution operator for a heteronuclear spin-$\frac{1}{2}$ pair. The process uncovers that the emergent propagator is factored into a product of rotations due to equivalent RF pulses and free precession intervals (under both static and MAS conditions), in compliance with the CTP theory requirements.[1] We have provided expressions for the effect of these phases on all 16 PO's and found both theoretically and experimentally that their dependencies remain intact under both static and MAS conditions. We can also use this treatment to evaluate the effect of the CP RF phases on any of the 16 single transition operators $|p\rangle\langle q|$ via Eq. (11), if needed. This work may also help to incorporate the CP block with arbitrary and nonstandard ϕ_I and ϕ_S in pulse sequences and combine the resulting signals to design experiments for desired outcomes. For example, suppose the pulse sequence actually starts with CP, followed by a block of pulses, RF irradiations, delays etc. Again, Eq. (11) can be used to compute the entire density matrix (not just $\mathbf{S}_x$ and $\mathbf{S}_y$) at the end of the 1st CP stage for further calculations. E.g., if the next stage is a 2nd CP involving a 3rd spin-1/2 nucleus, we

may ’feed’ the result of the 1st stage as the σ(1) parameter to Eq. (6), albeit treating the two active nuclei as the spin-1/2 ’CP pair’, to evaluate the entire density matrix after the 2nd stage.

Additionally, as mentioned earlier, since the phase dependence due to CP Hamiltonian, given by Eq. (20), coincides with that due to an “ordinary” RF pulse, similar applications may become available utilizing the present work. Knowledge on the effect of arbitrary phases will guide the development of novel pulse sequences to expand the capability of NMR technique. For long we had difficulty locating a systematic description of the CP phase dependence, which stimulated us to undertake this endeavor. We think a complete presentation can be of some interest and might benefit many ssNMR researchers. A possible future direction is to find the equivalent of Eq. (20), which was derived for a spin-1/2 pair, for the case of CP from ^{1}H to quadrupolar nuclei (static as well as MAS and ultra-fast MAS), which may or may not have an analytical solution as it is an open-ended physics question. However, a systematic numerical analysis approach may be mounted, if analytical solution is out of reach.

6.7 Acknowledgements

This work was supported by the National Science Foundation MCB-1942665.

REFERENCES

1. Ernst, R. R.; Bodenhausen, G.; Wokaun, A., *Principles of nuclear magnetic resonance in one and two dimensions*. Oxford: Clarendon press: 1987; Vol. 14.

2. Wokaun, A.; Ernst, R. R., Selective detection of multiple quantum transitions in NMR by two-dimensional spectroscopy. *Chem. Phys. Lett.* **1977,** *52* (3), 407-412.

3. Bodenhausen, G.; Kogler, H.; Ernst, R. R., Selection of coherence-transfer pathways in NMR pulse experiments. *J. Magn. Reson. (1969)* **1984,** *58* (3), 370-388.

4. Pines, A.; Gibby, M. G.; Waugh, J. S., Proton-enhanced nuclear induction spectroscopy. A method for high resolution NMR of dilute spins in solids. *J. Chem. Phys.* **1972,** *56* (4), 1776-1777.

5. Pines, A.; Gibby, M. G.; Waugh, J. S., Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **1973,** *59* (2), 569-590.

6. Hartmann, S. R.; Hahn, E. L., Nuclear double resonance in the rotating frame. *Phys. Rev.* **1962,** *128* (5), 2042.

7. Engelke, F.; Steuernagel, S., Cross Polarization in Rotating Solids: Spin-1/2 Nuclei * Update based on original article by Frank Engelke, Encyclopedia of Magnetic Resonance, ©1996, John Wiley & Sons Ltd. In *eMagRes*, 2011.

8. Bryce, D. L.; Bernard, G. M.; Gee, M.; Lumsden, M. D.; Eichele, K.; Wasylishen, R. E., Practical aspects of modern routine solid-state multinuclear magnetic resonance spectroscopy: one-dimensional experiments. *Can. J. Anal. Sci. Spectrosc.* **2001,** *46* (2), 46-81.

9. Lee, D.; Duong, N. T.; Lafon, O.; De Paëpe, G., Primostrato Solid-State NMR Enhanced by Dynamic Nuclear Polarization: Pentacoordinated Al^{3+} Ions Are Only Located at the Surface of Hydrated γ-Alumina. *J. Phys. Chem. C* **2014,** *118* (43), 25065-25076.

10. Shekar, S. C.; Lee, D.-K.; Ramamoorthy, A., An experimental strategy to dramatically reduce the RF power used in cross polarization solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2001,** *123* (30), 7467-7468.

11. Nowacka, A.; Bongartz, N. A.; Ollila, O. H. S.; Nylander, T.; Topgaard, D., Signal intensities in ^{1}H–^{13}C CP and INEPT MAS NMR of liquid crystals. *J. Magn. Reson.* **2013,** *230*, 165-175

12. Kolodziejski, W.; Klinowski, J., Kinetics of cross-polarization in solid-state NMR: A guide for chemists. *Chem. Rev.* **2002,** *102* (3), 613-628.

13. Stejskal, E. O.; Schaefer, J.; Waugh, J. S., Magic-angle spinning and polarization transfer in proton-enhanced NMR. *J. Magn. Reson. (1969)* **1977,** *28* (1), 105-112.

14. Wu, X. L.; Zilm, K. W., Cross polarization with high-speed magic-angle spinning. *J. Magn. Reson., Ser. A* **1993,** *104* (2), 154-165.

15. Wi, S.; Gan, Z.; Schurko, R.; Frydman, L., Cross-polarization phenomena in the NMR of fast spinning solids subject to adiabatic sweeps. *J. Chem. Phys.* **2015,** *142* (6), 064201.

16. Paluch, P.; Pawlak, T.; Amoureux, J.-P.; Potrzebowski, M. J., Simple and accurate determination of X–H distances under ultra-fast MAS NMR. *J. Magn. Reson.* **2013,** *233*, 56-63.

17. Paluch, P.; Trébosc, J.; Nishiyama, Y.; Potrzebowski, M. J.; Malon, M.; Amoureux, J. P., Theoretical study of CP-VC: A simple, robust and accurate MAS NMR method for analysis of dipolar C–H interactions under rotation speeds faster than ca. 60kHz. *J. Magn. Reson.* **2015,** *252*, 67-77.

18. Matsunaga, T.; Matsuda, I.; Yamazaki, T.; Ishii, Y., Decoherence optimized tilted-angle cross polarization: A novel concept for sensitivity-enhanced solid-state NMR using ultra-fast magic angle spinning. *J. Magn. Reson.* **2021,** *322*, 106857.

19. Venkatesh, A.; Hung, I.; Boteju, K. C.; Sadow, A. D.; Gor'kov, P. L.; Gan, Z.; Rossini, A. J., Suppressing ^{1}H Spin Diffusion in Fast MAS Proton Detected Heteronuclear Correlation Solid-State NMR Experiments. *Solid State Nucl. Magn. Reson.* **2020,** *105*, 101636.

20. Baldus, M.; Meier, B. H., Broadband polarization transfer under magic-angle spinning: application to total through-space-correlation NMR spectroscopy. *J. Magn. Reson.* **1997,** *128* (2), 172-193.

21. Harris, K. J.; Lupulescu, A.; Lucier, B. E. G.; Frydman, L.; Schurko, R. W., Broadband adiabatic inversion pulses for cross polarization in wideline solid-state NMR spectroscopy. *J. Magn. Reson.* **2012,** *224*, 38-47.

22. Shekar, S. C.; Lee, D.-K.; Ramamoorthy, A., Chemical shift anisotropy and offset effects in cross polarization solid-state NMR spectroscopy. *J. Magn. Reson.* **2002,** *157* (2), 223-234.

23. Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G., Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **1998,** *95* (6), 1197-1207

24. Johnson, R. L.; Schmidt-Rohr, K., Quantitative solid-state 13C NMR with signal enhancement by multiple cross polarization. *J. Magn. Reson.* **2014,** *239*, 44-49.

25. Wang, Z.; Hanrahan, M. P.; Kobayashi, T.; Perras, F. A.; Chen, Y.; Engelke, F.; Reiter, C.; Purea, A.; Rossini, A. J.; Pruski, M., Combining fast magic angle spinning dynamic nuclear polarization with indirect detection to further enhance the sensitivity of solid-state NMR spectroscopy. *Solid State Nucl. Magn. Reson.* **2020,** *109*, 101685.

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

27. Schmidt-Rohr, K.; Spiess, H. W., *Multidimensional solid-state NMR and polymers*. Elsevier: 2012.

28. Duer, M. J., *Introduction to solid-state NMR spectroscopy*. Oxford, UK: 2004.

29. Apperley, D. C.; Harris, R. K.; Hodgkinson, P., *Solid-state NMR: Basic principles and practice*. Momentum Press: 2012.

30. Meier, B. H., Cross polarization under fast magic angle spinning: thermodynamical considerations. *Chem. Phys. Lett.* **1992,** *188* (3-4), 201-207.

31. Müller, L.; Kumar, A.; Baumann, T.; Ernst, R. R., Transient oscillations in NMR cross-polarization experiments in solids. *Phys. Rev. Lett.* **1974,** *32* (25), 1402.

32. Müller, L.; Ernst, R. R., Coherence transfer in the rotating frame: Application to heteronuclear cross-correlation spectroscopy. *Mol. Phys.* **1979,** *38* (3), 963-992.

33. Levitt, M. H.; Suter, D.; Ernst, R. R., Spin dynamics and thermodynamics in solid-state NMR cross polarization. *J. Chem. Phys.* **1986,** *84* (8), 4243-4255

34. Zhang, S.; Wu, X., A new model for cross polarization in the rotating frame in solids. *Chem. Phys. Lett.* **1989,** *156* (4), 333-336

35. Levitt, M. H., Heteronuclear cross polarization in liquid-state nuclear magnetic resonance: Mismatch compensation and relaxation behavior. *J. Chem. Phys.* **1991,** *94* (1), 30-38

36. Hediger, S.; Meier, B. H.; Kurur, N. D.; Bodenhausen, G.; Ernst, R. R., NMR cross polarization by adiabatic passage through the Hartmann—Hahn condition (APHH). *Chem. Phys. Lett.* **1994,** *223* (4), 283-288

37. Ernst, M.; Meier, B. H.; Tomaselli, M.; Pines, A., Time-reversal of cross-polarization in nuclear magnetic resonance. *J. Chem. Phys.* **1998,** *108* (23), 9611-9613.

38. Shekar, S. C.; Ramamoorthy, A., The unitary evolution operator for cross-polarization schemes in NMR. *Chem. Phys. Lett.* **2001,** *342* (1-2), 127-134.

39. Marica, F.; Snider, R. F., An analytical formulation of CPMAS. *Solid State Nucl. Magn. Reson.* **2003,** *23* (1-2), 28-49

40. Marks, D.; Vega, S., A Theory for Cross-Polarization NMR of Nonspinning and Spinning Samples. *J. Magn. Reson., Ser. A* **1996,** *118* (2), 157-172.

41. Chakraborty, A; Fernando, L.D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latge, J.P.; Wang, T. *Nat. Commun.* **2021** *12*, 6346.

42. Kang, X.; Kirui, A.; Muszyński, A.; Dickwella Widanage, M. C.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9* (1), 1-12

43. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR.

Nat. Commun. **2019,** *10* (1), 1-9

44. Heise, H.; Seidel, K.; Etzkorn, M.; Becker, S.; Baldus, M., 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations. *J. Magn. Reson.* **2005,** *173* (1), 64-74.

45. Abragam, A., *The principles of nuclear magnetism*. Oxford university press: 1961.

46. Slichter, C. P., *Principles of magnetic resonance*. Springer Science & Business Media: 2013; Vol. 1

47. Levitt, M. H., *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons: 2013.

48. Dirac, P. A. M., *The principles of quantum mechanics*. Oxford university press: 1981.

49. Shekar, S. C.; Rong, P.; Jerschow, A., Irreducible spherical tensor analysis of quadrupolar nuclei. *Chem. Phys. Lett.* **2008,** *464* (4-6), 235-239.

50. Jaccard, G.; Wimperis, S.; Bodenhausen, G., Multiple-quantum NMR spectroscopy of S= 3/2 spins in isotropic phase: a new probe for multiexponential relaxation. *J. Chem. Phys.* **1986,** *85* (11), 6282-6293

51. Shekar, S. C.; Tang, J. A.; Jerschow, A., Dynamics of I=3\over2 nuclei in isotropic slow motion, anisotropic and partially ordered phases. *Concepts Magn. Reson., Part A* **2010,** *36* (6), 362-387

52. States, D.; Haberkorn, R.; Ruben, D., A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson. (1969)* **1982,** *48* (2), 286-292.

53. Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. D., Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **1989,** *85* (2), 393-399.

54. Akutsu, H.; Kyogoku, Y., Phase transitions of phosphatidylcholine bilayers as revealed by cross-polarization efficiency of 31P-NMR. *Biochimica et Biophysica Acta* **1984,** *774* (2), 293-295.

55. Akutsu, H., Detection of slow-motional changes in the phosphatidylcholine bilayers by 1H-31P cross-polarization dynamics. *J. Magn. Reson. (1969)* **1986,** *66* (2), 250-263.

56. Salmon, A.; Wiedemann, T., The effect of lung surfactant apolipoproteins B/C on the chemical shift anisotropy of sn-2-[1-13C]dipalmitoylphosphatidylcholine. *Chem. Phys. Lipids* **1991,** *58* (1), 55-62.

57. Mavromoustakos, T.; Theodoropoulou, E., A combined use of 13C-cross polarization/magic angle spinning, 13C-magic angle spinning and 31P-nuclear magnetic resonance spectroscopy with differential scanning calorimetry to study cannabinoid-membrane interactions. *Chem. Phys. Lipids* **1998,** *92* (1), 37-52.

58. Tsvetkova, N. M.; Phillips, B. L.; Crowe, L. M.; Crowe, J. H.; Risbud, S. H., Effect of Sugars on Headgroup Mobility in Freeze-Dried Dipalmitoylphosphatidylcholine Bilayers: Solid-State 31P NMR and FTIR Studies. *Biophys. J.* **1998,** *75* (6), 2947-2955.

59. Harris, R. K.; Wasylishen, R. E.; Duer, M. J.; Shekar, S. C.; Jerschow, A., *Tensors in NMR in 'NMR crystallography'*. John Wiley & Sons: 2009; Vol. 4

CHAPTER 7: CARBOHYDRATE-AROMATIC INTERFACE AND MOLECULAR ARCHITECTURE OF LIGNOCELLULOSE

Research paper reprinted with permission: *Nat. Commun.* 13, 538 (2022)

Authors: Alex Kirui[#], Wancheng Zhao[#], Fabien Deligey[#], Hui Yang, Xue Kang, Frederic Mentink-Vigier, and Tuo Wang

[#]*Equal contribution*

7.1 Abstract

Plant cell walls constitute the majority of lignocellulosic biomass and serve as a renewable resource of biomaterials and biofuel. Extensive interactions between polysaccharides and the aromatic polymer lignin make lignocellulose recalcitrant to enzymatic hydrolysis, but this polymer network remains poorly understood. Here we interrogate the nanoscale assembly of lignocellulosic components in plant stems using solid-state nuclear magnetic resonance (NMR) and dynamic nuclear polarization (DNP) approaches. We show that the extent of glycan-aromatic association increases sequentially across grasses, hardwoods, and softwoods. Lignin principally packs with the xylan in a non-flat conformation via non-covalent interactions and partially binds the junction of flat-ribbon xylan and cellulose surface as a secondary site. All molecules are homogeneously mixed in softwoods; this unique feature enables water retention even around the hydrophobic aromatics. These findings unveiled the principles of polymer interactions underlying the heterogeneous architecture of lignocellulose, which will guide the rational design of more digestible plants and more efficient biomass-conversion pathways.

7.2 Introduction

With solar energy and carbon dioxide transformed into carbohydrate-rich cell walls, terrestrial plants constitute eighty percent of the biomass distributed in the biosphere[1]. The

secondary cell wall is a lignocellulosic composite deposited once the cellular expansion has ceased, which has evolved into a major source of biopolymers and biofuels[2, 3]. Lignification mechanically strengthens secondary walls, however, the presence of these intractable polyphenols and their association with carbohydrate components contributes to the biomass recalcitrance that renders the feedstock resistant to enzymatic hydrolysis during its conversion to liquid transportation fuel[4, 5]. To cost-effectively access structural polysaccharides for ethanol fermentation, vast efforts have been dedicated to tailoring plants to produce more digestible walls and optimizing deconstruction procedures in biorefineries[6-8]. These efforts have not yet reached the full potential due to our limited understanding of cell wall architecture.

The secondary cell wall is assembled by carbohydrate and aromatic constituents, with remarkable complexity and variability. Each elementary cellulose microfibril contains eighteen 1,4-β-glucan chains, which are held together by a hydrogen-bonding network[9, 10]. The exact organization of these glucan chains is unresolved, but recent density functional theory (DFT) calculations suggest a six-layered organization, likely with 2, 3, 4, 4, 3, and 2 chains in each layer (**Figure 7.1a**)[11]. Elementary microfibrils frequently coalesce, forming large fibrils that often span across tens of nanometers[12, 13]. Hemicelluloses, such as xylan, glucuronoxylan, arabinoxylan, and glucomannan, are highly variable in their monosaccharide composition and linkage pattern. Xylan is among the most found hemicelluloses, and its backbone comprises β-1,4-xylose units in a wide range of conformations, with substitutions by arabinose (Ara) or glucuronic acid (GlcA), and modifications by acetyl (Ac) groups. Lignin contains guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H) phenolic residues, which are interconnected by different types of covalent linkers such as β-O-4 ether-O-aryl, β-β' resinol, and β-5' phenylcoumaran[14, 15].

Conceptually, the mechanical scaffold of crystalline cellulose is dispersed in a matrix formed by hemicellulose and lignin[16]. Our understanding of cell wall organization is supported by many studies that employed diffraction methods to reveal the spatial arrangement of cellulose microfibrils, imaging techniques to map out cell wall meshes and the microscopic distribution of lignin, and solution NMR spectroscopy to identify lignin-carbohydrate linkages[17-19]. However, the interface between lignin and polysaccharides, the focus of this study, is not yet well understood. This is partially due to the hardly accessible length scale (angstrom to nanometer) and the requirement of both chemical and atomic resolutions. In addition, only a small number of molecules reside on this lignin-carbohydrate interface, which needs to be deconvoluted from the bulk of the cell wall. As both lignin and polysaccharide exist in the solid state, conventional separation methods often perturb their structures and interactions, making it difficult to investigate this polymer interface.

Recently, multidimensional solid-state NMR spectroscopy of *Arabidopsis* and *Zea mays* (maize) has spotlighted a structure-function relationship of the molecules involved in the lignocellulosic interface[20, 21]. Lignin tends to form hydrophobic and disordered nanodomains, the surface of which binds the xylan in a three-fold helical screw conformation (three sugar residues per helical turn: a non-flat structure; **Figure 7.1a**) through non-covalent interactions. The three-fold domain is connected to its two-fold flat-ribbon region, which is coating the smooth surface of cellulose microfibrils[22-25]. To generalize these structural principles, we need to examine other plant systems to evaluate three critical aspects: (i) the conformational bias of hemicellulose's function, (ii) the absence of cellulose-lignin contact, and (iii) the self-aggregating nature of aromatic polymers.

The combination of solid-state NMR and DNP methods has allowed us to unveil the structural and chemical principles underlying the formation of lignocellulosic materials. We investigated the ^{13}C-labeled stems of two hardwoods, eucalyptus (*Eucalyptus grandis*) and poplar (*Populus x canadensis*), and the softwood spruce (*Picea abies*). These plants are non-food energy candidates for the development of second-generation biofuels to reduce our dependence on grain crops[26]. Despite lignin's preference for binding non-flat xylan, direct contacts are also observed between the aromatics and the junctions of cellulose surface and flat-ribbon xylan, which is probably enforced by molecular crowding. In these woody plants, molecules experience a highly homogeneous mixing on the nanoscale, which inevitably involves polymer interpenetration and entanglement. Consequently, the counter-intuitive hydration of aromatics is observed, where water is retained and stabilized by the polysaccharides closely packed to lignin. These molecular insights have brought our understanding of lignocellulose architecture to an unprecedented level of detail, allowing us to envision better biomass-conversion schemes for sustainable energy production.

7.3 Results

7.3.1 Polymorphic Structure of Carbohydrates in Woody Stems

We obtained ^{13}C-labeled plant stems by providing poplar, eucalyptus (gum tree), and spruce with $^{13}CO_2$. Free from chemical treatments, the hydrated lignocellulosic materials were directly analyzed using solid-state NMR. Therefore, all biomolecules have fully retained their chemical structure and physical packing. Atomic-level information of polysaccharide structure was obtained using two-dimensional (2D) ^{13}C-^{13}C correlation experiments (**Figure S7.1**), which spotted 69 allomorphs of monosaccharide residues in three samples (**Tables S7.1** and **S7.2**). These sugar units, with a wide range of linkages and conformations, were mainly found in cellulose and

four hemicelluloses, including glucuronoxylan, arabinoxylan, galactoglucomannan, and a very low amount of xyloglucan (**Figure 7.1a**).

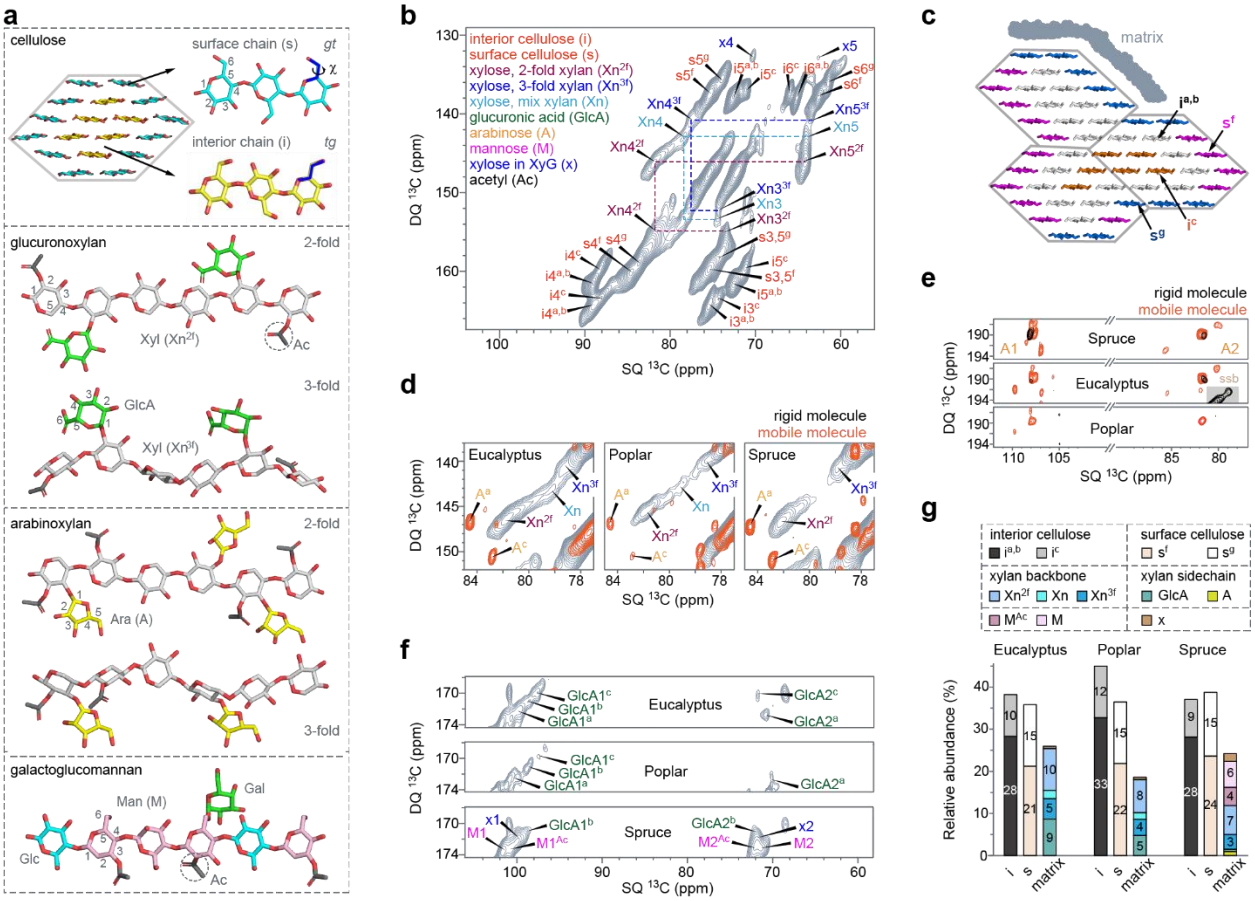

Figure 7.1 Solid-state NMR resolves the polymorphic structure of polysaccharides using intact wood cells. **a**, Representative structures of cellulose and major types of hemicellulose (xylan and galactoglucomannan). The cross-section of an elementary cellulose microfibril is shown, with a close view of an individual glucan chain and the hydroxymethyl conformation of surface (gauche-trans, gt) and internal (trans-gauche, tg) glucan chains. Xylan has two-fold or three-fold helical screw conformation in the backbone, with glucuronic acid (GlcA) or arabinose (Ara or A) sidechains. Acetyl motifs (Ac) are present in hemicelluloses. **b**, Representative 2D ^{13}C J-INADEQUATE spectrum of eucalyptus, which is based on ^{13}C cross polarization selecting rigid

Figure 7.1 (cont'd)

molecules. Dash lines track the carbon connectivity of xylan conformers. **c**, A model of three elementary cellulose microfibrils fitting the NMR observables. The hydrophilic (s^f) and hydrophobic (s^g) surfaces, the embedded chains (i^c), and the middle layer ($i^{a,b}$) are color-coded. **d**, Carbon-4 regions measured using ^{13}C direct polarization (orange spectra selecting mobile molecules) and cross polarization (grey spectra selecting rigid molecules). Both two- and three-fold xylan backbones are rigid while arabinose is mobile. **e**, Arabinose carbon-1 and carbon-2 regions. The only rigid arabinose is the terminal one in xylan sidechain. **f**, Eucalyptus has a high content of GlcA and spruce is rich in mannose (M). **g**, Composition of the rigid molecules in cell walls determined by peak volumes from 2D spectra. Source Data files are provided.

The remarkable resolution is evidenced by the narrow ^{13}C linewidths of 0.5-0.9 ppm, which allowed us to inspect the multifaceted polymorphism of polysaccharide structures. Locally, the structural variation of cellulose happens to the glucose hydroxymethyl conformations defined by the O5−C5−C6−O6 (χ) torsional angle (**Figure 7.1a**)[27]. The surface (s) and internal (i) chains primarily and respectively adopt gauche-trans (gt, χ=+60°) and trans-gauche (tg, χ=180°) conformations[27], which resulted in well-resolved signals (**Figure 7.1b**). Upon bundling, the averaged structure of a fibril was restrained using the number of glucan chains residing in different environments, including the hydrophilic (s^f) or hydrophobic (s^g) surface, the inaccessible core (i^c), and the middle layer ($i^{a,b}$) sandwiched in between (**Figure 7.1c**)[28]. A satisfactory fit includes three elementary microfibrils, with 15, 12, 6, and 21 chains for the s^f, s^g, i^c, and $i^{a,b}$ forms, respectively. Assuming uniform interfibrillar association, this averaged structure only designates the minimal bundle size, without considering loose packing. Other arrangements disagree with spectral observables (**Figure S7.2**).

Hemicellulose structure is highly complex as evidenced by the peak multiplicity. The backbones of hardwood xylan encompass two-fold (Xn^{2f}), three-fold (Xn^{3f}), and mixed (Xn) conformations (**Figure 7.1d**). The mixed form is absent in spruce; therefore, softwood xylan has higher homogeneity. Xylan was found to be rigid, suggestive of a close association with the stiff cellulose microfibrils. This observation differs from previous findings in *Arabidopsis*, where three-fold xylan remained mobile due to its spatial separation from cellulose[23]. Four types of arabinose (A) signals were identified, likely induced by the varied linkage sites at carbon 1, 2, 3, and 5 (**Figure 7.1e**). Only a single type is partially rigid, which is attributed to the terminal arabinose of xylan sidechains in secondary cell walls, whereas the mobile ones are from pectic polymers in primary walls. Hardwoods showed a high content of glucuronic (GlcA), while the softwood spruce exhibited unique signals of mannoses (M), one of the major constituents of galactoglucomannan (**Figure 7.1f**)[29].

Analysis of the spectral intensities led to the molar composition of rigid polysaccharides (**Figure 7.1g**). Around three quarters (74-81%) are cellulose, making it the most abundant polymer in woody stems (**Table S7.3**). Xylan makes up 18-26% of hardwood polysaccharides. Spruce has equal shares of xylan and mannan, each accounting for 10-12%. Consistent across these samples, the amount of two-fold xylan has doubled that of the three-fold counterpart, likely due to their promoted interactions with cellulose. Xylan sidechains are predominantly GlcA (therefore, glucuronoxylan) in hardwoods but mainly Ara (that is, arabinoxylan) in spruce.

7.3.2 Complex Structure and Linkage of Lignin

Wood lignin mainly contains guaiacyl (G) and syringyl (S) residues, with a single and two methoxyl groups, respectively (**Figure 7.2a, b**). The unsubstituted ring, *p*-hydroxyphenyl (H), was not observed in solid-state NMR spectra due to its low abundance in these plants. The aromatic

signals are dispersed over an extensive range of chemical shifts. For instance, four types of S/S' residues and four forms of G units were identified in eucalyptus and spruce, respectively. The multiplicity should be triggered by the varied oxidation states (for example, S' is a Cα-oxidized form of S unit), the assorted linkages, and presumably, the wide-ranging conformation and packing in native solids.

Aromatic rings are interconnected by a diverse array of covalent linkers. For example, β-ethers (A), with a characteristic one-bond cross peak between carbons β and γ at (85, 59) ppm (**Figure 7.2a**), have the β-O-4 linkages that are readily cleavable during degradation. Other signals emerged from resinol (B) that has β-β connections, phenylcoumaran (C) that encompasses β-5 and α-O-4 linkages and spirodienone (E) that contains β-1 and α-O-α links (**Figure 7.2c**). These signals are detectable by two experimental schemes based on ^{1}H-^{13}C cross polarization (CP) and ^{13}C direct polarization (DP) methods, indicating the distribution of these linkers in the rigid and mobile phases of lignin (**Figure S7.3**). Despite the low intensity, we managed to detect several cross peaks between these linkers and aromatic carbons (**Figure 7.2a**). For example, the connection between the carbon α of β-ethers and the carbon 1 of S residues yielded a cross peak at (72, 135) ppm in eucalyptus and poplar. Similar correlations were also found between the carbon 1 of S residues and the carbon α of B and E linkers. In spruce, analogous connections were observed for G units. These junctions typically evade solution-NMR characterization because the aromatic carbon 1 of lignin is non-protonated. These results demonstrated the feasibility of using solid-state NMR to characterize lignin structure and linkage, but the technical capability still requires further development.

Hardwood samples are rich in S residues (63-86 mol%) while the softwood only contains G units (**Figure 7.2d**). This observation agrees with our freshly collected solution NMR data of

ball-milled and dissolved lignocellulose (**Figure S7.4** and **Table S7.4**), average S/G ratios found in the literature (**Table S7.5**) as well as the quantification using deconvoluted 1D solid-state NMR spectra (**Figure S7.5** and **Table S7.6**). Both solid-state and solution NMR results (**Figure 7.2e**) suggest that these woody plants contain a large amount of β-ethers[30]. Solution NMR spectra also show considerable signals of both resinol and phenylcoumaran in poplar and spruce (**Figure 7.2e**), while the phenylcoumaran peaks become very weak in eucalyptus[15, 30-36]. Since these linkers respond differently to degradation, such analysis helps identify the plant candidates for saccharification.

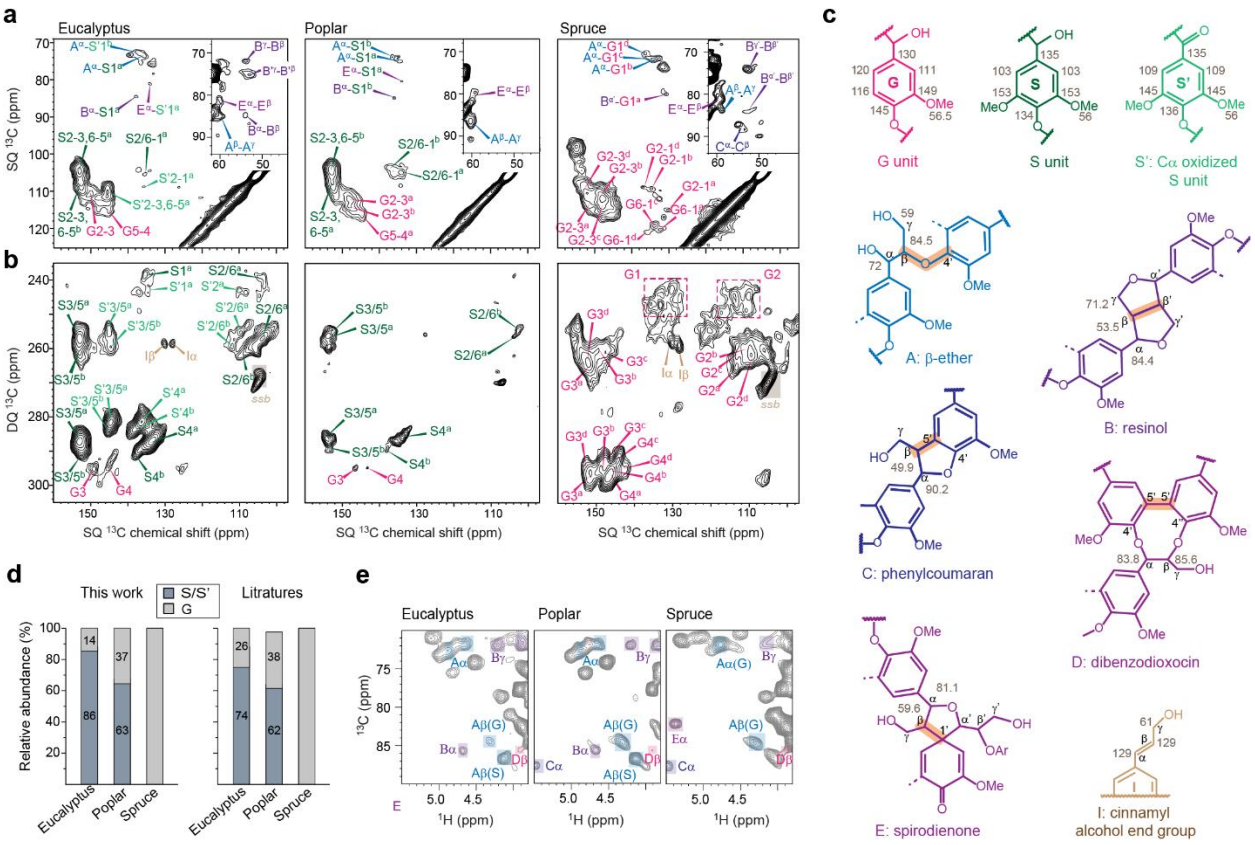

Figure 7.2 Chemical structure of lignin determined using solid-state and liquid NMR. Lignin signals are resolved using the aromatic regions of **a**, 2D ^{13}C-^{13}C RFDR spectra and **b**, 2D ^{13}C J-INADEQUATE spectra of intact plant cells of eucalyptus, poplar, and spruce. **c**, The structure of

Figure 7.2 (cont'd)

lignin units and linkages resolved in wood stems. The observed ^{13}C chemical shifts are labeled for each carbon site. Key inter-residual linkages are highlighted. **d**, Composition of lignin units quantified using the integration of one-bond cross peaks in 2D ^{13}C RFDR spectra (left panel) and using solution NMR as reported in the literature (right panel)[30, 32, 33, 35, 37-39]. Eucalyptus and poplar are S-rich while spruce is G-rich. Source Data files are provided. **e**, Molar fractions of the major linkage types quantified using the integration of one-bond cross peaks in 2D ^{13}C RFDR spectra. **e**, Representative lignin linkages detected by 2D ^{1}H-^{13}C HSQC spectra of ball-milled biomass in d_6-DMSO.

7.3.3 Distinct Patterns of Lignin-Carbohydrate Packing Across Plant Species

The supramolecular architecture of lignocellulose differs dramatically in hardwood and softwood. In both eucalyptus and poplar, the use of a long-mixing period (1.0 s) in 2D ^{13}C-^{13}C correlation spectra has generated many long-range intermolecular cross peaks that are absent in the short-mixing (0.1 s) spectrum (**Figure 7.3a** and **Figure S7.6**). Puzzlingly, these two spectra showed a comparable pattern in spruce, which signified that ^{13}C magnetization was already equilibrated among polymers within 0.1 s. This is not caused by variations of spin diffusion coefficients or polymer dynamics in hardwood and softwood as validated experimentally (**Figure S7.7** and **S7.8**). Therefore, the rapid equilibrium observed in spruce indicates that lignin and polysaccharides are well-mixed on the nanoscale in spruce but stay apart in hardwoods.

The spatial association of polymers was pinpointed by 272 intermolecular cross peaks, each of which represents a sub-nanometer, physical contact between two carbons from different molecules (**Figure 7.3b** and **Figure S7.9**). For example, the methyl carbon in the acetyl group of xylan (Ac^{Me}) exhibited cross peaks with the carbon 3 or 5 of S-lignin (S3/5) at (21, 153) ppm and

with the carbon 3 of G-lignin (G3) at (21, 148) ppm (**Figure 7.3b**). These interactions happen between i) acetyl groups and lignin, ii) xylan and lignin, iii) acetyl and cellulose, iv) acetyl and lignin methyl, v) lignin methoxyl group and cellulose, and vi) different lignin residues (**Figure S7.10**)[21].

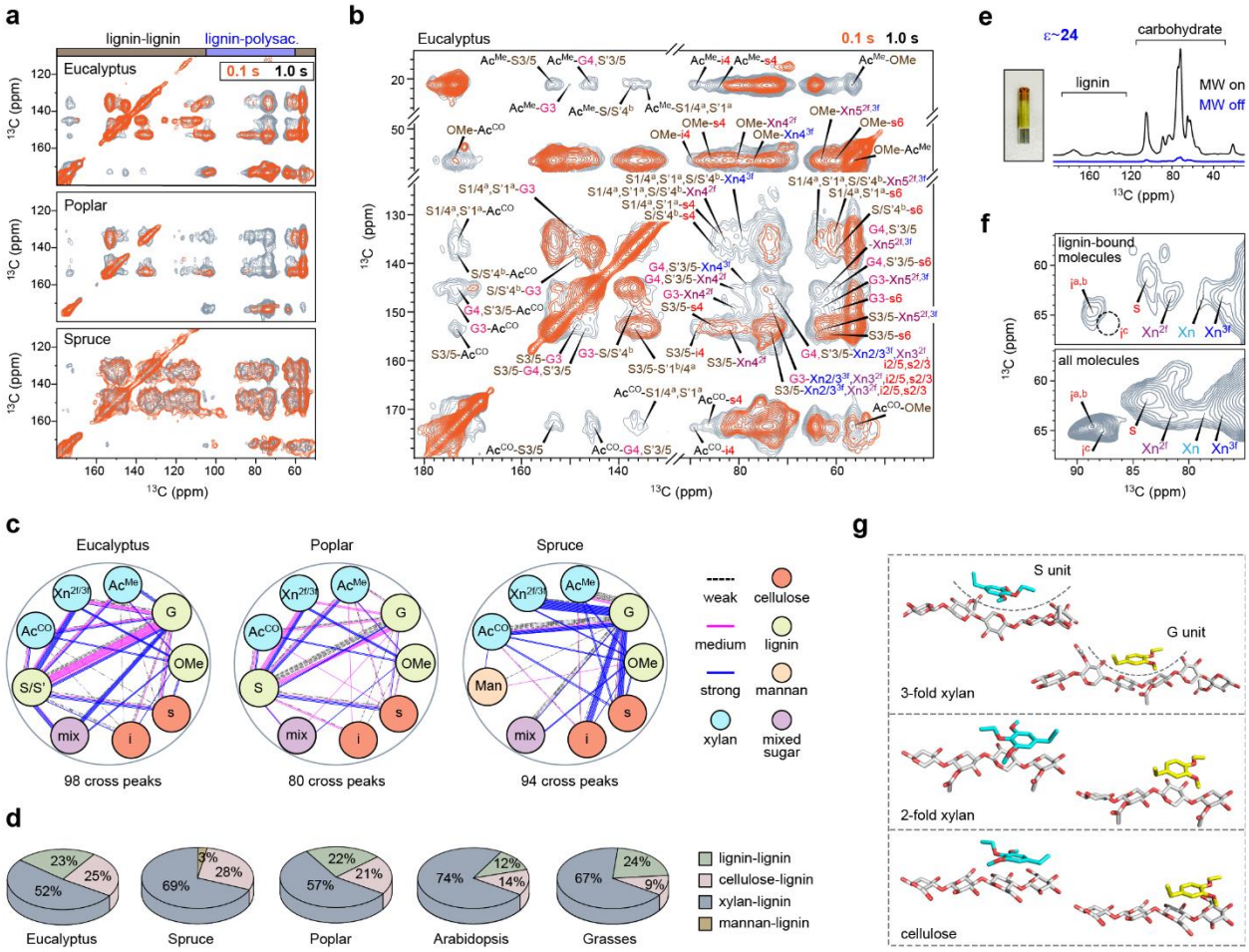

Figure 7.3 DNP-assisted detection of long-range interactions defines lignin-carbohydrate packing. **a**, Overlay of long-range (1.0 s mixing; grey) and short-range (0.1 s mixing; orange) 2D ^{13}C-^{13}C correlation spectra. The long- and short-range spectra showed similar spectral patterns in spruce: softwood polymers are homogeneously mixed on the molecular level. **b**, Dipolar-gated 2D ^{13}C-^{13}C correlation spectra showing 98 intermolecular cross peaks in eucalyptus. **c**, Summary of 272 intermolecular interactions identified in eucalyptus, poplar, and spruce. **d**, The count of

Figure 7.3 (cont'd)

lignin-lignin, cellulose-lignin, and xylan-lignin interactions in eucalyptus, poplar, spruce, *Arabidopsis*, and grasses (maize, rice, and switchgrass). Lignin-cellulose interactions are scarce in *Arabidopsis* and grasses but become abundant in woods. **e**, DNP enhancing NMR sensitivity of eucalyptus by 24-fold. Inset shows a picture of the plant material in a sapphire DNP-NMR rotor. **f**, Comparison of the equilibrium spectrum (bottom) of eucalyptus that detects all components with the aromatic-edited spectrum (top) that only shows lignin-bound molecules. The lignin-bound polysaccharides include the two-fold, three-fold, and mixed conformers of xylan as well as the surface and interior cellulose. The deeply embedded glucan chains in cellulose (i^c) are absent as highlighted using the dash line circle. **g**, DFT energy-minimized structures showing the possible packing between lignin units and polysaccharides.

A statistical view of the number and intensities of intermolecular contacts (**Tables S7.7-S7.10**) designated xylan as the major interactor with lignin, which is manifested by the extensive correlations between G/S and xylan carbons, including the carbonyl and methyl carbons of acetyl (Ac^{CO} and Ac^{Me}) and xylose ring carbons (**Figure 7.3c**). This result echoes with the recent findings in *Arabidopsis* and commelinid monocots (grasses, for example, maize and switchgrass)[24], validating the principal role of xylan-lignin interaction in stabilizing lignocellulose. Mannan has a small number of resolvable sites; therefore, it only showed a few cross peaks with lignin. However, the equilibrated pattern in **Figure 7.3a** supports a widespread association of mannan and lignin in spruce.

Unexpectedly, wood lignin also exhibited extensive correlations with cellulose (**Figure 7.3c**), which accounts for 20-30% of all intermolecular cross peaks (**Figure 7.3d**). Such interactions are particularly abundant in spruce, consistent with the homogeneous mixing of

molecules in this sample. In contrast, cellulose-lignin contacts only account for 9-14% of polymer interactions in *Arabidopsis* and grasses[24]. Thus, cellulose-lignin interaction is a unique feature of woody plants, which should contribute, at least in part, to the mechanical strengths of their stiff stems.

7.3.4 Visualization of the Polysaccharide-Lignin Interface

The unanticipated lignin-cellulose interactions were verified using the sensitivity-enhancing DNP technique[40-42]. The eucalyptus stem showed a 24-fold enhancement of NMR intensity (**Figure 7.3e**), which was achieved by transferring the polarization of electrons in the stable biradical AMUPol to the protons of biopolymers. This sensitivity enhancement shortens the measurements by 576 times (a 1.5-year experiment in a day), allowing us to take snapshots of the lowly populated boundaries between polysaccharides and lignin. Eucalyptus cellulose was found in the lignin-bound part of polysaccharides (**Figure 7.3f**), but such signals were missing in maize[24]. Only the surface glucan chains of cellulose (s) or those chains right underneath the surface layer ($i^{a,b}$) correlated with lignin, while the embedded core (i^{c}) lacked such interaction. The three-fold xylan (Xn^{3f}) is still favorable for binding lignin, showing stronger signals than the two-fold (Xn^{2f}) and the mixed (Xn) conformers.

Favorable lignin-polysaccharide interactions were revealed using quantum mechanical geometry optimizations conducted using the DFT method in continuum solvation models (**Figure 7.3g**). The three-fold xylan formed a packet to enclose aromatics. Conversely, the flat chains from cellulose surface or two-fold xylan relied on their pyranose/furanose rings to align with an aromatic unit. Examination of recently reported DFT structures[43] showed that S-units preferentially aligned to carbohydrate rings as stabilized by the two methoxyl groups on both sides of the aromatic ring (**Figure S7.11**). The G unit, on the other hand, typically has its single methoxyl

group closer to the carbohydrates. Such orientational preference supports the strong interactions experimentally observed between lignin methoxyl groups and xylan acetyls (**Figure S7.10**). Consequently, the methoxyl-rich S-residues correlate better with xylan in space. Therefore, non-covalent interactions between these polar groups are essential to the existence of lignin-xylan complex.

7.3.5 Landscape of Biomolecular Hydration and Dynamics

Biopolymers have sophisticated dynamics and variable water-association in their native environments. Among the three plants, poplar turned out to be the worst hydrated sample as shown by its slowest water-to-polymer ^{1}H polarization transfer buildup curves, which were consistent for both cellulose and xylan (i4 and Xn^{3f}1, **Figure 7.4a**), as well as lignin (**Figure S7.12**). The relative intensities (S/S_0) of a hydration heatmap reflect the degree of water association (**Figure 7.4b** and **Figure S7.13**). Plots of the S/S_0 ratios against 127 carbon sites show that the hydration level increases sequentially from poplar to eucalyptus and spruce (**Figure 7.4c**). Within each sample, polymer hydration generally increases from cellulose to xylan, and then to mannan, if present.

Surprisingly, wood lignin retained high S/S_0 ratios, typically on the same range as those of xylan, indicating a high level of hydration (**Figure 7.4c**). This is intriguing because these aromatic polymers are perceived as relatively hydrophobic, which by expectation, should hinder water retention. For example, maize lignin was previously found to self-aggregate, forming nanodomains that are largely separated from water[24]. The water contact observed here should originate from lignin's tight association with those carbohydrates that kept water localized. Likewise, wood biopolymers should experience frequent entanglement and interpenetration instead of domain separation.

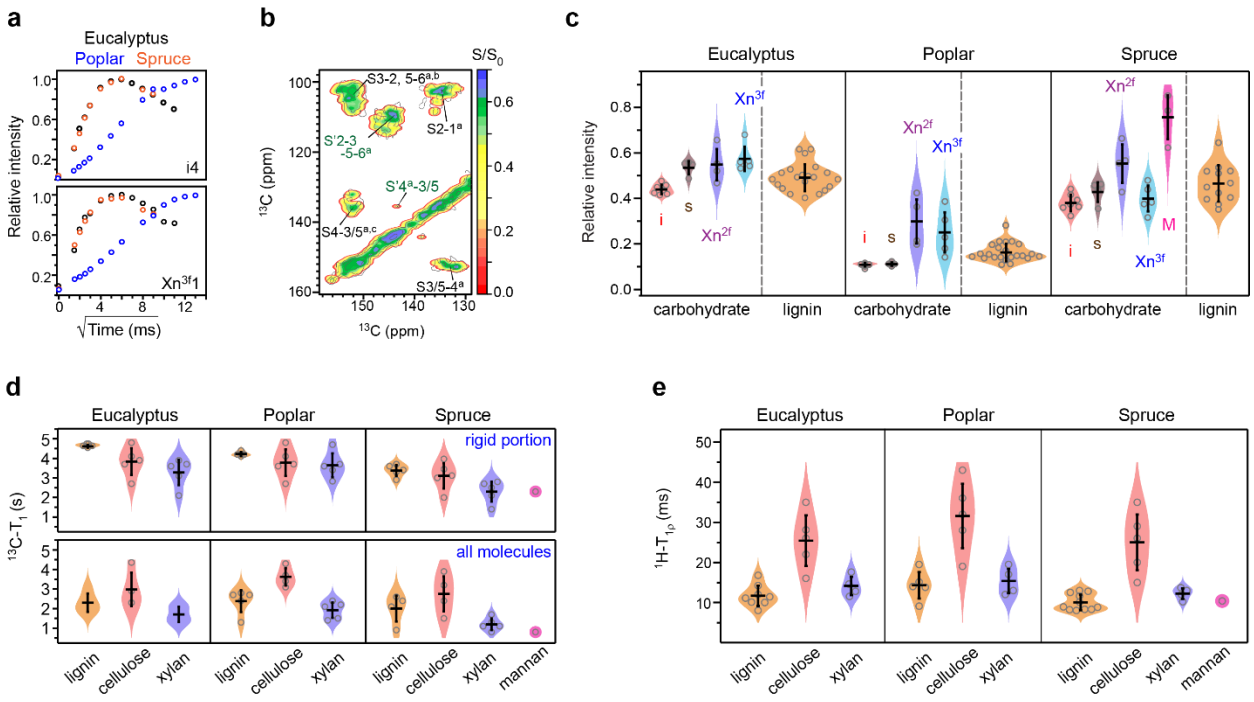

Figure 7.4 Site-specific hydration and dynamic landscape of biomolecules. **a**, Representative water-to-biomolecule ^{1}H polarization transfer buildup curves. The data of interior cellulose carbon 4 (i4) and three-fold xylan carbon 1 ($Xn^{3f}1$) are compared across plants. Poplar has slow buildup due to the limited water contact. **b**, Hydration map on top of a 2D spectrum plotting the ratio (S/S_0) of water-edited intensity (S) to the equilibrium intensity (S_0). A larger S/S_0 ratio indicates better polymer hydration. **c**, Distribution of the relative water-edited intensities (S/S_0) of polysaccharides and lignin in eucalyptus, poplar, and spruce. Molecules with better water retention show higher water-edited intensities. Poplar is poorly hydrated. **d**, ^{13}C-T_1 relaxation times of rigid (blue rectangles) and all molecules (colored as yellow, red, purple, and magenta for different molecules) in three woody plants, which represent nanosecond timescale motions. **e**, ^{1}H-$T_{1\rho}$ relaxation times reflecting microsecond timescale dynamics. The average value and standard deviation (error bars) are presented for each violin plot in panels **c-e**, the dataset of which are tabulated in **Tables S7.11-7.15**. Source Data files are provided.

Polymer dynamics were examined using ^{13}C-T_1 and ^{1}H-$T_{1\rho}$ relaxation measurements, which generated 150 relaxation curves (**Figure S7.14** and **S7.15**). When all molecules were considered, the ^{13}C-T_1 relaxation time decreased in the order of cellulose, lignin, xylan, and mannan, if any (**Figure 7.4d**). The short ^{13}C-T_1 time constants of lignin and hemicellulose revealed the efficient ^{13}C-T_1 relaxation in these non-cellulosic polymers, and furthermore, their enhanced motion on the nanosecond timescale. Such differences became indistinguishable if only rigid molecules were focused on, for example, all rigid molecules of eucalyptus showed ^{13}C-T_1 time constants of 4-5 s. This observation contradicts previous results in which maize lignin showed substantially longer ^{13}C-T_1 than any polysaccharide[24]. This can be comprehended using the efficient spin-exchange between lignin and carbohydrates (mediated by the better molecular mixing), which has averaged the ^{13}C-T_1 relaxation times in wood. On the microsecond timescale, cellulose had the longest ^{1}H-$T_{1\rho}$ relaxation times (30-40 ms) due to the restricted motion in the massively hydrogen-bonded microfibrils (**Figure 7.4e**). Both lignin and hemicellulose had short ^{1}H-$T_{1\rho}$ times of 10-15 ms, like the counterparts in maize[24]. Compared to G residues, S units exhibited slower ^{1}H-$T_{1\rho}$ relaxation, indicating attenuated dynamics due to interactions with polysaccharides (**Figure S7.15**). Among the three samples, spruce displayed the shortest ^{13}C-T_1 and ^{1}H-$T_{1\rho}$ relaxation times, revealing a unique profile of molecular dynamics in softwood.

7.3.6 Effects of Rehydration on Water-Retention and Polymer Dynamics

Water molecules are important for stabilizing the nanostructure of cell wall biomaterials, and the complete removal of water could potentially cause irreversible changes. In a recent NMR study of the softwood *Pinus radiata*, oven-drying (for complete dehydration) and rehydration were found to promote polymer association (e.g. xylan-cellulose and lignin-cellulose packing) and irreversibly alter the dynamics and conformation of mannan in the secondary cell walls[44]. Similarly,

the lyophilized-rehydrated grass sorghum (*Sorghum bicolor* L. Moench) has shown permanent changes of hemicellulose, with enhanced mobility for the three-fold xylan backbone and some arabinose residues[45]. Differently, dehydration only caused reversible structural changes to the primary cell walls of *Arabidopsis*, likely due to the high content of pectic polymers that can efficiently associate with water molecules[46].

We measured the never-dried samples of spruce and eucalyptus, followed by re-examination of the same materials after lyophilization and rehydration. In spruce, the lignocellulosic complex was not dramatically affected by the rehydration process, which differs from lipids and proteins (**Figure 7.5a**). The structure of the carbohydrates and lignin, as well as their distribution in dynamically distinct domains, can be efficiently retained in the rehydrated biomass, occasionally with minor intensity variations that are well below 10% for each carbohydrate or lignin carbon site (**Figure 7.5a**). The retention of water molecules can also be fully restored, as shown by the identical spectral patterns in 1D water-edited spectra (**Figure 7.5b**) and the comparable 2D heatmap representation for both carbohydrates and aromatics in the never-dried and rehydrated samples (**Figure 7.5c**). The nanosecond-timescale motion and the subnanometer polymer packing (reflected by spin-exchange) are unaffected since there is no deviation of ^{13}C-T_1 relaxation within the well-controlled error margin (**Figure 7.5d**). Both lignin and carbohydrates showed a 10-20% increase in the ^{1}H-$T_{1\rho}$ relaxation time constants; therefore, the only notable change happened to the microsecond scale motions of biopolymers (**Figure 7.5e**). Eucalyptus behaves differently: this plant fully restored the structural and dynamical features but moderately improved polymer-water association after rehydration (**Figure S7.16**).

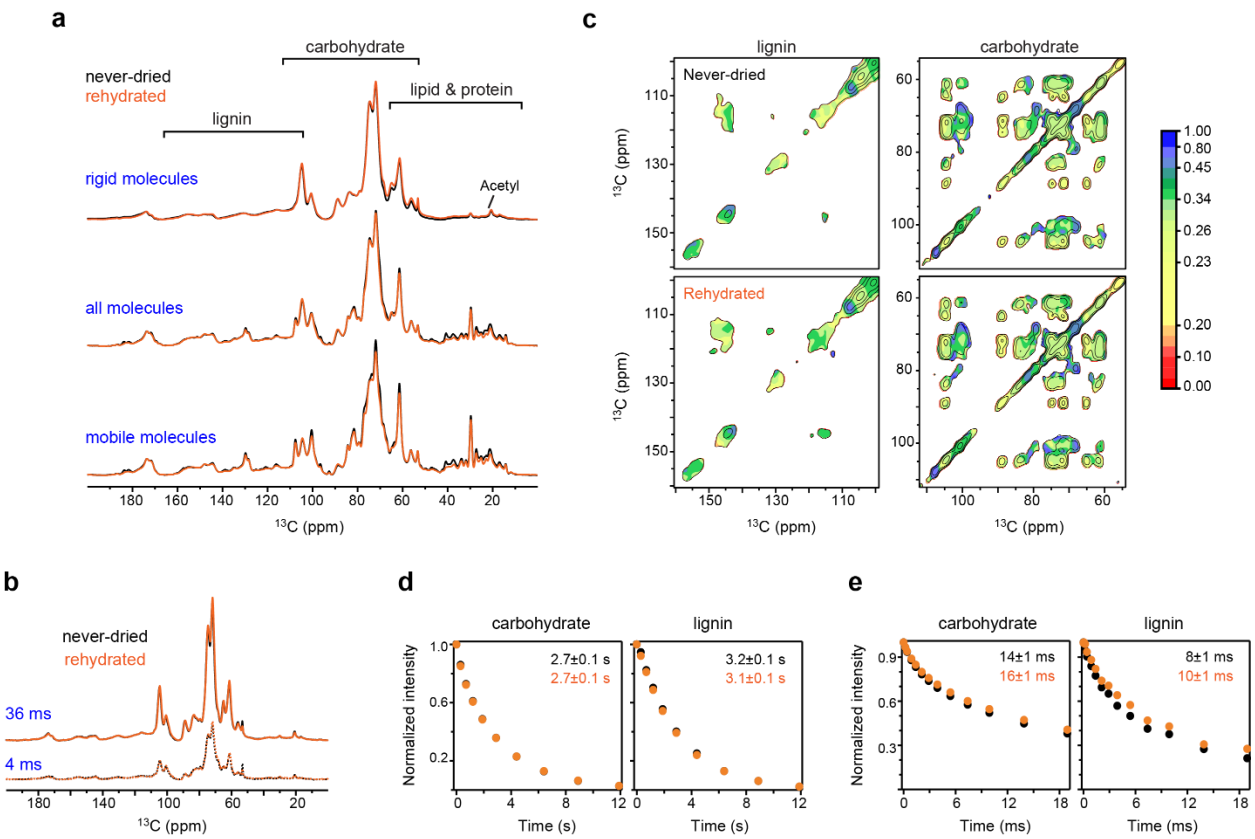

Figure 7.5 Molecular-level dynamics and water association after dehydration and rehydration. A never-dried spruce sample was first measured, and then freeze-dried and rehydrated for direct comparison. **a,** Overlay of spectra collected using never-dried (black) and lyophilized and then rehydrated (orange) spruce stem samples showing high similarity for polysaccharides and lignin. From top to bottom are the CP, quantitative DP, and short recycle-delay DP spectra. The key regions of lignin, carbohydrate, and lipid are marked. **b,** 1D water-to-carbohydrate polarization transfer spectra revealing similar water association for never-dried and rehydrated samples. The well-hydrated molecules are selected using a short ^{1}H mixing time of 4 ms while the equilibrated state is measured using 36 ms ^{1}H mixing. **c,** Water-edited 2D hydration map showing comparable water-edited intensities for both lignin and carbohydrate regions in the never-dried and the rehydrated samples. **d,** Average ^{13}C-T_1 relaxation plot, using the integrals of spectral regions for carbohydrates (left) and lignin (right). **e,** Average $T_{1\rho}$ relaxation of carbohydrates and lignin in

Figure 7.5 (cont'd)

never-dried (black) and rehydrated (orange) samples. Error bars are standard deviations of the fitting derived parameters.

7.4 Discussion

The abundant molecular-level evidence presented four novel features of lignocellulosic materials, each exploring a structural or chemical foundation of the supramolecular architecture (**Figure 7.6**). First, even though three-fold xylan is favored for binding lignin, in wood, other xylan conformers can also coexist with lignin in part. Accompanying with this functional resemblance is the hydration equivalence of these xylan forms (**Figure 7.4c**). These results have extended the conceptual model of lignocellulose derived from maize, in which lignin mainly interacts with the three-fold xylan, and *vice versa*. Actually, interactions between three-fold xylan and cellulose have been reported in a grass *Sorghum*[45], which also implies the interchangeable roles of xylan conformers. Uniquely, wood xylan is mainly in two-fold helical screw, which is energetically unfavorable unless forced by the binding to cellulose surface. Therefore, the unanticipated proximity of lignin to the two-fold xylan, as well as its associated cellulose surface, might be a destined consequence of spatial crowding in densely packed lignocellulosic materials.

Second, the packing between cellulose and lignin is plant-specific and only serves as a secondary interaction. In *Arabidopsis* and grasses, lignin and cellulose are spatially separated[24]. The situation has changed for woody biomass, where lignin and cellulose become colocalized. Although *Arabidopsis* is widely used as a model system for investigating the biophysical traits of hardwoods, it becomes apparent that these plants differ in their cell wall organization. Our DFT results and a recent molecular simulation study[47] consistently suggest that lignin is mainly docked on the hydrophobic surface of cellulose, with aligned phenyl and pyranose rings. The increased

coverage of cellulose surface by aromatic polymers might reduce enzyme accessibility and contribute to biomass recalcitrance.

Third, the structural feature of aromatic polymer needs reconsideration. Lignin nanodomains observed in maize are absent in woody plants due to the promoted polymer mixing on the nanoscale (**Figure 7.3**). Accordingly, polymer entanglements and inter-penetration, rather than superficial contact between domain surfaces, should govern lignin-carbohydrate interactions in wood[48]. Once well-mixed with polysaccharides, even the aromatics could effectively retain and immobilize water molecules (**Figure 7.4c**). The thermodynamical driving force of this peculiar phenomenon awaits inputs from modeling methods.

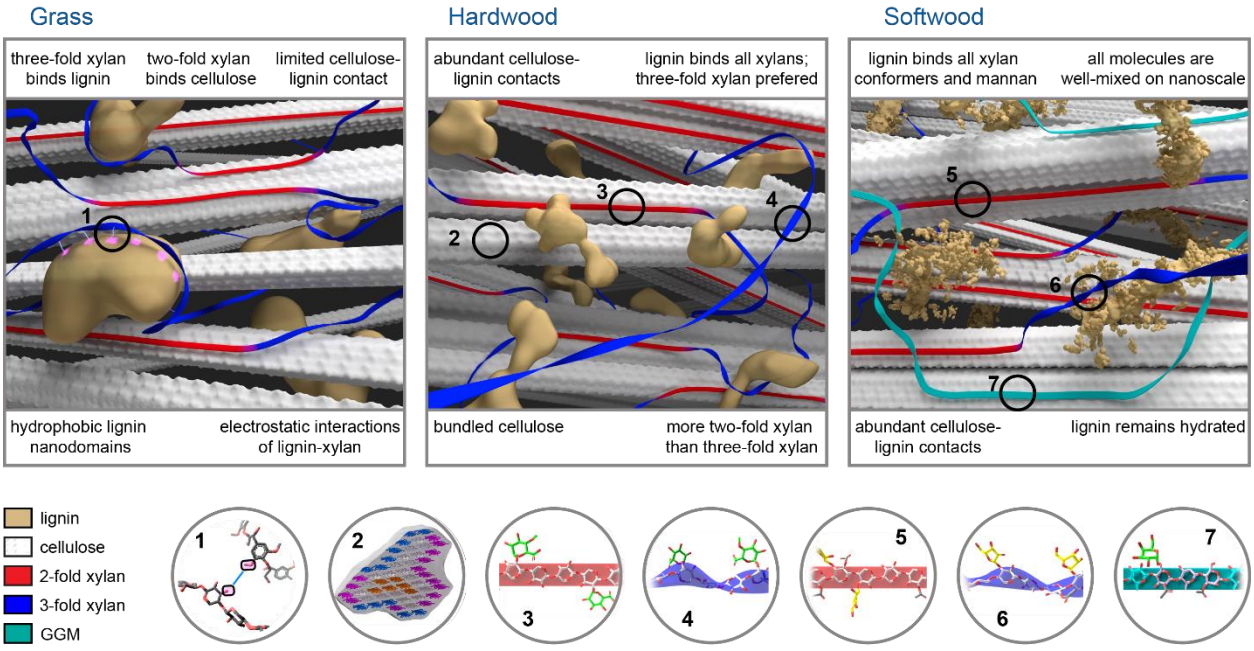

Figure 7.6 Comparative schemes of secondary cell walls in grass, hardwood, and softwood. The figure shows the spatial arrangement of lignin (yellow), cellulose (white fibrils), two-fold xylan (red flat ribbons), three-fold xylan (blue twisted ribbons), mixed forms of xylan (magenta), and mannan (GGM; green) in secondary plant cell walls. Numbered spheres highlight the structural features of (1) lignin-xylan interaction in all plants, (2) cellulose bundles in woody plants,

Figure 7.6 (cont'd)

(3-6) twofold and threefold xylan with different sidechains, and (7) galactoglucomannan. The molecular fraction of polysaccharides is considered in the depiction, but the illustration may not be strictly to scale. The model of grass cell wall is generated using the data recently reported[24], for comparison with the models of woody plants. Lignin and carbohydrates are much better mixed in woody plants than in grass, resulting in the binding of lignin to both three-fold and two-fold xylan as well as to cellulose microfibrils. The structural assembly of woody cell walls is thus different from the domain-separation scheme of grass cell walls.

Finally, the ultrastructure of softwood is featured by homogeneous molecular mixing (**Figure 7.3a**). This structural hallmark has revised the prevailing model in which well-aligned and partially crystalline glucomannans were perceived as being sandwiched between an internal cylinder of cellulose microfibrils and an outside tubular domain of lignin, with xylan forming an outermost layer[49]. Actually, our results better align with a recent study reporting that glucomannan and xylan could potentially associate with both cellulose and lignin[29].

Although the focus of this exploratory study is to understand the chemical and physical principles underlying polymer packing and lignocellulose architecture, the structural foundation and methodology established here would inspire more in-depth investigations for understanding the structural diversity and mechanical properties presented by numerous plant species and mutants, various cell types, and different growth stages[50]. These structural insights will guide the utilization of forestry resources for the production of biomaterials and biofuels[51], and the spectroscopic toolbox will stimulate structural investigations of polymer assemblies in other organisms, such as bacteria, fungi, and algae, as well as bio-inspired materials[52-60].

7.5 Methods

7.5.1 Plant Material.

Uniformly ^{13}C-enriched stems (97% ^{13}C) of eucalyptus (*Eucalyptus grandis*; age 16 weeks), poplar wood (*Populus x canadensis*; 27 weeks), and spruce (*Picea abies*; 16 weeks) were obtained for structural characterization from IsoLife (The Netherlands). Eucalyptus and spruce were obtained from seeds while poplar was obtained from stem cutting. Poplar plants were saplings of close to 1 m height at harvest. Immediately after removing the plants from the growth chamber, plant shoots were dissected into leaves and stems. The stem-sections were split, cut, and debarked after freeze-drying. Debarking was strived to be complete, which was conducted through the longitudinal cutting of the bark followed by separation from the xylem at the cambium interface. Debarked stems were hydrated for NMR analysis. The material was briefly hand-grinded using a pestle and a mortar, resulting in small pieces on the range of a few mm across. The homogeneity helps to avoid potential issues during magic-angle spinning. The technical details of isotope labeling are presented in the **Supplementary Methods**.

To test the effect of freeze-drying, we obtained never-dried ^{13}C-labeled spruce and eucalyptus wood from IsoLife, without lyophilization. The debarking process is finished manually by carefully removing the outside soft bark layer with tweezers. Around 104 mg of debarked spruce was cut into mm scale pieces, packed into 4-mm rotor, and measured under 10 kHz MAS on a 400 MHz Bruker NMR spectrometer. The never-dried samples were first measured freshly. After ssNMR measurements, overnight lyophilization was conducted, giving 43.5 mg of freeze-dried spruce stem. The sample is then sufficiently rehydrated (with a final weight of 107 mg), and all ssNMR experiments were conducted again on the rehydrated material for direct comparison with never-dried samples (**Figure 7.5**). Besides, never-dried (96 mg) eucalyptus stem, and the

dehydrated and then rehydrated material (94mg including 68 wt% H_2O) were also examined (**Figure S7.16**).

7.5.2 Solid-State NMR Experiments for Assignment.

The plant stems were packed into MAS rotors for measurements on 600 MHz and 400 MHz Bruker spectrometers using 3.2-mm and 4-mm MAS probes, respectively. Most experiments were collected under 10-14 kHz MAS at 293 K unless otherwise stated. ^{13}C chemical shifts were referenced to the tetramethylsilane (TMS) scale. Radiofrequency field strengths were 80-100 kHz for ^{1}H decoupling, 62.5 kHz for ^{1}H CP contact pulse, and 50–62.5 kHz for ^{13}C. The technical parameters of all solid-state NMR experiments were summarized in **Tables S7.16** and **S7.17** for the initial batch of rehydrated samples, and in **Table S7.18** for the more recently obtained never-dried and the subsequent freeze-dried and rehydrated samples.

To assign NMR signals, 2D correlation spectra were recorded using the refocused CP J-INADEQUATE sequence[61], which was coupled with ^{13}C DP for detecting mobile molecules or ^{1}H-^{13}C CP for detecting rigid polymers. A 2D ^{13}C radio frequency-driven recoupling (RFDR) correlation experiment was measured under 10 kHz MAS and 83 kHz ^{1}H decoupling to assign intra-residue cross peaks[62]. A recoupling time of 1.6 ms was used to detect one-bond ^{13}C–^{13}C cross peaks. An additional set of ^{13}C-DP PDSD experiments were conducted to examine the structure of lignin in the mobile phase, which were selected through the short recycle delay of 2 s. The standard flow chart of resonance assignment is presented in **Figure S7.17**. All chemical shifts are validated by comparison with literature-reported values (**Table S7.19**).

To enhance aromatic signals, we measured a dipolar-coupling-gated version of the proton-driven spin diffusion (PDSD) experiment[24, 63, 64]. A dipolar-dephasing period of 48 μs was employed to efficiently suppress the signals of protonated carbons, allowing the preferential

detection of non-protonated carbons in the aromatic lignin. This period was asymmetric with respect to the π pulse in the Hahn echo, containing two undecoupled delays of 32 μs and 16 μs. A 0.1-s PDSD mixing was used to report intramolecular cross peaks in all samples. A 20-ms mixing was also used for spruce. The observed chemical shifts were compared with values deposited in the Complex Carbohydrate Magnetic Resonance Database to facilitate resonance assignment[65].

7.5.3 Solid-State NMR Experiments for Structural Analysis.

The 2D gated PDSD experiment was also conducted to determine lignin-carbohydrate packing, with a long mixing time (1 s) for intermolecular polarization transfer. We have identified 98, 80, and 94 cross peaks for eucalyptus, poplar, and spruce, respectively. These 272 correlations were categorized as 112 strong, 75 medium, and 85 weak restraints according to their relative intensities (area of a peak relative to that of the 1D cross-section), with cutoffs set to >4.0% (strong), 2.0–4.0% (medium), and <2.0 (weak) as tabulated in **Tables S7.7-S7.10**.

Polymer hydration was determined using water-edited 2D ^{13}C–^{13}C correlation experiment at 277 K (**Figure 7.4 a, b**)[66, 67], which generated 44, 48, and 37 datasets for eucalyptus, poplar, and spruce, respectively. This experiment used a ^{1}H-T_2 relaxation filter of 120 μs × 2 to suppress the polysaccharide signals to <5% and retain >85% of water magnetization. Water-polarization was transferred to spatially proximal polymers using a 4-ms ^{1}H mixing period, followed by a 1-ms CP for ^{13}C detection. A 100-ms DARR mixing was used for both water-edited and control spectra. 1D buildup curves were measured at 277 K using a ^{1}H-T_2 filter of 120 μs × 2 and a ^{1}H mixing of 0-81 ms for spruce, 0-121 ms for eucalyptus, and 0- 169 ms for poplar.

To probe dynamics, we measured ^{13}C spin-lattice relaxation (T_1) and ^{1}H rotating frame spin-lattice relaxation ($T_{1\rho}$). ^{13}C-T_1 was measured using Torchia-CP [68] and standard ^{13}C DP inversion recovery schemes. Torchia-CP preferentially detected rigid molecules, with a z-filter of

0-12 s. A 30-s recycle delay was used in the DP inversion recovery experiment for quantitatively detecting all molecules, with a z-period of 0-35 s. $^{1}H–T_{1\rho}$ was measured using an effective ^{1}H spin-lock field of 62.5 kHz with a duration of 0-19 ms. Relaxation data were fit using single exponential functions (**Tables S7.14** and **S7.15**). Also, the dipolar-chemical shift (DIPSHIFT) experiment[69] was conducted under 7.5 kHz MAS, to report dipolar order parameters of biopolymers in three wood samples. Frequency-Switched Lee Goldburg (FSLG) ^{1}H homonuclear decoupling sequence was utilized, the scaling factor was verified to be 0.577.

7.5.4 Solution NMR Sample Preparation and Experiments.

To validate the lignin assignment obtained using the solid-state method, we conducted solution NMR analysis, which has well-documented chemical shifts of lignin polymers. A summary of the lignin chemical shifts is included in **Table S7.4**. Wood samples were ball-milled for 1 h with a motor running at 1750 rpm. About 30 mg of each powdered wood sample was dissolved in 10 mL DMSO-d_6 with constant stirring using a magnetic stirrer for 2 h at 60°C for solution NMR experiments. To identify linkages in lignin, we conducted 2D ^{1}H-^{13}C HSQC spectra conducted on a Bruker Avance III 500 MHz ^{1}H Larmor frequency and equipped with a 5 mm *z*-gradient Prodigy TCI probe. Both the ^{1}H and the ^{13}C chemical shifts gave indications of the linkage types using the chemical shifts reported previously[39, 70-74].

7.5.5 MAS-DNP Experiment.

About 30-35 mg of eucalyptus stems were impregnated into 50 µL DNP matrix solution containing 10 mM AMUPol[75] radical in d_8-glycerol/D_2O/H_2O (60/30/10 Vol%). The material was softly ground for 10-15 min to allow radical distribution and packed into a 3.2-mm sapphire rotor. The NMR sensitivity was enhanced by 24-fold with microwave irradiation on a 600 MHz/395 GHz MAS-DNP spectrometer under 10 kHz MAS. DNP analysis has been applied to various

carbohydrate and plant systems using different protocols[76-78]. Recently, we have demonstrated that the radical can effectively and homogeneously polarize all molecules in cell wall materials[79]. Lignin had a slightly shorter DNP buildup time (1.8 s) compared with polysaccharides (2.2 s). The recycle delay was set to 2.3 s to facilitate the selection of lignin against carbohydrates in the 2D experiments. A lignin-edited 2D ^{13}C-^{13}C experiment was measured to detect the signals of lignin-bound carbohydrates, with a 0.5-s PDSD mixing to transfer lignin polarization to carbohydrates, followed by a 20-ms PDSD mixing for 2D ^{13}C-^{13}C correlation[24]. For comparison, a 20-ms PDSD spectrum was measured as a control.

7.5.6 DFT Calculation.

DFT calculations were used to investigate the interactions between the S or G units in lignin and cellulose, or xylan with two- and three-fold conformations. Tetramers of β-1,4-linked glucose and xylose were built to represent cellulose and two-fold xylan, respectively. Hexamers of β-1,4-linked xylose were built to represent three-fold xylan, based on a crystal structure of three-fold xylopentaose (PDB ID: 1GNY).[80] Six models were constructed: G unit-cellulose, G unit-xylan (two-fold), G unit-xylan (three-fold), S unit-cellulose, S unit-xylan (two-fold), and S unit-xylan (three-fold). Previously, it has been shown that for lignin-cellulose and lignin-xylan interactions, the stacked configurations tend to have the largest interaction energies[43]. Therefore, the initial starting configurations of the six models were constructed by positioning the G/S units parallel to the surface of the glucose/xylose residues, ~4 Å away. The geometry optimizations were conducted using the DFT method, M05-2X/6-31+G*[81] with the integral equation formalism for the polarizable continuum model (IEFPCM)[82] solvation model in Gaussian 09[83]. Although other methods may be available, this exchange-correlation functional with dispersion corrections performs reasonably well for carbohydrate and carbohydrate-aromatic interactions[11, 84, 85].

7.6 Acknowledgments

The solid-state and solution NMR analyses of wood cell walls were conducted by AK, XK, and WZ, who were supported by the U.S. Department of Energy (grant no. DE-SC0021210). DFT calculation and compositional analysis conducted by HY and FD were supported as part of the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under award no. DE-SC0001090. *The National High Magnetic Field Laboratory is supported by National Science Foundation through NSF/DMR-1644779 and the State of Florida.* The MAS-DNP system at NHMFL is funded in part by NIH S10 OD018519 and NSF CHE-1229170. The authors thank Dr. Daniel Cosgrove for the helpful discussion.

REFERENCES

1. Bar-On, Y. M.; Phillips, R.; Milo, R., The biomass distribution on Earth. *P Natl Acad Sci USA* **2018,** *115* (25), 6506-6511.

2. Somerville, C.; Youngs, H.; Taylor, C.; Davis, S. C.; Long, S. P., Feedstocks for Lignocellulosic Biofuels. *Science* **2010,** *329* (5993), 790-792.

3. Moon, R. J.; Martini, A.; Nairn, J.; Simonsen, J.; Youngblood, J., Cellulose nanomaterials review: structure, properties and nanocomposites. *Chem Soc Rev* **2011,** *40* (7), 3941-3994.

4. Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D., Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007,** *315*, 804-805.

5. McCann, M. C.; Carpita, N. C., Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. *J. Exp. Bot.* **2015,** *66* (14), 4109-18.

6. Loque, D.; Scheller, H. V.; Pauly, M., Engineering of plant cell walls for enhanced biofuel production. *Curr Opin Plant Biol* **2015,** *25*, 151-161.

7. Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E., Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **2014,** *344* (6185), 709-+.

8. Cai, Y.; Zhang, K.; Kim, H.; Hou, G.; Zhang, X.; Yang, H.; Feng, H.; Miller, L.; Ralph, J.; Liu, C. J., Enhancing digestibility and ethanol yield of Populus wood via expression of an engineered monolignol 4-O-methyltransferase. *Nature communications* **2016,** *7*, 11989.

9. Purushotham, P.; Ho, R.; Zimmer, J., Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science* **2020,** *369* (6507), 1089-1094.

10. Jarvis, M., Chemistry: cellulose stacks up. *Nature* **2003,** *426* (6967), 611-612.

11. Yang, H.; Kubicki, J., A density functional theory study on the shape of the primary cellulose microfibril in plants: effects of C6 exocyclic group conformation and H-bonding. *Cellulose* **2020,** *27*, 2389-2402.

12. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *P Natl Acad Sci USA* **2011,** *108* (47), E1195-E1203.

13. Thomas, L. H.; Altaner, C. M.; Forsyth, V. T.; Mossou, E.; Kennedy, C. J.; Martel, A.; Jarvis, M. C., Nanostructural deformation of high-stiffness spruce wood under tension. *Scientific reports* **2021,** *11* (1), 453.

14. Vanholme, R.; Demedts, B.; Morreel, K.; Ralph, J.; Boerjan, W., Lignin biosynthesis and

structure. *Plant Physiol.* **2010,** *153* (3), 895-905.

15. Wang, H. M.; Ma, C. Y.; Li, H. Y.; Chen, T. Y.; Wen, J. L.; Cao, X. F.; Wang, X. L.; Yuan, T.-Q.; Sun, R.-C., Structural Variations of Lignin Macromolecules from Early Growth Stages of Poplar Cell Walls. *ACS Sustainable Chem. Eng.* **2020,** *8* (4), 1813-1822.

16. Cosgrove, D. J.; Jarvis, M. C., Comparative structure and biomechanics of plant primary and secondary cell walls. *Front Plant Sci* **2012,** *3*.

17. Du, X.; Perez-Boada, M.; Fernandez, C.; Rencoret, J.; del Rio, J. C.; Jimenez-Barbero, J.; Li, J.; Gutierrez, A.; Martinez, A. T., Analysis of lignin-carbohydrate and lignin-lignin linkages after hydrolase treatment of xylan-lignin, glucomannan-lignin and glucan-lignin complexes from spruce wood. *Planta* **2014,** *239*, 1079-1090.

18. Giummarella, N.; Pu, Y.; Ragauskas, A. J.; Lawoko, M., A critical review on the analysis of lignin carbohydrate bonds. *Green Chem.* **2019,** *21*, 1573-1595.

19. Zhang, T.; Vavylonis, D.; Durachko, D. M.; Cosgrove, D. J., Nanoscale movements of cellulose microfibrils in primary cell walls. *Nat Plants* **2017,** *3* (5).

20. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-State NMR Spectroscopy. *Nat. Rev. Methods Primers* **2021,** *1*, 2.

21. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: A critical review. *Solid state nuclear magnetic resonance* **2020,** *107*, 101660.

22. Grantham, N. J.; Wurman-Rodrich, J.; Terrett, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat Plants* **2017,** *3* (11), 859-865.

23. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature communications* **2016,** *7*, 13902.

24. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature communications* **2019,** *10* (1), 347.

25. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry-Us* **2015,** *54* (14), 2335-2345.

26. Simmons, B. A.; Loque, D.; Blanch, H. W., Next-generation biomass feedstocks for biofuel production. *Genome biology* **2008,** *9* (12), 242.

27. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct Determination of

Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

28. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

29. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nature communications* **2019,** *10* (1), 4978.

30. Rencoret, J.; Gutierrez, A.; Nieto, L.; Jimenez-Barbero, J.; Faulds, C. B.; Kim, H.; Ralph, J.; Martinez, A. T.; Del Rio, J. C., Lignin Composition and Structure in Young versus Adult Eucalyptus globulus Plants. *Plant Physiol.* **2011,** *155*, 667-682.

31. Rinaldi, R.; Jastrzebski, R.; Clough, M. T.; Ralph, J.; Kennema, M.; Bruijnincx, P. C. A.; Weckhuysen, B. M., Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis. *Angew Chem Int Edit* **2016,** *55* (29), 8164-8215.

32. Rico, A.; Rencoret, J.; del Rio, J. C.; Martinez, A. T.; Gutierrez, A., In-Depth 2D NMR Study of Lignin Modification During Pretreatment of Eucalyptus Wood with Laccase and Mediators. *Bioenergy Res.* **2015,** *8*, 211-230.

33. Ibarra, D.; Chavez, M. I.; Rencoret, J.; Del Rio, J. C.; Gutierrez, A.; Romero, J.; Camarero, S.; Martinez, M. J.; Jimenez-Barbero, J.; Martinez, A. T., Structural modification of eucalypt pulp lignin in a totally chlorine-free bleaching sequence including a laccase-mediator stage. *Holzforschung* **2007,** *61*, 634-646.

34. Li, M.; Cao, S.; Meng, X.; Studer, M.; Wyman, C. E.; Ragauskas, A. J.; Pu, Y., The effect of liquid hot water pretreatment on the chemical–structural alteration and the reduced recalcitrance in poplar. . *Biotechnology for biofuels* **2017,** *10*, 237.

35. Samuel, R.; Foston, M.; Jaing, N.; Cao, S.; Allison, L.; Studer, M.; Wyman, C.; Ragauskas, A. J., HSQC (heteronuclear single quantum coherence) 13C–1H correlation spectra of whole biomass in perdeuterated pyridinium chloride–DMSO system: An effective tool for evaluating *Fuel* **2011,** *90* (9), 2836-2842.

36. Samuel, R.; Cao, S.; Das, B. K.; Hu, F.; Pu, Y.; Ragauskas, A. J., Investigation of the fate of poplar lignin during autohydrolysis pretreatment to understand the biomass recalcitrance. *RSC Adv.* **2013,** *3* (16), 5305-5309.

37. Rencoret, J.; Marques, G.; Gutierrez, A.; Nieto, L.; Santos, J. I.; Jimenez-Barbero, J.; Martinez, A. T.; del Rio, J. C., HSQC-NMR analysis of lignin in woody (Eucalyptus globulus and Picea abies) and non-woody (Agave sisalana) ball-milled plant materials at the gel state 10th EWLP, Stockholm, Sweden, August 25–28, 2008. *Holzforschung* **2009,** *63*, 691-698.

38. Wen, J. L.; Sun, S. L.; Yuan, T. Q.; Xu, F.; Sun, R. C., Structural elucidation of lignin polymers of Eucalyptus chips during organosolv pretreatment and extended delignification.

Journal of agricultural and food chemistry **2013,** *61* (46), 11067-11075.

39. Yuan, T.-Q.; Sun, S.-N.; Xu, F.; Sun, R.-C., Characterization of Lignin Structures and Lignin–Carbohydrate Complex (LCC) Linkages by Quantitative ^{13}C and 2D HSQC NMR Spectroscopy. *Journal of agricultural and food chemistry* **2011,** *59*, 10604-10614.

40. Su, Y.; Andreas, L.; Griffin, R. G., Magic angle spinning NMR of proteins: high-frequency dynamic nuclear polarization and ^{1}H detection. *Annu. Rev. Biochem.* **2015,** *84*, 465-97.

41. Mentink-Vigier, F.; Akbey, U.; Oschkinat, H.; Vega, S.; Feintuch, A., Theoretical aspects of Magic Angle Spinning - Dynamic Nuclear Polarization. *J Magn Reson* **2015,** *258*, 102-120.

42. Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Coperet, C.; Emsley, L., Dynamic nuclear polarization surface enhanced NMR spectroscopy. *Accounts of chemical research* **2013,** *46* (9), 1942-51.

43. Yang, H.; Watts, H. D.; Gibilterra, V.; Weiss, T. B.; Petridis, L.; Cosgrove, D. J.; Kubicki, J. D., Quantum Calculations on Plant Cell Wall Component Interactions. *Interdisciplinary sciences, computational life sciences* **2019,** *11* (3), 485-495.

44. Cresswell, R.; Dupree, R.; Brown, S. P.; Pereira, C. S.; Skaf, M. S.; Sorieul, M.; Dupree, P.; Hill, S., Importance of Water in Maintaining Softwood Secondary Cell Wall Nanostructure. *Biomacromolecules* **2021,** *in press*.

45. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose-xylan interaction dominates in sorghum secondary cell walls. *Nature communications* **2020,** *11* (1), 6081.

46. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis thaliana Primary Cell Walls: Evidence from Solid-State NMR. *Plant Physiol.* **2015,** *168*, 871-884.

47. Vermaas, J. V.; Crowley, M. F.; Beckham, G. T., A Quantitative Molecular Atlas for Interactions Between Lignin and Cellulose. *ACS Sustainable Chem. Eng.* **2019,** *7* (24), 19570-19583.

48. Langan, P.; Petridis, L.; O'Neill, H. M.; Pingali, S. V.; Foston, M.; Nishiyama, Y.; Schulz, R.; Lindner, B.; Hanson, B. L.; Harton, S.; Heller, W. T.; Urban, V.; Evans, B. R.; Gnanakaran, S.; Ragauskas, A. J.; Smith, J. C.; Davison, B. H., Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chem.* **2014,** *16* (1), 63-68.

49. Terashima, N.; Yoshida, M.; Hafren, J.; Fukushima, K.; Westermark, U., Proposed supramolecular structure of lignin in softwood tracheid compound middle lamella regions. *Holzforschung* **2012,** *66* (8), 907-915.

50. Berglund, J.; Mikkelsen, D.; Flanagan, B. M.; Dhital, S.; Gaunitz, S.; Henriksson, G.; Lindstrom, M. E.; Yakubov, G. E.; Gidley, M. J.; Vilaplana, F., Wood hemicelluloses exert distinct biomechanical contributions to cellulose fibrillar networks. *Nature communications* **2020,**

11 (1), 4692.

51. Ramage, M. H.; Burridge, H.; Busse-Wicher, M.; Fereday, G.; Reynolds, T.; Shah, D. U.; Wu, G. L.; Yu, L.; Fleming, P.; Densley-Tingley, D.; Allwood, J.; Dupree, P.; Linden, P. F.; Scherman, O., The wood from the trees: The use of timber in construction. *Renew Sust Energ Rev* **2017,** *68*, 333-359.

52. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

53. Ehren, H. L.; Appels, F. V. W.; Houben, K.; Renault, M. A. M.; Wosten, H. A. B.; Baldus, M., Characterization of the cell wall of a mushroom forming fungus at atomic resolution using solid-state NMR spectroscopy. *Cell surface* **2020,** *6*, 100046.

54. Chrissian, C.; Camacho, E.; Kelly, J. E.; Wang, H.; Casadevall, A.; Stark, R. E., Solid-state NMR spectroscopy identifies three classes of lipids in Cryptococcus neoformans melanized cell walls and whole fungal cells. *The Journal of biological chemistry* **2020,** *295* (44), 15083-15096.

55. Martin, R. W.; Kelly, J. E.; Kelz, J. I., Advances in instrumentation and methodology for solid-state NMR of biological assemblies. *J. Struct. Biol.* **2019,** *206*, 73-89.

56. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206* (1), 66-72.

57. Mao, H.; Tang, J.; Chen, J.; Wang, J.; Hou, K.; Peng, Y.; Halat, D. M.; Xiao, L.; Zhang, R.; Lv, X.; Yang, A.; Cui, Y.; Reimer, J. A., Designing hierarchical nanoporous membranes for highly efficient gas adsorption and storage. *Sci. Adv.* **2020,** *6*, eabb0694.

58. Thongsomboon, W.; Serra, D. O.; Possling, A.; Hadjineophytou, C.; Hengge, R.; Cegelski, L., Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. *Science* **2018,** *359* (6373), 334-338.

59. Chakraborty, A.; Fernando, L. D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latge, J. P.; Wang, T., A molecular vision of fungal cell wall organization by functional genomics and solid-state NMR. *Nature communications* **2021,** *12*, 6346.

60. Poulhazan, A.; Dickwella Widanage, M. C.; Muszynski, A.; Arnold, A.; Warschawsi, D.; Parastoo, A.; Isabelle, M.; Wang, T., Identification and Quantification of Glycans in Whole Cells: Architecture of Microalgal Polysaccharides Described by Solid-State Nuclear Magnetic Resonance. *J Am Chem Soc* **2021**, in press.

61. Lesage, A.; Auger, C.; Caldarelli, S.; Emsley, L., Determination of through-bond carbon-carbon connectivities in solid-state NMR using the INADEQUATE experiment. *J Am Chem Soc* **1997,** *119* (33), 7867-7868.

62. Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W. G.; Lansbury, P. T.; Griffin, R.

G., Homonuclear radio frequency-driven recoupling in rotating solids. *J Chem Phys* **1998,** *108* (22), 9463-9479.

63. Williams, J. K.; Schmidt-Rohr, K.; Hong, M., Aromatic spectral editing techniques for magic-angle-spinning solid-state NMR spectroscopy of uniformly (13)C-labeled proteins. *Solid state nuclear magnetic resonance* **2015,** *72*, 118-26.

64. Schmidt-Rohr, K.; Fritzsching, K. J.; Liao, S. Y.; Hong, M., Spectral editing of two-dimensional magic-angle-spinning solid-state NMR spectra for protein resonance assignment and structure determination. *Journal of biomolecular NMR* **2012,** *54* (4), 343-353.

65. Kang, X.; Zhao, W.; Dickwella Widanage, M. C.; Kirui, A.; Ozdenvar, U.; Wang, T., CCMRD: A Solid-State NMR Database for Complex Carbohydrates. *Journal of biomolecular NMR* **2019,** *in press*.

66. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water-polysaccharide interactions in the primary cell wall of *Arabidopsis thaliana* from polarization transfer solid-state NMR. *J Am Chem Soc* **2014,** *136* (29), 10399-10409.

67. Ader, C.; Schneider, R.; Seidel, K.; Etzkorn, M.; Becker, S.; Baldus, M., Structural Rearrangements of Membrane Proteins Probed by Water-Edited Solid-State NMR Spectroscopy. *J Am Chem Soc* **2009,** *131* (1), 170-176.

68. Torchia, D. A., Measurement of Proton-Enhanced ^{13}C-T_1 Values by a Method Which Suppresses Artifacts. *J Magn Reson* **1978,** *30* (3), 613-616.

69. Munowitz, M. G.; Griffin, R. G.; Bodenhausen, G.; Huang, T. H., Two-dimensional rotational spin-echo nuclear magnetic-resonance in solids: correlation of chemical-shift and dipolar interactions. *J Am Chem Soc* **1981,** *103*, 2529-2533.

70. Terashima, N.; Akiyama, T.; Ralph, S.; Evtuguin, D.; Neto, C. P.; Parkas, J.; Paulsson, M.; Westermark, U.; Ralph, J., 2D-NMR (HSQC) difference spectra between specifically C-13-enriched and unenriched protolignin of Ginkgo biloba obtained in the solution state of whole cell wall material. *Holzforschung* **2009,** *63* (4), 379-384.

71. Yelle, D. J.; Ralph, J.; Frihart, C., Characterization of nonderivatized plant cell walls using high-resolution solution-state NMR spectroscopy. *Magn. Reson. Chem.* **2008,** *46*, 508-517.

72. Mansfield, S. D.; Kim, H.; Lu, F. C.; Ralph, J., Whole plant cell wall characterization using solution-state 2D NMR. *Nat Protoc* **2012,** *7* (9), 1579-1589.

73. Grabber, J. H.; Ralph, J.; Hatfield, R. D., Model studies of ferulate-coniferyl alcohol cross-product formation in primary maize walls: implications for lignification in grasses. *Journal of agricultural and food chemistry* **2002,** *50* (21), 6008-16.

74. Grabber, J. H.; Ralph, J.; Hatfield, R. D., Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *Journal of agricultural and food chemistry* **2000,** *48* (12), 6106-13.

75. Sauvee, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew Chem Int Edit* **2013,** *52* (41), 10858-10861.

76. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paepe, G., Rapid Natural-Abundance 2D 13C-13C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew Chem Int Edit* **2012,** *51* (47), 11766-11769.

77. Perras, F. A.; Luo, H.; Zhang, X.; Mosier, N. S.; Pruski, M.; Abu-Omar, M. M., Atomic-Level Structure Characterization of Biomass Pre- and Post-Lignin Treatment by Dynamic Nuclear Polarization-Enhanced Solid-State NMR. *The journal of physical chemistry. A* **2017,** *121* (3), 623-630.

78. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of Pretreated Wood Fibers Using Dynamic Nuclear Polarization. *J Phys Chem C* **2019,** *123* (50), 30407-30415.

79. Zhao, W.; Kirui, A.; Deligey, F.; Mentink-Vigier, F.; Zhou, Y.; Zhang, B.; Wang, T., Solid-state NMR of unlabeled plant cell walls: high-resolution structural analysis without isotopic enrichment. *Biotechnology for biofuels* **2021,** *14* (1), 14.

80. Szabo, L.; Jamal, S.; Xie, H.; Charnock, S. J.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J., Structure of a family 15 carbohydrate-binding module in complex with xylopentaose. Evidence that xylan binds in an approximate 3-fold helical conformation. *The Journal of biological chemistry* **2001,** *276* (52), 49061-5.

81. Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A., 6-31G* basis set for third-row atoms. *J. Comput. Chem.* **20001,** *22*, 976-984.

82. Cances, E.; Mennucci, B.; Tomasi, J., A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. . *J Chem Phys* **1997,** *107*, 3032-3041.

83. Frisch, M. e. a., (Gaussian, Inc., Wallingford CT, 2009).

84. Raju, R. K.; Ramraj, A.; Hillier, I. H.; Vicent, M. A.; Burton, N. A., Carbohydrate–aromatic π interactions: a test of density functionals and the DFT-D method. *Phys. Chem. Chem. Phys.* **2009,** *11*, 3411-3416.

85. Yang, H.; Zimmer, J.; Yingling, Y. G.; Kubicki, J. D., How Cellulose Elongates-A QM/MM Study of the Molecular Mechanism of Cellulose Polymerization in Bacterial CESA. *J. Phys. Chem. B* **2015,** *119*, 6525-6535.

86. Smakman, G.; Rinie, H., Energy metabolism of Plantago lanceolata, as affected by change in root temperature *Physiol. Plant.* **1982,** *56*, 33-37.

87. De Visser, R.; Vianden, H.; Schnyder, H., Kinetics and relative significance of remobilized

and current C and N incorporation in leaf and root growth zones of Lolium perenne after defoliation: assessment by 13C and 15N steady-state labelling. . *Plant Cell Environ.* **1997,** *20*, 37-46.

88. Wang, H. M.; Ma, C. Y.; Li, H. Y.; Chen, T. Y.; Wen, J. L.; Cao, X. F.; Wang, X. L.; Yuan, T.-Q.; Sun, R.-C., Structural Variations of Lignin Macromolecules from Early Growth Stages of Poplar Cell Walls. *ACS Sustainable Chem. Eng.* **2020,** *8* (4), 1813-1822.

89. Rencoret, J.; Gutierrez, A.; Nieto, L.; Jimenez-Barbero, J.; Faulds, C. B.; Kim, H.; Ralph, J.; Martinez, A. T.; Del Rio, J. C., Lignin Composition and Structure in Young versus Adult Eucalyptus globulus Plants. *Plant Physiol.* **2011,** *155*, 667-682.

90. Purushotham, P.; Ho, R.; Zimmer, J., Architecture of a catalytically active homotrimeric plant cellulose synthase complex. *Science* **2020,** *369* (6507), 1089-1094.

91. Yang, H.; Kubicki, J., A density functional theory study on the shape of the primary cellulose microfibril in plants: effects of C6 exocyclic group conformation and H-bonding. *Cellulose* **2020,** *27*, 2389-2402.

92. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

93. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

94. Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calve, S.; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z. H.; Hoatson, G., Modelling one- and two-dimensional solid-state NMR spectra. *Magn Reson Chem* **2002,** *40* (1), 70-76.

95. Yang, H.; Watts, H. D.; Gibilterra, V.; Weiss, T. B.; Petridis, L.; Cosgrove, D. J.; Kubicki, J. D., Quantum Calculations on Plant Cell Wall Component Interactions. *Interdisciplinary sciences, computational life sciences* **2019,** *11* (3), 485-495.

96. Rencoret, J.; Marques, G.; Gutierrez, A.; Nieto, L.; Santos, J. I.; Jimenez-Barbero, J.; Martinez, A. T.; del Rio, J. C., HSQC-NMR analysis of lignin in woody (Eucalyptus globulus and Picea abies) and non-woody (Agave sisalana) ball-milled plant materials at the gel state 10th EWLP, Stockholm, Sweden, August 25–28, 2008. *Holzforschung* **2009,** *63*, 691-698.

97. Rico, A.; Rencoret, J.; del Rio, J. C.; Martinez, A. T.; Gutierrez, A., In-Depth 2D NMR Study of Lignin Modification During Pretreatment of Eucalyptus Wood with Laccase and Mediators. *Bioenergy Res.* **2015,** *8*, 211-230.

98. Ibarra, D.; Chavez, M. I.; Rencoret, J.; Del Rio, J. C.; Gutierrez, A.; Romero, J.; Camarero, S.; Martinez, M. J.; Jimenez-Barbero, J.; Martinez, A. T., Structural modification of eucalypt pulp lignin in a totally chlorine-free bleaching sequence including a laccase-mediator stage. *Holzforschung* **2007,** *61*, 634-646.

99. Wen, J. L.; Sun, S. L.; Yuan, T. Q.; Xu, F.; Sun, R. C., Structural elucidation of lignin polymers of Eucalyptus chips during organosolv pretreatment and extended delignification. *Journal of agricultural and food chemistry* **2013,** *61* (46), 11067-11075.

100. Yuan, T.-Q.; Sun, S.-N.; Xu, F.; Sun, R.-C., Characterization of Lignin Structures and Lignin–Carbohydrate Complex (LCC) Linkages by Quantitative ^{13}C and 2D HSQC NMR Spectroscopy. *Journal of agricultural and food chemistry* **2011,** *59*, 10604-10614.

101. Li, M.; Cao, S.; Meng, X.; Studer, M.; Wyman, C. E.; Ragauskas, A. J.; Pu, Y., The effect of liquid hot water pretreatment on the chemical–structural alteration and the reduced recalcitrance in poplar. . *Biotechnology for biofuels* **2017,** *10*, 237.

102. Samuel, R.; Foston, M.; Jaing, N.; Cao, S.; Allison, L.; Studer, M.; Wyman, C.; Ragauskas, A. J., HSQC (heteronuclear single quantum coherence) 13C–1H correlation spectra of whole biomass in perdeuterated pyridinium chloride–DMSO system: An effective tool for evaluating *Fuel* **2011,** *90* (9), 2836-2842.

103. Samuel, R.; Cao, S.; Das, B. K.; Hu, F.; Pu, Y.; Ragauskas, A. J., Investigation of the fate of poplar lignin during autohydrolysis pretreatment to understand the biomass recalcitrance. *RSC Adv.* **2013,** *3* (16), 5305-5309.

104. Lagerquist, L.; Pranovich, A.; Sumerskii, I.; von Schoultz, S.; Vahasalo, L.; Willfor, S.; Eklund, P., Structural and thermal analysis of softwood lignins from a pressurized hot water extraction biorefinery process and modified derivatives. . *Molecules* **2019,** *24* (2), 335.

105. Michalak, L.; Knutsen, S. H.; Aarum, I.; Westereng, B., Effects of pH on steam explosion extraction of acetylated galactoglucomannan from Norway spruce. *Biotechnology for biofuels* **2018,** *11* (1), 1-12.

106. Du, X.; Perez-Boada, M.; Fernandez, C.; Rencoret, J.; del Rio, J. C.; Jimenez-Barbero, J.; Li, J.; Gutierrez, A.; Martinez, A. T., Analysis of lignin-carbohydrate and lignin-lignin linkages after hydrolase treatment of xylan-lignin, glucomannan-lignin and glucan-lignin complexes from spruce wood. *Planta* **2014,** *239*, 1079-1090.

107. Wang, T.; Salazar, A.; Zabotina, O.; Hong, M., Structure and Dynamics of Brachypodium Primary Cell Walls Polysaccharides from 2D 13C Solid-State NMR Spectroscopy. *Biochemistry-Us* **2014,** (53), 2840-2854.

108. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature communications* **2019,** *10* (1), 347.

109. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose-xylan interaction dominates in sorghum secondary cell walls. *Nature communications* **2020,** *11* (1), 6081.

110. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell

walls revealed by solid-state NMR. *Nature communications* **2016,** *7*, 13902.

111. Duan, P.; Kaser, S. J.; Lyczakowski, J. J.; Phyo, P.; Tryfona, T.; Dupree, P.; Hong, M., Xylan Structure and Dynamics in Native Brachypodium Grass Cell Walls Investigated by Solid-State NMR Spectroscopy. *ACS Omega* **2021,** *6* (23), 15460-15471.

112. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nature communications* **2019,** *10* (1), 4978.

113. Phyo, P.; Wang, T.; Xiao, C.; Anderson, C.; Hong, M., Effects of Pectin Molecular-Weight Changes on the Structure, Dynamics and Polysaccharide Interactions of Primary Cell Walls of Arabidopsis thaliana: Insights from Solid-State NMR. *Biomacromolecules* **2017,** *18*, 2937-2950.

114. Ibarra, D.; Chávez, M. I.; Rencoret, J.; Del Río, J. C.; Gutiérrez, A.; Romero, J.; Camarero, S.; Martínez, M. J.; Jiménez-Barbero, J.; Martínez, A. T., Lignin Modification during Eucalyptus globulus Kraft pulping followed by totally chlorine-free bleaching: A two-dimensional nuclear magnetic resonance, fourier transform infrared, and pyrolysis− gas chromatography/mass spectrometry study. *Journal of agricultural and food chemistry* **2007,** *55* (9), 3477-3490.

115. Rencoret, J.; Neiva, D.; Marques, G.; Gutiérrez, A.; Kim, H.; Gominho, J.; Pereira, H.; Ralph, J.; Del Río, J. C., Hydroxystilbene glucosides are incorporated into Norway spruce bark lignin. . *Plant Physiol.* **2019,** *180*, 1310-1321.

APPENDIX

S7.1 Supplementary Methods

S7.1.1 Isotope Labeling.

Plant samples from eucalyptus, poplar, and spruce were produced under identical growth conditions in custom-designed, air-tight, high-irradiance labeling chambers of the Experimental Soil Plant Atmosphere System at IsoLife (Wageningen, The Netherlands). Under regulated environmental conditions, these plants were grown hydroponically in a closed atmosphere containing 97 atom% $^{13}CO_2$ from germinated seed or rooted stem cutting till harvest: photosynthetic photon flux density (PPFD) 800 µmol m^{-2} s^{-1} (top of plants); CO_2 concentration (Day); 400 ppm (v/v); 15 h (eucalyptus), 21 h (poplar), and 16 h (spruce) day lengths; day/night temperature of 22/16 °C (eucalyptus), 22/15 °C (poplar), and 24/20 °C (spruce). The day/night relative humidity is 75/75% for eucalyptus, 70/80% for poplar, and 70/75% for spruce. Minerals and water were supplied as aerated modified Hoagland nutrient solutions with micronutrients and iron[86,87], maintaining nitrogen concentration between 25 and 200 mg/L; pH close to 5; EC between 0.4 – 0.7 mS/cm; 25% of total nitrogen was supplied as ammonium. Plant shoots were dissected into leaves and stems after removing the plants from the growth chamber. After freeze-drying, these plant stems were kept stored at -20°C, in the dark with silica-gel drying bags. All plants were rehydrated for solid-state NMR analysis.

S7.1.2 Effects of Plant Age on Cell Walls.

Given our ever-increasing knowledge on plant secondary cell wall, impacts of the plant age at which the material is harvested relative to its natural growth cycle can be considered. Our 6.5-months old poplar can be put in perspective of 3-, 6- and 18-month-old poplar samples reported in literature[88]. Overall, there is no significant change in cellulose to hemicellulose proportions, and

cellulose crystallinity has become a stable parameter after 6 months. However, a slight increase in lignin content with age has been consistently reported for both poplar and eucalyptus[89]. Solution NMR data have demonstrated that lignin structure undergoes substantial evolution over a longer period of time, in both unit nature. For example, S lignin is deposited at later stages and the proportions of different linkage types. In-depth analyses of polysaccharide structure, lignin composition and linkage types, the consequential effect on lignocellulose architecture is thus of significant interest for future solid-state NMR research.

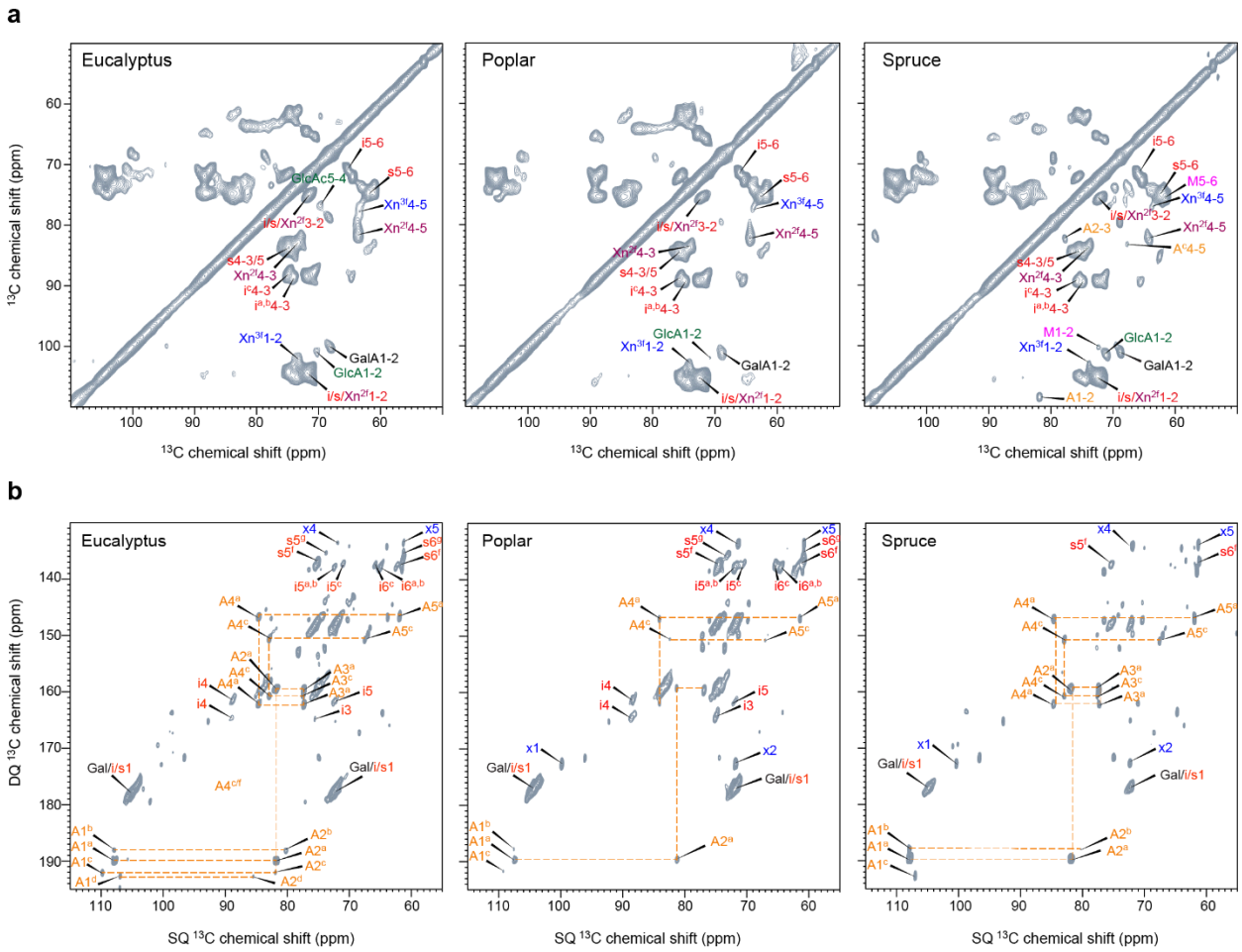

Figure S7.1 2D ^{13}C-^{13}C spectra of polysaccharides in softwood and hardwood. **a**, 2D ^{13}C-^{13}C RFDR spectra showing the rigid components of eucalyptus, poplar, and spruce. Uniquely, spruce has mannan and arabinose signals in the CP-based RFDR spectrum. **b**, 2D ^{13}C DP J-INADEQUATE spectra of eucalyptus, poplar, and spruce. The combination of ^{13}C DP with a short recycle delay of 2 s selectively detects the mobile molecules. Arabinose (A) signals dominate the spectra, and the four major types of arabinose are distinguished by the superscript, from A^a to A^d. The carbon connectivity of arabinose is highlighted by yellow lines. Poplar has relatively weak signals of arabinose. The ^{13}C chemical shifts are summarized in Tables S7.1 and S7.2.

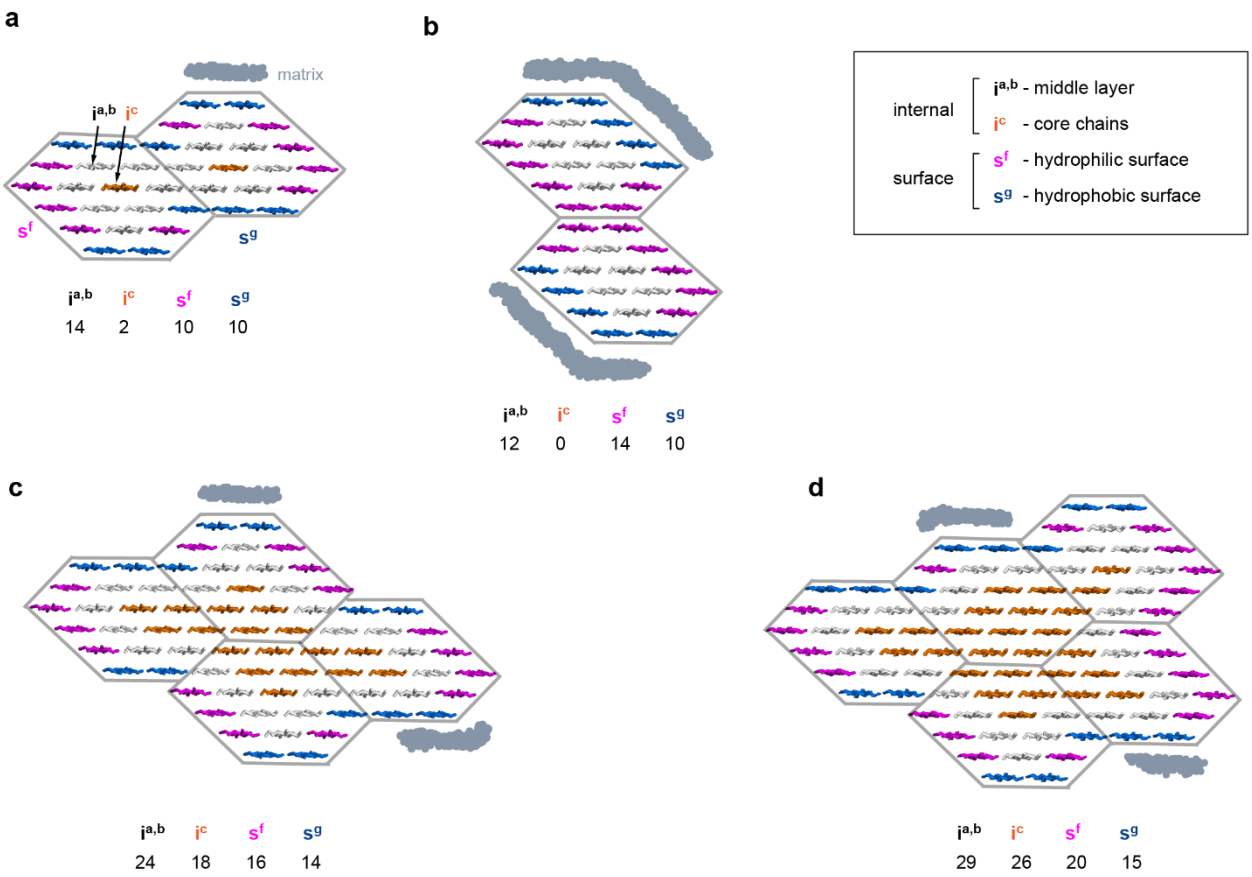

Figure S7.2 Cellulose bundling models that violate the NMR constraints. The cartoon illustrations include **a and b,** two elementary microfibrils, **c,** four elementary microfibrils, and **d,** five elementary microfibrils. Each elementary microfibril is depicted to contain 18 glucan chains following the current biochemical evidence[90, 91]. The number of hydrophobic surface chains (s^g), hydrophilic surface chains (s^f), middle layer internal chains ($i^{a,b}$), and the deeply embedded core chains (i^c) are labeled[92, 93]. The matrix polymers are added to better fit the NMR constraints, but major violations still exist for all these models.

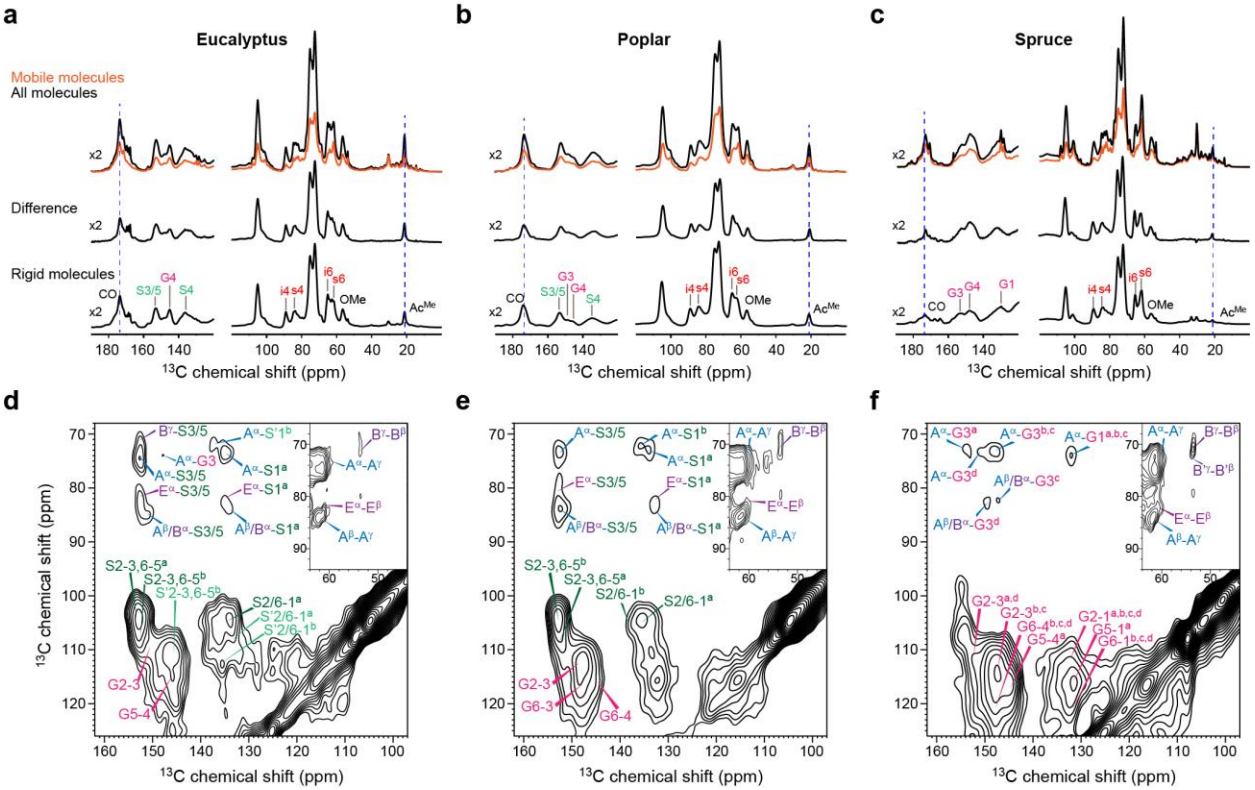

Figure S7.3 Lignin distributed in the mobile and rigid phases. The top row is the quantitative-detected DP (orange), mobile molecule detected DP, rigid component detected CP spectra of **a**, eucalyptus, **b**, poplar, and **c**, spruce. All the three difference spectra of ^{13}C DP spectra with long and short recycle delays showing peaks of the rigid components, are comparable to the corresponding CP spectra. Both 2-s DP and CP spectra give well-resolved aromatic peaks, indicating dynamically heterogeneous lignin components. The bottom row is DP-PDSD spectra (1.7s recycle delay) of **d**, eucalyptus, **e**, poplar, and **f**, spruce. Abundant signals of aromatics and linkers are identified, suggesting that these linkers are also revolved at mobile phases of lignin.

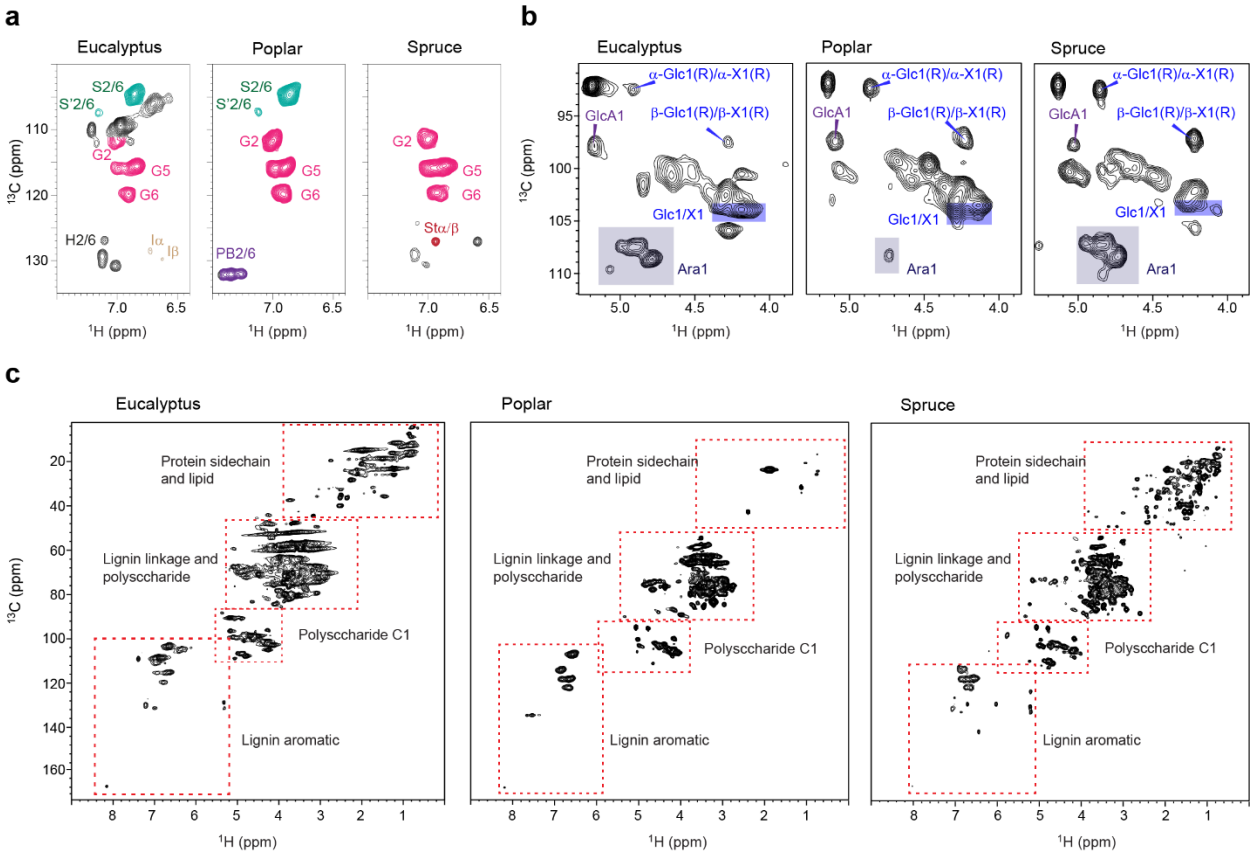

Figure S7.4 Solution NMR of lignin and polysaccharides in woody stems. 2D ^{1}H-^{13}C HSQC spectra showing selected regions of **a**, mainly lignin signals and **b**, mainly polysaccharide peaks. Arabinose (Ara) is more predominant in spruce than in eucalyptus and poplar. **c,** Full HSQC spectra of the three samples. The ^{13}C and ^{1}H chemical shifts are documented in Table S7.4 and supported by a literature surveil compiled in Table S7.5.

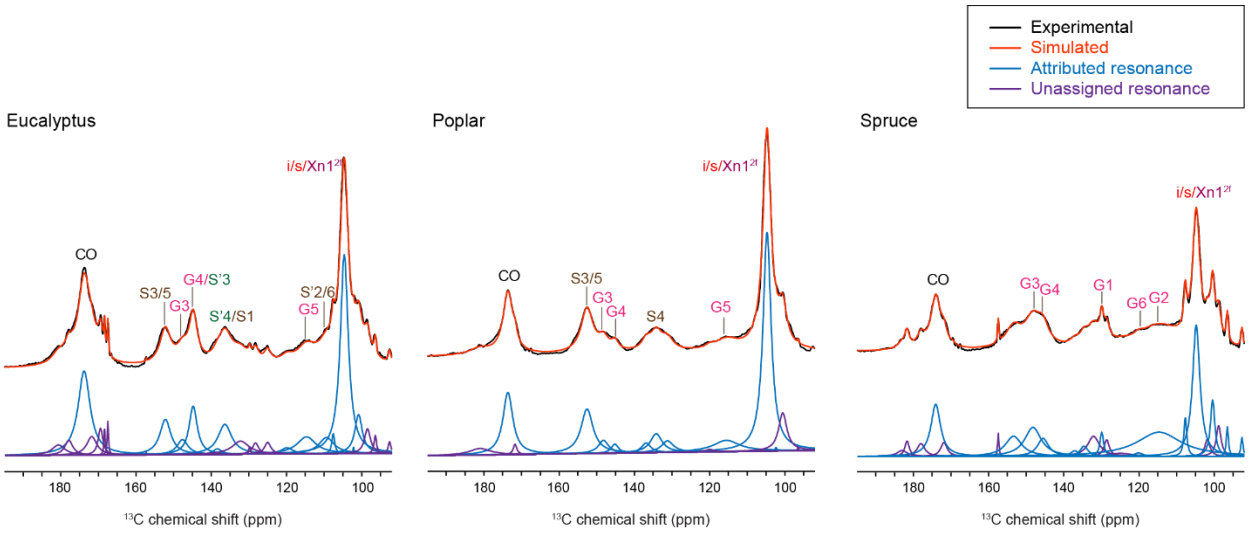

Figure S7.5 Spectral deconvolution reveals the lignin composition. Deconvolution is conducted on 1D quantitative ^{13}C DP spectra collected with long recycle delays of 35-40 s. The experimentally measured spectra (black) is overlaid with the simulated spectra (red) for eucalyptus, poplar, and spruce. The bottom panels show the deconvoluted peaks. Attributed resonances are plotted in blue and uncertain resonances are in purple. Deconvolution is conducted using the DMfit software[94], using a Lorentzian model on a 92-195 ppm chemical shift window. The fitting parameters are summarized in Table S7.6.

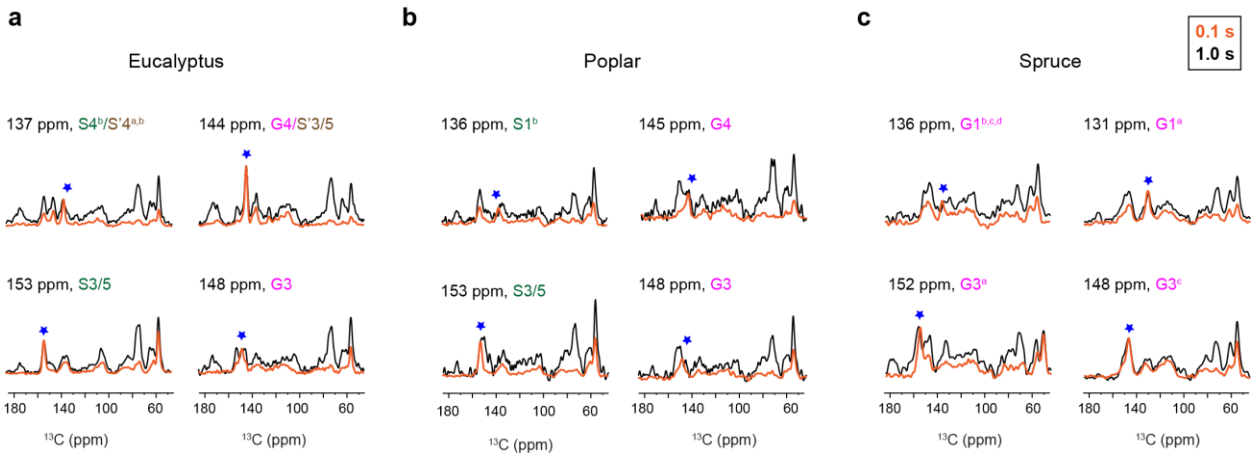

Figure S7.6 1D ^{13}C cross sections of lignin region reveal the polymer mixing patterns. Representative cross sections are obtained from dipolar gated 2D ^{13}C-^{13}C correlation spectra (Figure 7.3a) of **a,** eucalyptus, **b,** poplar, and **c,** spruce. For eucalyptus and poplar, the first column shows the cross sections from S-lignin units and the second column shows those from G units. All the cross sections are normalized by the diagonal peaks (blue asterisks). The more similar the spectral pattern between 0.1 s and 1.0 s, the more homogeneous the polymers are mixed on the sub-nanometer scale. The spruce cross sections are more equilibrated.

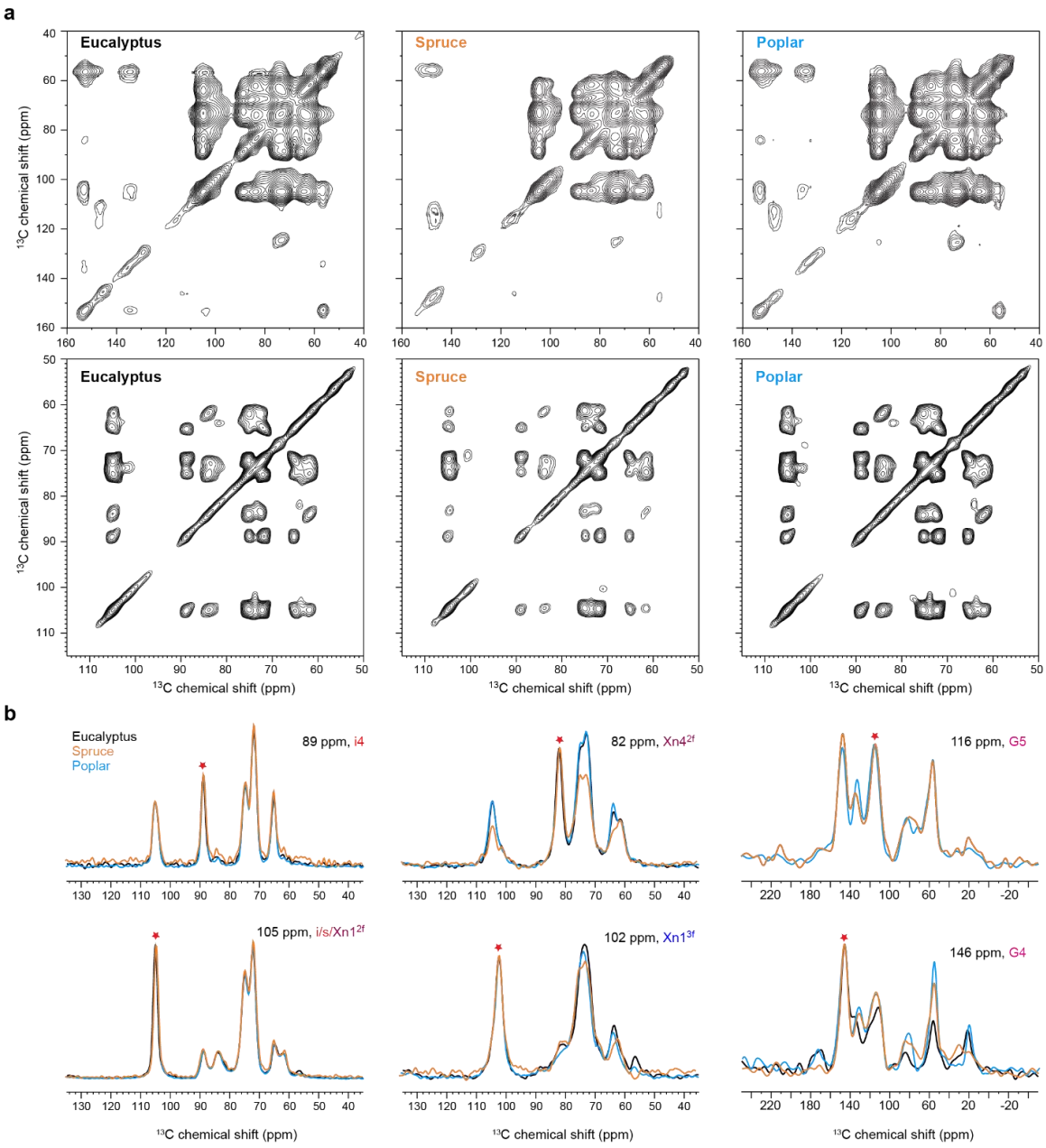

Figure S7.7 Comparable spin diffusion rates of biopolymers in three wood samples. **a**, Aromatics regions (top row) and carbohydrates regions (second row) of 0.1s PDSD spectra of three woody plants. **b**, Representative cross peaks for cellulose, xylan, and lignin are obtained from 0.1s PDSD spectra, which are normalized by the diagonal peaks (red asterisks). The similar spectral patterns among three wood samples indicates the spin diffusion rates of lignin and carbohydrates are comparable in all three woody stems.

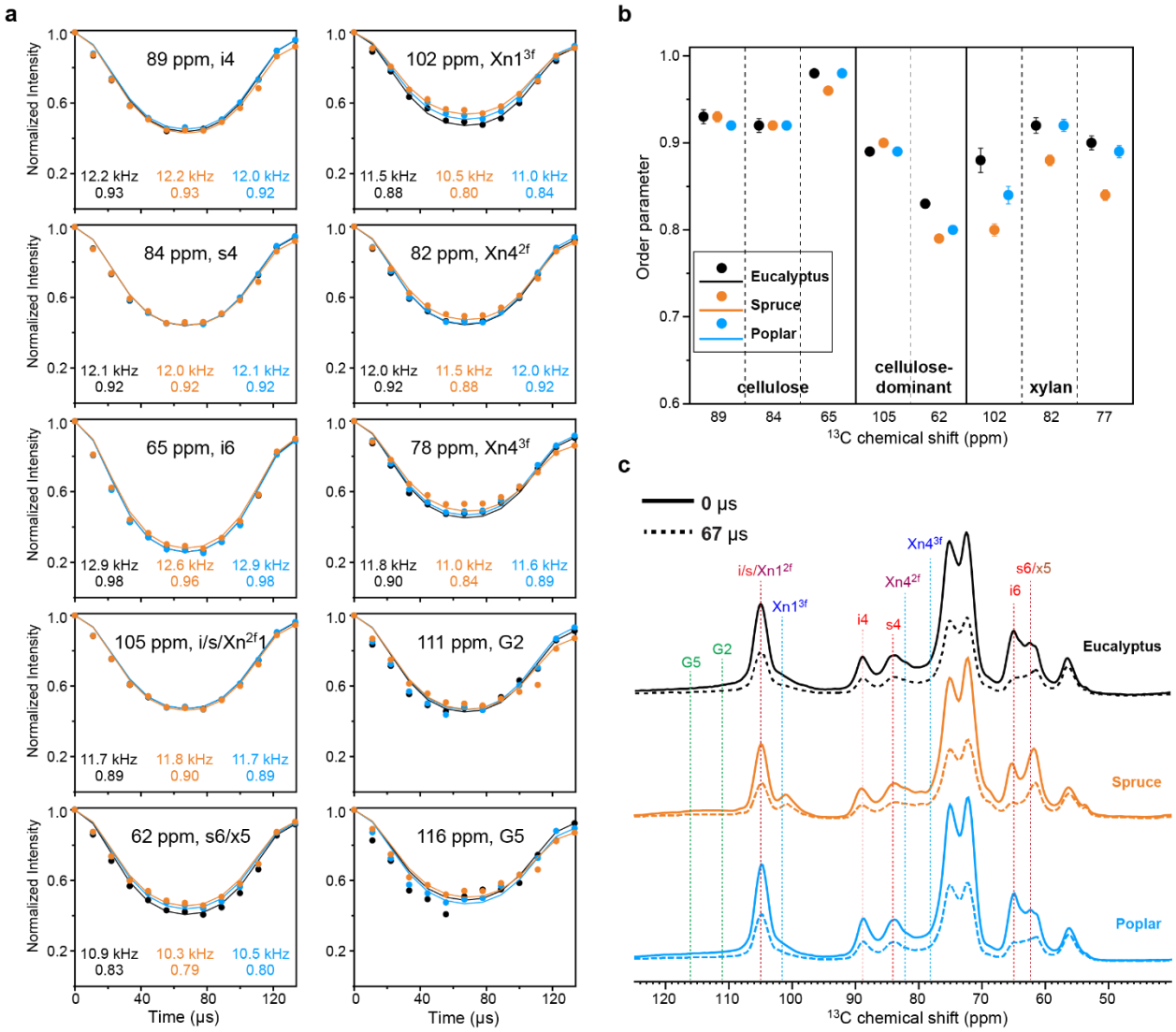

Figure S7.8 ^{13}C–^{1}H dipolar order parameters of biopolymers in wood cell walls. **a**, Dipolar dephasing curves of carbohydrates and lignin processed from DIPSHIFT spectra. The spectra were collected under 7.5 kHz MAS with a 0.577 scaling factor for C-H bond. **b**, Summary of C–H dipolar order parameters with error bars. Spruce shows a relatively smaller order parameter for xylan. **c**, The first (0 μs dipolar dephasing) and middle (67 μs dipolar dephasing) slices of the 2D DIPSHIFT experiment imply how much the spectra are suppressed at a half rotor period. Source data are provided as a Source Data file.

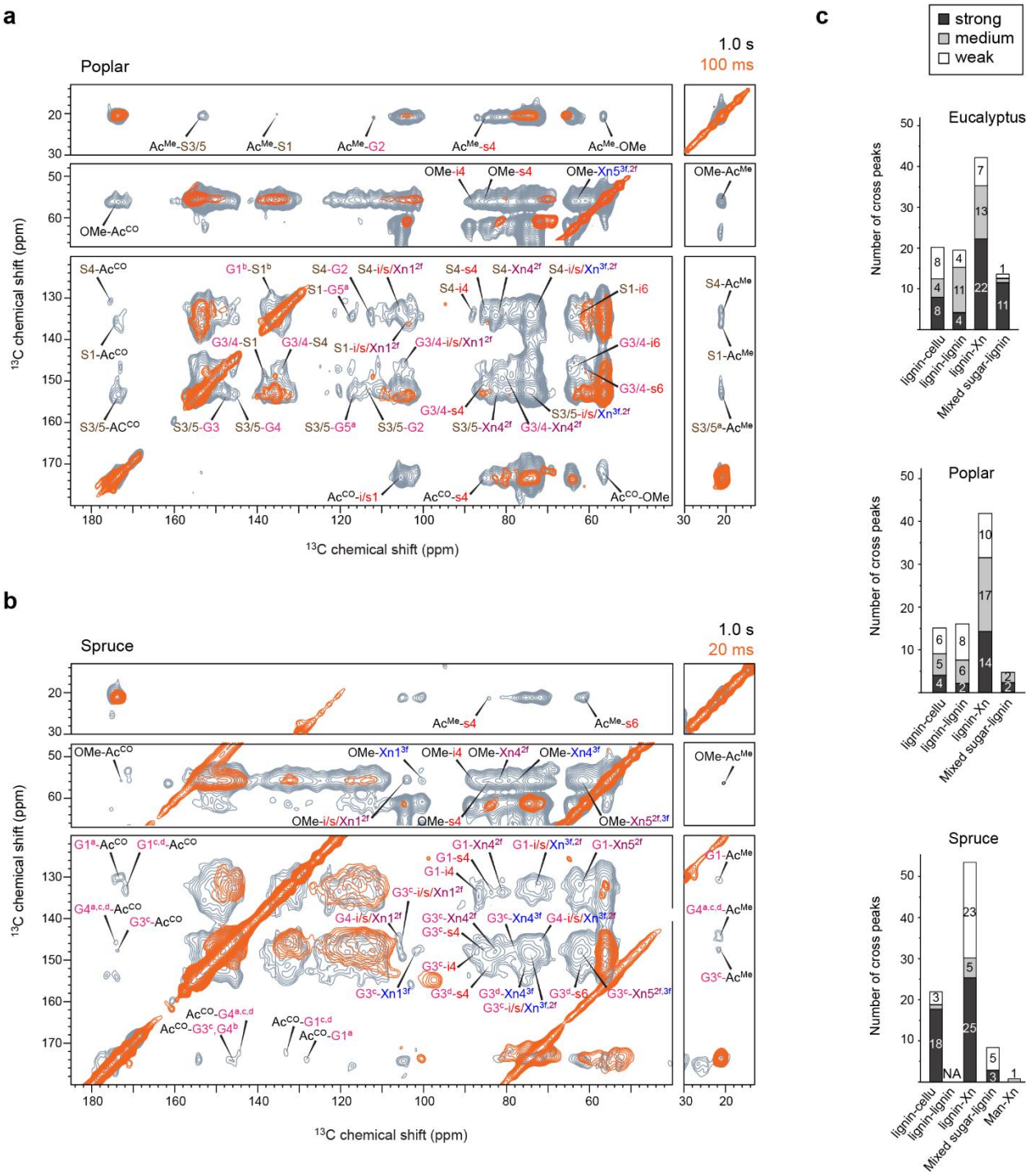

Figure S7.9 Intermolecular interactions of polymers in woody plants. Dipolar-gated 2D ^{13}C-^{13}C correlation spectra measured with 1.0 s and 100 ms (poplar), 1.0 s and 20 ms (spruce) are overlaid for **a,** poplar and **b,** spruce. 100 ms (poplar) and 20 ms (spruce) spectra mainly detect the intramolecular correlations, and the 1 s spectra show many intermolecular cross peaks. Only the

Figure S7.9 (cont'd)

intermolecular cross peaks are labeled. **c**, The bar diagrams show the number of interactions between different molecules in the woods. The interactions are categorized according to the peak intensity. The details of the short-range and long-range cross peaks are summarized in Tables S7.7-7.10. Source data are provided as a Source Data file.

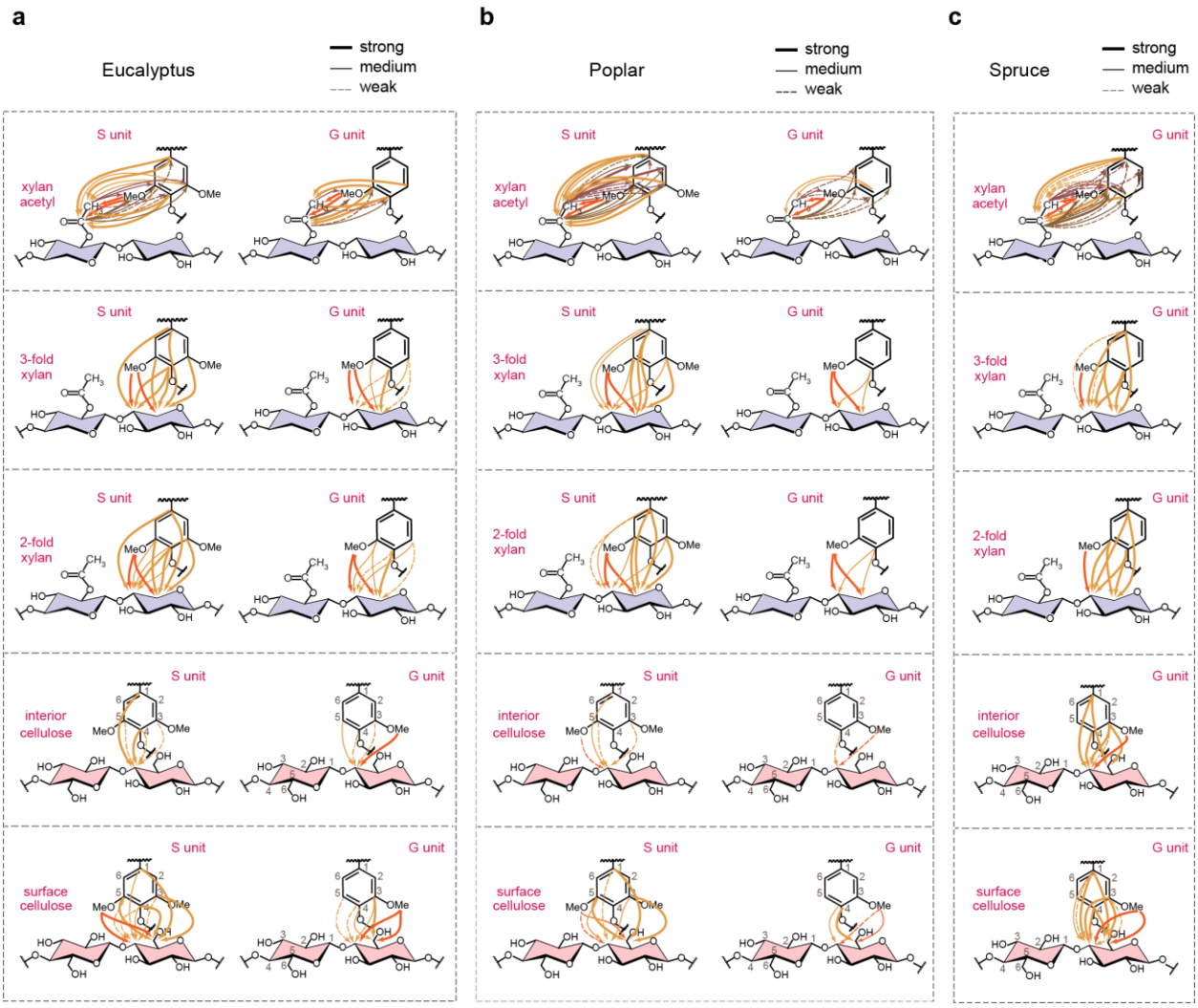

Figure S7.10 NMR restraints of polymer packing in secondary cell walls. Thick lines, thin lines, and dash lines are used to represent strong, medium, and weak cross peaks observed in solid-state NMR spectra, respectively. Three major types of interactions happen between xylan acetyl and lignin (brown), between lignin methyl ether and carbohydrates (orange), and between lignin aromatics and carbohydrates (yellow). The details of these cross peaks are tabulated in Tables S7.7-7.10.

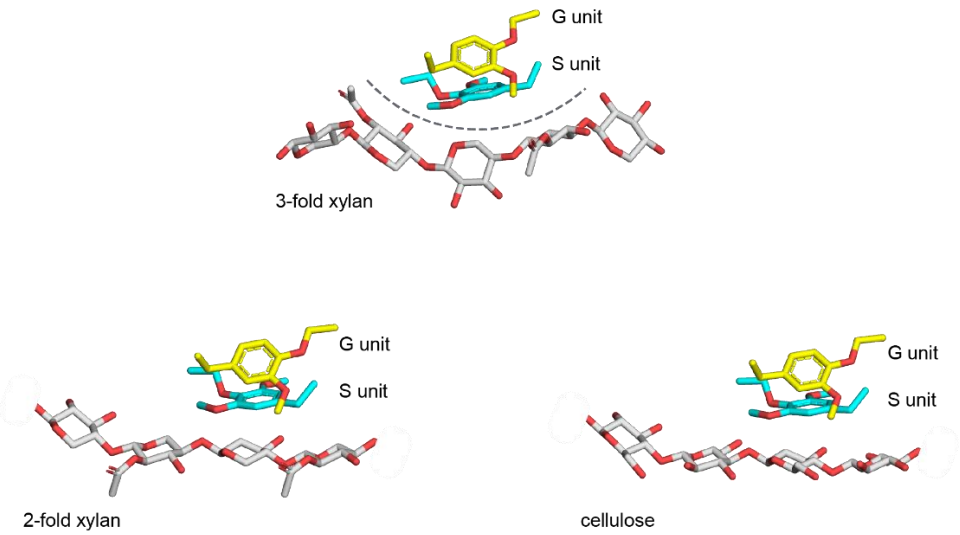

Figure S7.11 DFT structures of lignin-carbohydrate packing. The structure includes the complex formed between G (yellow) or S (cyan) unit of lignin and three types of polysaccharides. The structures were adapted from an earlier study.[95]

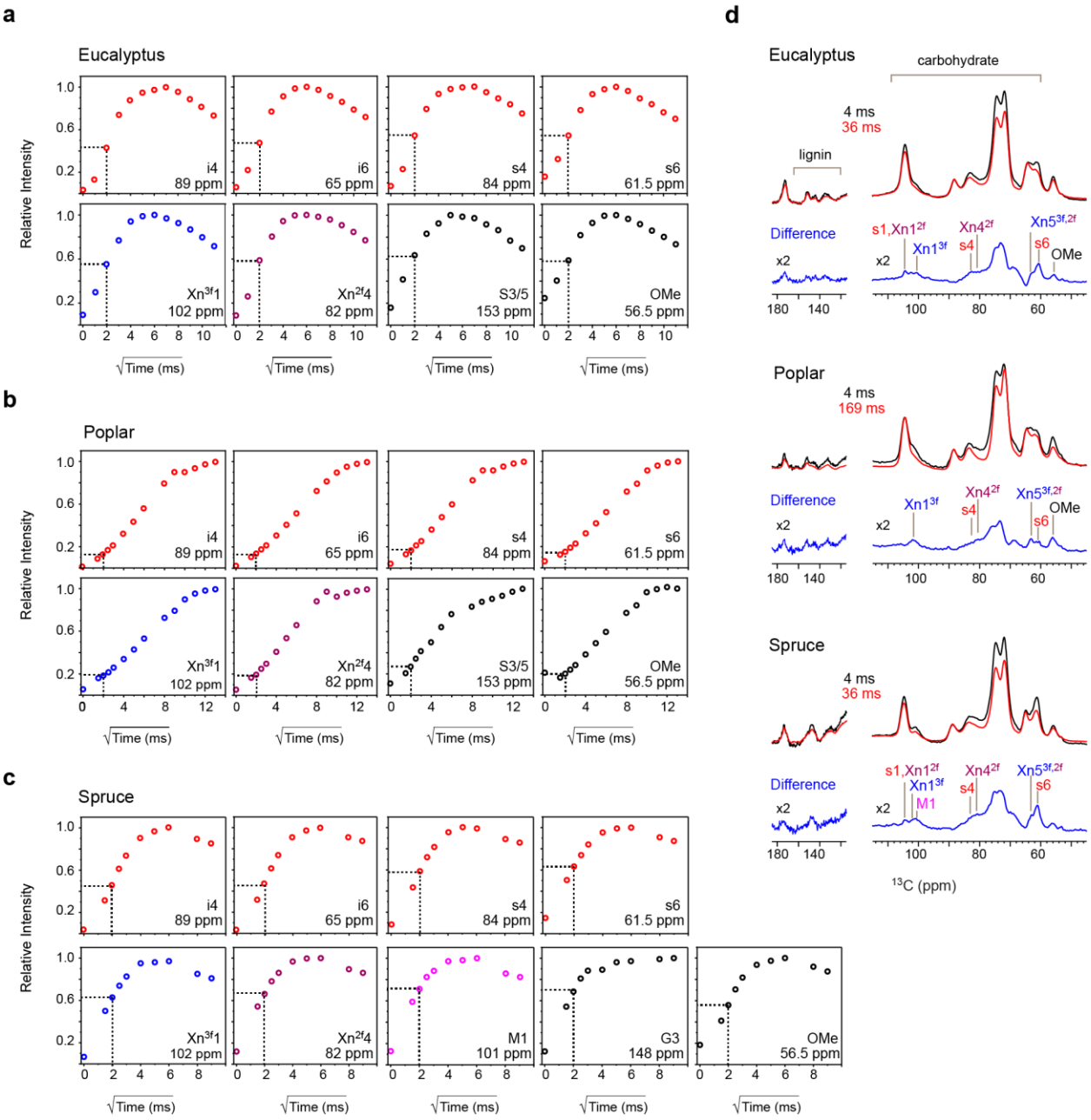

Figure S7.12 Buildup curves for water-to-polysaccharides/lignin. The water ^{1}H spin diffusion curves for cellulose, xylan, and lignin for **a,** Eucalyptus, **b,** poplar and **c,** spruce. Dash lines indicate the intensities of 4-ms ^{1}H mixing. Poplar has the slowest spin diffusion from water among the three woods, with ~15-20% of the equilibrium intensity detected at 4-ms ^{1}H mixing. Eucalyptus and spruce have the fastest spin diffusion with ~45-60% of the equilibrium intensity detected at 4-ms ^{1}H mixing. **d,** Water-edited spectra measured with different mixing times. The 4-ms spectrum

Figure S7.12 (cont'd)

detects the hydrated molecules and the 36-ms (eucalyptus and spruce) and 169-ms (poplar) spectra report equilibrium intensities. The difference spectrum only shows well-hydrated molecules. Source data are provided as a Source Data file.

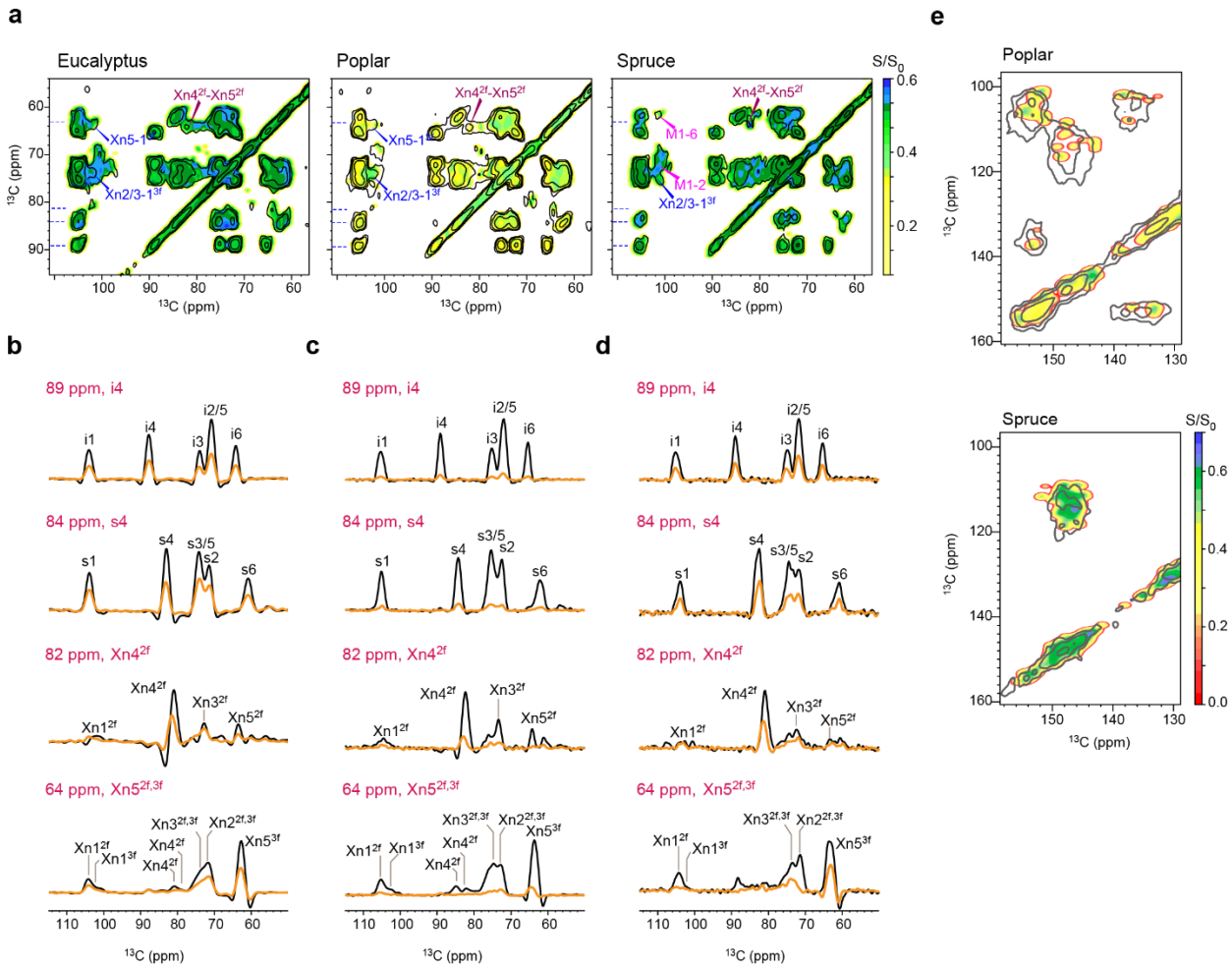

Figure S7.13 Water-edited 2D ^{13}C-^{13}C correlation spectra of polysaccharides. Overlay of water-edited and equilibrium spectra showing hydration maps with S/S_0 values for **a**, Eucalyptus, **b**, poplar, and **c**, spruce. The blue dash lines indicate the positions at which the ^{13}C cross sections are extracted. The representative 1D ^{13}C cross sections of cellulose (interior: i4; surface: s4) and xylan (Xn4 and Xn5) are shown for **a**, Eucalyptus, **b**, poplar, and **c**, spruce. The water-edited spectra and the equilibrium spectra are plotted in orange and black, respectively. The 3-fold and 2-fold xylans have enhanced intensity in the water-edited spectra, indicating better interactions with water molecules. **e**, Hydration maps of lignin regions with S/S_0 values for poplar (top) and spruce (bottom). Spruce is well hydrated compared to poplar as it shows higher S/S_0 values. The water-edited intensities are summarized in Tables S7.11-7.13.

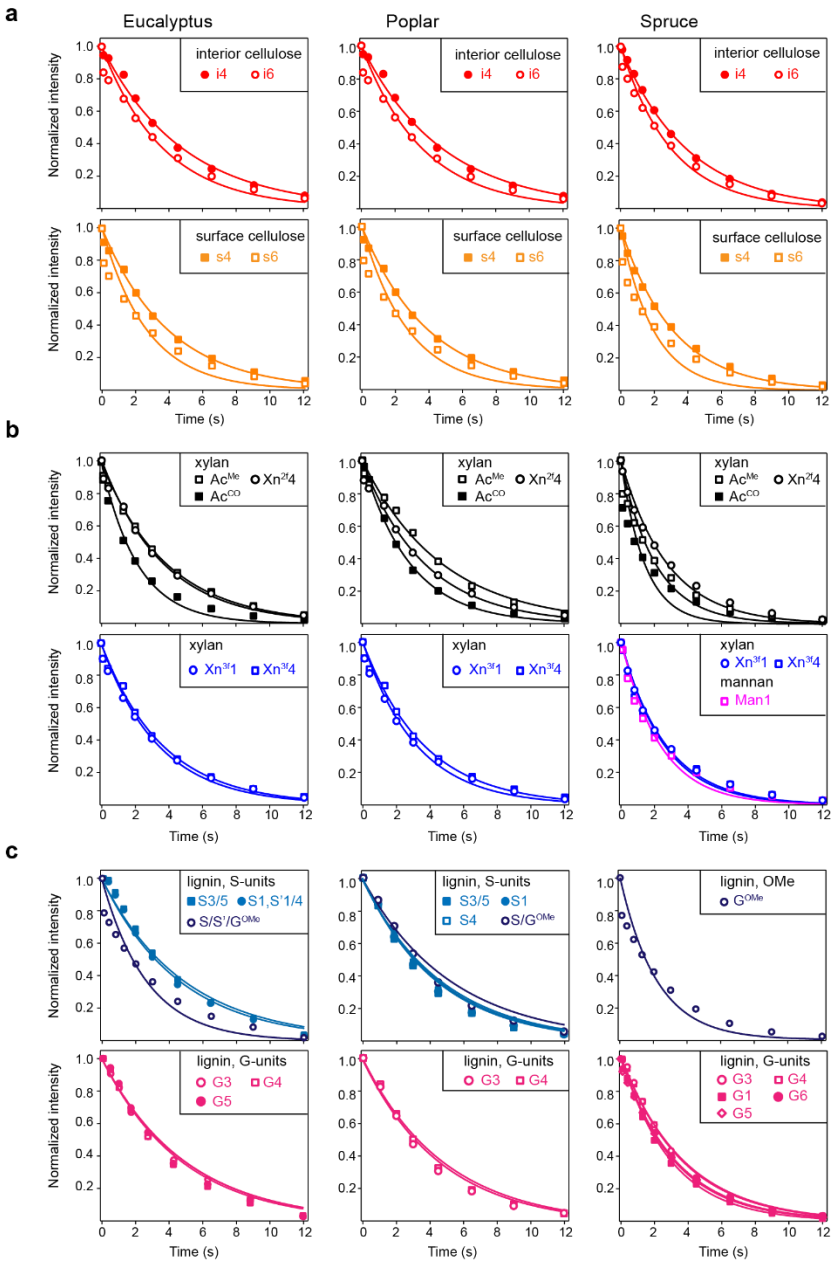

Figure S7.14 The ^{13}C-T_1 relaxation curves of polysaccharides and lignin. The ^{13}C-T_1 relaxation curves of **a**, cellulose (interior and surface glucan chains), **b**, hemicellulose (2-fold and 3-fold xylan; mannan), and **c**, lignin (S and G units) are shown. The left, middle, and right columns are for eucalyptus, poplar, and spruce, respectively. The data are fitted using a single exponential equation. The fit parameters are summarized in Table S7.14. Source data are provided as a Source Data file.

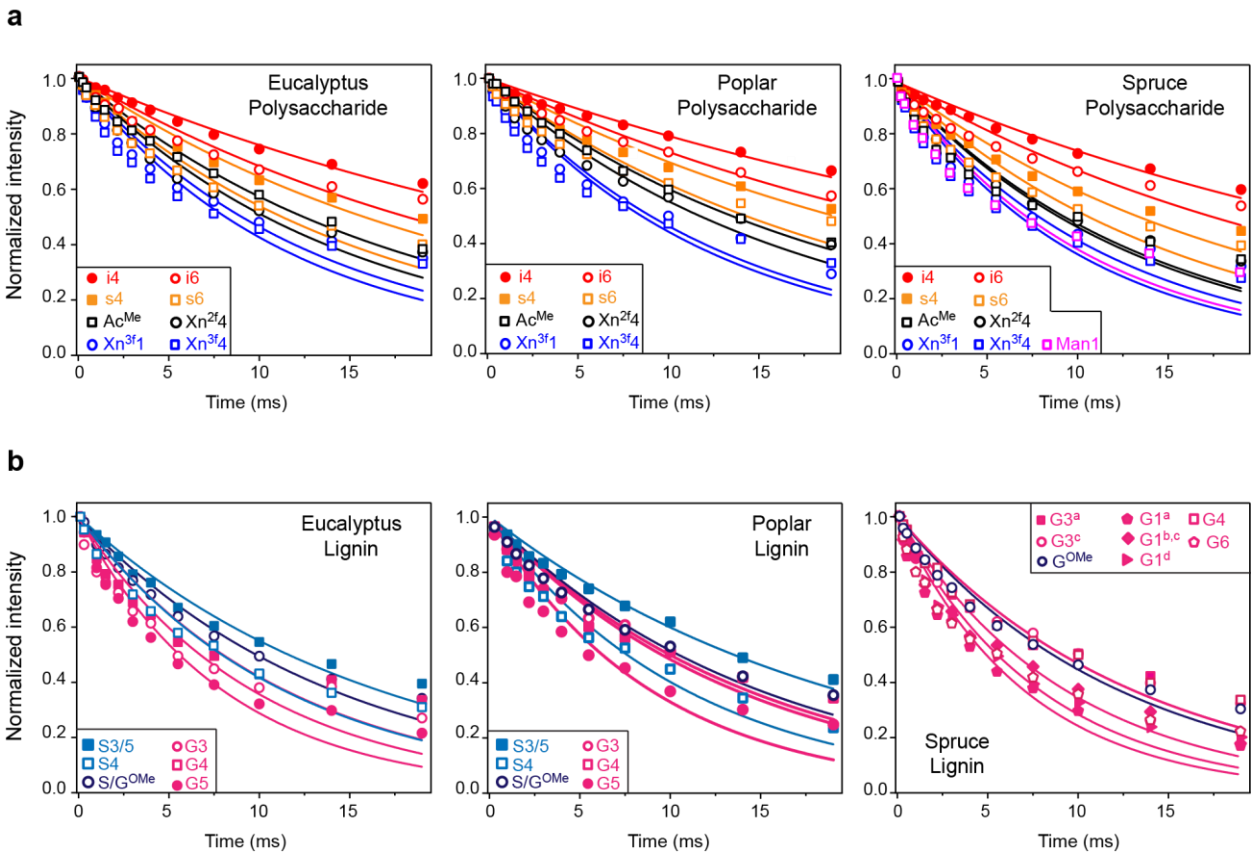

Figure S7.15 ^{1}H-T$_{1\rho}$ relaxation curves of polysaccharides and lignin in woody plants. The ^{1}H-T$_{1\rho}$ relaxation curves of **a**, polysaccharides and **b**, lignin data are shown for eucalyptus, poplar, and spruce. The data are fitted using a single exponential equation. Xylan shows faster relaxation times compared to cellulose. Within lignin, the G residue has faster ^{1}H-T$_{1\rho}$ relaxation than the S unit. The fit parameters are summarized in Table S7.15. Source data are provided as a Source Data file.

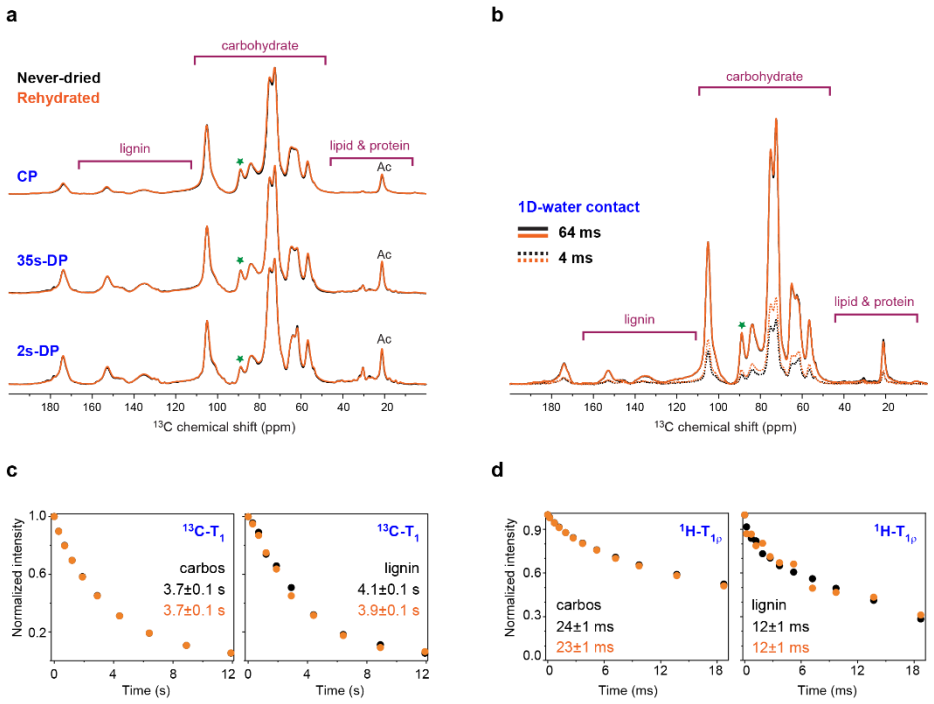

Figure S7.16 Effect of freeze-drying and rehydration on hardwood eucalyptus. **a**, CP (rigid components selection), quantitative DP and DP (mobile components selection) normalized to i4 peak at 89 ppm. Distinct chemical shift intervals for lignin, carbohydrate, and lipid are highlighted. **b**, T_2-filtered (water-edited) 1D experiments. Signal maxima observed after 64 ms of spin diffusion, in comparison to the previous 36 ms and 26% signal difference in the absence of diffusion upon rehydration are consistent with both a sample less prone to deep hydration and a slight betterment of its surface water access after freeze-drying. **c**, Average ^{13}C T_1 relaxation curves, using an evolution of integrals of spectral regions established in panel (a). **d**, Average ^{1}H-$T_{1\rho}$ relaxation using the same two integrated intervals. Error bars are standard deviations of the fitting derived parameters. Source data are provided as a Source Data file. All these spectra and fitted relaxation parameters are very close for the fresh and rehydrated samples, demonstrating that structure and dynamics are not significantly altered by this sample preparation method. Their overall consistency with the main text eucalyptus sample also validates experimental reproducibility.

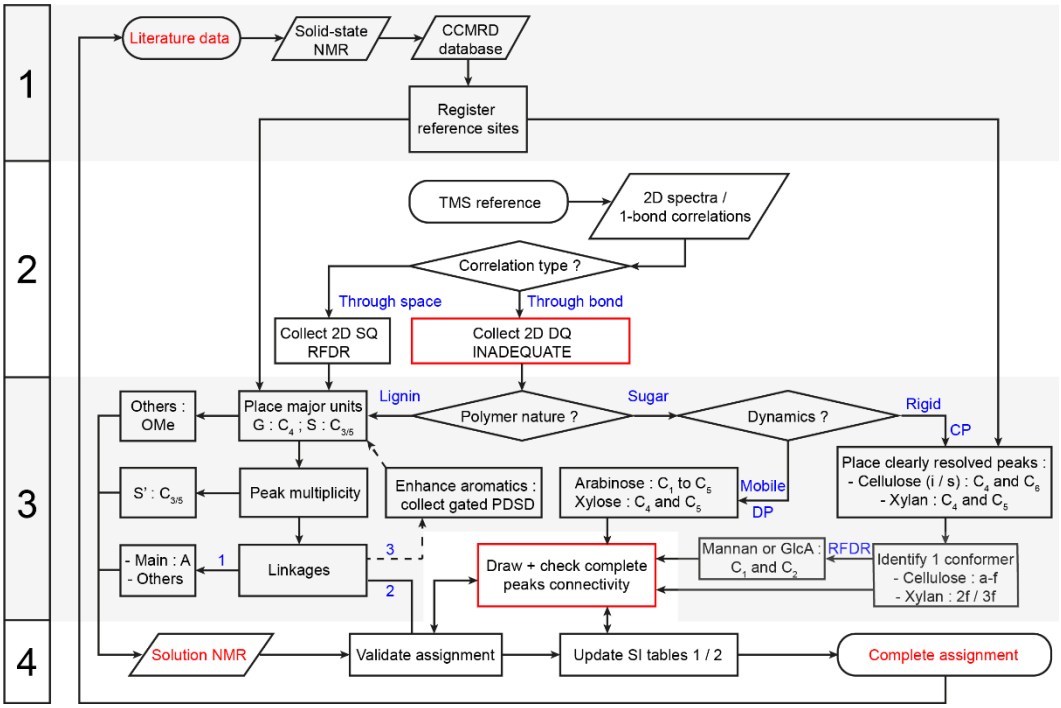

Figure S7.17 Representative flow chart for resonance assignment. Shading and numbering on the left highlight the four main steps, while red coloring denotes essential sub-steps of the protocol followed for assigning chemical shifts. 1) Collection of reference chemical shifts for expected resolved peaks from solid-state NMR literature and databases. For each carbon site, we report associated references in Table S7.19. 2) Acquisition of 2D solid-state NMR spectra. Subsequent assignments rely primarily on the resolution provided by the DQ dimension of INADEQUATE spectra. 3) Sequential assignment (boxes from right to left), from well-established cellulose and xylan conformers to mobile primary cell wall polysaccharides, ending with lignin and its linkages. For each polymer, we indicate a well-established site from which assignment can be easily initiated, while less resolved sites are either identified via cross-peaks connections or from literature reported values. 4) Safeguarding steps taken to provide accountability in the final assignments. Particularly for the complex lignin assignment, a good match is demanded with solution NMR spectra, as summarized in Tables S7.5 and S7.19.

Table S7.1. ^{13}C chemical shifts of rigid polysaccharides and lignin in wood secondary cell walls. The rigid components are identified from ^{13}C CP-based INADEQUATE and RFDR spectra. Superscripts are used to denote different allomorphs. Unidentified signals are indicated as (-). Not applicable (/)

	Lignin	C1	C2	C3	C4	C5	C6	OMe	C7 (α)	C8(β)	C9(γ)	Plant
G	guaiacyl	-	111.1	148.9	144.9	115.9	-	57.0	/	/	/	Eucalyptus
G^{a}		130.7	113.4	152.2	145.3	116.3	-	56.7	/	/	/	Spruce
G^{b}		133.5	110.5	149.7	146.1	-	120.2	57.0	/	/	/	Spruce
G^{c}		134.2	113.5	147.6	144.1	-	119.2	57.0	/	/	/	Spruce
G^{d}		136.0	109.5	151.6	142.2	-	121.1	57.0	/	/	/	Spruce
G^{a}		-	112.1	148.2	144.9	116.1	-	56.6				Poplar
G^{b}		-	113.8	148.2	144.9	116.1	-	56.6	/	/	/	Poplar
S^{a}	syringyl	134.9	102.8	153.3	133.5	153.3	102.8	57.0	/	/	/	Eucalyptus
S^{b}		-	109.7	153.2	137.2	153.2	109.7	57.0	/	/	/	Eucalyptus
S^{a}		134.2	104.0	153.1	134.6	153.1	104.0	56.6	/	/	/	Poplar
S^{b}		137.1	102.5	152.7	135.8	152.7	102.5	56.6	/	/	/	Poplar
S'^{a}	Cα-oxidized syringyl	134.1	109.2	144.8	136.3	144.8	109.2	56.5	/	/	/	Eucalyptus
S'^{b}		136.5	112.5	145.5	138.5	145.5	112.5	56.5	/	/	/	Eucalyptus
A(S)	β-O-4'	/	/	/	/	/	/	/	72.3	84.5	59.4	Eucalyptus
A (G)		/	/	/	/	/	/	/	72.4	84.3	59.9	Spruce
A(S)		/	/	/	/	/	/	/	72.0	86.3	60.4	Poplar
B	γ-acylated β-β'	/	/	/	/	/	/	/	84.4	53.5	71.2	Eucalyptus
B'	β-β'	/	/	/	/	/	/	/		53.5	74.4	Eucalyptus
B		/	/	/	/	/	/	/		53.9	70.3	Spruce
B'		/	/	/	/	/	/	/	83.9	52.9	74.1	Spruce
B		/	/	/	/	/	/	/	84.8	52.6	70.5	Poplar
C'	β-5'	/	/	/	/	/	/	/	90.2	49.9		Spruce
C'	γ-acylated β-5'	/	/	/	/	/	/	/	86.7	49.6		Spruce
C		/	/	/	/	/	/	/	88.1	54.2		Spruce
E	β'-1'	/	/	/	/	/	/	/	81.1	59.6		Eucalyptus
		/	/	/	/	/	/	/	81.4	60.0		Spruce

Table S7.1 (cont'd)

		/	/	/	/	/	/	/	80.3	59.7		Poplar
	Carbohydrate	C1	C2	C3	C4	C5	C6	Ac^{CO}	Ac^{Me}			
Xn^{2f}	2-fold xylan	104.8	72.2	74.0	82.2	64.0	/	174.0	21.0			Eucalyptus
Xn^{3f}	3-fold xylan	102.5	73.4	74.1	77.4	63.3	/	174.0	21.0			
Xn	xylan mix	-	-	74.2	78.5	64.1	/	174.0	21.0			Eucalyptus
Xn^{2f}	2-fold xylan	104.8	72.1	73.2	82.0	63.9	/	174.0	21.0			Poplar
Xn^{3f}	3-fold xylan	102.2	73.6	74.2	77.9	64.1	/	174.0	21.0			
Xn^{2f}	2-fold xylan	104.8	72.2	73.3	82.2	63.9	/	174.0	21.0			Spruce
Xn^{3f}	3-fold xylan	103.5	73.9	74.2	77.8	65.4	/	174.0	21.0			
x	xylose in XyG	-	-	74.1	70.4	62.5	/	/	/			All samples
$GlcA^{a}$	GlcA in GAX	101.5	71.3	79.8	69.9	76.6	-					All samples
$GlcA^{b}$	GlcA in GAX	99.2	69.9	-	-	-	-					All samples
$GlcA^{c}$	GlcA in GAX	98.5	71.6	80.8	-	-	-					All samples
GalA	galacturonic acid	100.2	68.1	-	79.6	-	-					All samples
M	mannan	101.2	71.3	75.1	-	75.1	62.1					Spruce
M^{Ac}	acetylated mannan	101.6	72.1	75.1	80.7	75.1	62.1					Spruce
A (t-Ara)	arabinose	108.1	82.1	-	85.6	62.1	/					Spruce
i^{a}	interior cellulose	104.9	72.1	75.2	88.7	72.3	65.0					Eucalyptus
i^{b}		104.9	72.1	75.2	88.7	72.3	65.0					
i^{c}		104.9	72.1	74.1	87.9	71.1	65.7					
s^{f}	surface cellulose	104.9	72.1	74.6	84.4	74.6	62.3					Eucalyptus
s^{g}		104.9	72.1	73.6	83.5	73.6	61.4					
i^{a}	interior cellulose	104.8	72.1	75.0	88.9	72.4	64.9					Poplar
i^{b}		104.8	72.1	75.0	88.9	72.4	64.9					
i^{c}		104.8	72.1	74.3	87.9	70.1	65.5					
s^{f}	surface cellulose	104.8	72.1	75.7	84.3	75.0	62.3					Poplar
s^{g}		104.8	72.1	75.2	83.4	73.7	61.3					
i^{a}	interior cellulose	104.8	72.2	75.3	89.0	72.4	65.0					Spruce
i^{b}		104.8	72.2	75.3	89.0	72.4	65.0					
s^{f}	surface cellulose	104.8	72.2	75.1	84.5	75.0	62.1					Spruce
s^{g}		104.8	72.2	74.3	84.3	74.3	61.4					

Table S7.2. ^{13}C chemical shifts of mobile molecules in woods. The mobile components are identified from ^{13}C DP-based INADEQUATE spectra measured with short recycle delays of 2 s. Superscripts are used to denote different allomorphs. Not applicable (/). Unidentified (-).

	carbohydrate	C1	C2	C3	C4	C5	C6	Plants
A[a]	arabinose	107.8	81.8	77.2	84.6	61.8	/	All samples
A[b]		107.8	80.1	-	-	-	/	All samples
A[c]		109.7	81.8	77.4	82.9	67.6	/	All samples
A[d]		106.9	85.3	-	-	-	/	Eucalyptus
Gal	galactose	104.9	72.3	-	-	-	-	All samples
α-Glc	α-glucose	92.7	-	-	-	-	-	Spruce
β-Glc	β-glucose	96.6		-	-	-	-	Spruce
x	xylose in XyG	100.3	72.3	-	71.9	61.3	/	All samples
i[a,b]	cellulose	-	-	75.5	88.9	72.4	64.9	Eucalyptus
i[c]		-	-	75.5	88.9	70.9	65.8	Eucalyptus
i[a,b]		-	-	75.5	88.9	72.4	64.9	Poplar
i[c]		-	-	75.5	88.9	70.9	65.8	Poplar
s[f]		-	-	-	-	75.0	61.4	All samples
s[g]		-	-	-	-	73.7	61.2	Eucalyptus
s[g]		-	-	-	-	73.7	61.2	Poplar
	lignin	C1	C2	C3	C4	C5	C6	Plants
G[a]	guaiacyl	136.3	-	144.1	144.1	-	-	Eucalyptus
G[b]		136.3	-	145.2	145.2	-	-	
S[a]	syringyl	136.7	-	152.9	-	152.9	-	
S[b]		135.5	-	151.7	-	151.7	-	
G	guaiacyl	130.1	-	147.8	146.4	-	-	Poplar
S[a]	syringyl	-	-	152.6	-	152.6	-	
S[b]		-	-	151.9	-	151.9	-	
G	guaiacyl	-	-	148.9	144.7	-	-	Spruce
G		-	-	152.3	-	152.3	-	

Table S7.3. Molar composition of polysaccharides in wood secondary cell walls. For each plant, the composition of cellulose, xylan, mannan, and the primary cell wall component xyloglucan (XyG) is given. Cellulose contains interior glucan chains ($i^{a,b}$: middle layer; i^{c}: embedded core chains) and surface glucan chains (s^{f}: hydrophilic surface; s^{g}: hydrophobic surface). Xylan contains the backbone (Xn^{2f}: two-fold xylan; Xn^{3f}: three-fold xylan; Xn: mixed conformation) and sidechains (GlcA and/or Ara). For Mannan, only the mannose residues (M: unacetylated; M^{Ac}: acetylated) are resolved, while the actual fraction of mannan should be higher than the value reported here due to the presence of Glc residues in the backbone. The sidechain α-xylose (x) is used to denote xyloglucan.

Eucalyptus

Cellulose 74.0%				Xylan 25.5%					Mannan -		XyG 0.5%
interior chains 38.2%		surface chains 35.8%		backbone 16.8%			sidechains 8.7%		- -		
$i^{a,b}$	i^{c}	s^{f}	s^{g}	Xn^{2f}	Xn	Xn^{3f}	GlcA	Ara	M	M^{Ac}	x
28.3%	9.8%	21.2%	14.6%	10.0%	2.0%	4.9%	8.7%	-	-	-	0.5%

Poplar

Cellulose 81.4%				Xylan 18.0%					Mannan -		XyG 0.6%
interior chains 45.0%		surface chains 36.4%		backbone 13.3%			sidechains 4.8%		-		
$i^{a,b}$	i^{c}	s^{f}	s^{g}	Xn^{2f}	Xn	Xn^{3f}	GlcA	Ara	M	M^{Ac}	x
32.7%	12.3%	21.9%	14.6%	7.9%	1.6%	3.8%	4.8%	-	-	-	0.6%

Spruce

Cellulose 75.8%				Xylan 11.8%					Mannan 10.5%		XyG 1.9%
interior chains 37.1%		surface chains 38.7%		backbone 10.4%			sidechains 1.5%				
$i^{a,b}$	i^{c}	s^{f}	s^{g}	Xn^{2f}	Xn	Xn^{3f}	GlcA	Ara	M	M^{Ac}	x
28.1%	8.9%	23.6%	15.2%	6.9%	-	3.5%	0.5%	1.0%	6.2%	4.3%	1.9%

The area of the following well-resolved peak pairs in CP J-INADEQUATE spectra are used:

$i^{a,b}$: the average of $i5^{a,b}$ - $i6^{a,b}$, $i4^{a,b}$ - $i5^{a,b}$, and $i3^{a,b}$ - $i4^{a,b}$.

Table S7.3 (cont'd)

i^{c} : the average of $i5^{c}$ - $i6^{c}$, $i4^{c}$ - $i5^{c}$, and $i3^{c}$ - $i4^{c}$.

s^{f} : the average of $s5^{f}$ - $s6^{f}$, $s4^{f}$ - $i5/3^{f}$. The integration of $s4^{f}$ - $i5/3^{f}$ has been doubled because 2 peak pairs are included (carbon 4-carbon 5 and carbon 4-carbon 3).

s^{g} : the average of $s5^{g}$ - $s6^{g}$, $s4^{g}$ - $i5/3^{g}$. The integration of $s4^{g}$ - $i5/3^{g}$ has been doubled because 2 peak pairs are included (carbon 4-carbon 5 and carbon 4-carbon 3).

Xn^{2f}: the average of $Xn4^{2f}$ and $Xn5^{2f}$.

Xn: the average of Xn4 and Xn5.

Xn^{3f}: the average of $Xn4^{3f}$ and $Xn5^{3f}$.

GlcA: the average of GlcA1 and GlcA2. The value reported are the sum of $GlcA^{a}$, $GlcA^{b}$ and $GlcA^{c}$.

Ara: the average of A1 and A2.

M: the average of M1 and M2.

M^{Ac}, the average of $M1^{Ac}$ and $M2^{Ac}$.

x: the average of x4 and x5.

Table S7.4. Solution NMR HSQC ^{1}H and ^{13}C chemical shifts of lignin aromatics and linkages.

Unidentified or unresolved (-).

		Eucalyptus		Poplar		Spruce	
		^{13}C (ppm)	^{1}H (ppm)	^{13}C (ppm)	^{1}H (ppm)	^{13}C (ppm)	^{1}H (ppm)
G2	guaiacyl	111.9	6.99	111.6	7.00	111.6	7.0
G5	guaiacyl	115.5 & 115.8	6.67 & 6.97	115.7 & 115.8	6.78 & 6.97	114.9 & 115.9	6.77 & 6.93
G6	guaiacyl	119.9	6.99	119.8	6.83	119.7	6.86
G'2	oxidized guaiacyl	112.2	7.53	112.4	7.53	112.2	7.51
G'6	oxidized guaiacyl	124.2	7.16	123.8	7.60	123.6	7.57
S2/6	syringyl	104.3	6.71	104.6	6.73	-	-
S'2/6	oxidized syringyl	107.3	7.32 & 7.12	107.2	7.24 & 7.09	-	-
H2/6		126.9	7.20	128.6	7.22	-	-
Aα (G)	β-O-4'	71.8	4.65	71.8	4.65	71.9	4.77
Aα (S)	β-O-4'	72.6	4.88	72.6	4.88		
Aβ(G)	β-O-4'	84.5	4.31	84.6	4.31	84.6	4.31
Aβ(S)	β-O-4'	86.7	4.14	86.7	4.14		
Aγ	β-O-4'	60.5	3.42 & 3.67	-	-	-	-
Bα	β-β' (resinol)	85.8	4.68	84.6	4.76	85.6	4.65
Bβ	β-β' (resinol)	53.8	3.14	54.3	3.10	53.9	3.14
Bγ	β-β' (resinol)	71.7	3.88 & 4.20	70.5	3.73	70.3	3.77
Cα	β-5' (phenylcoumaran)	87.7	5.49	87.7	5.49	87.6	5.49
Cβ	β-5' (phenylcoumaran)	53.8	3.14	54.1	3.49	54.2	3.42
Cγ	β-5' (phenylcoumaran)	62.7	3.44 & 3.59	-	-	-	-

Table S7.4 (cont'd)

Dα	5-5' (dibenzodioxocin)	83.8	4.90	-	-	-	-
Dβ	5-5' (dibenzodioxocin)	85.6	3.91	85.6	3.92	84.9	3.88
Eα	β-1' (spirodienone)	81.8	5.07	81.9	5.08	83.7	5.02
Eα'	β-1' (spirodienone)	86.9	4.41	87.0	4.42	-	-
Eβ	β-1' (spirodienone)	60.4	2.79	60.7	2.80	61.8	2.8
Eβ'	β-1' (spirodienone)	-	-	78.7	4.14	-	-
E2/2'	β-1' (spirodienone)	114.9	6.24	114.2	6.23	114.3	6.21
E6/6'	β-1' (spirodienone)	119.5	6.12	119.2	6.15	119.5	6.11
Iα	cinnamyl alcohol	128.3	6.46	129.0	6.45	129.5	6.48
Iβ	cinnamyl alcohol	129.7	6.25	129.5	6.27	129.1	6.32
Iγ	cinnamyl alcohol	62.4	4.11	62.3	4.12	62.5	4.13
Jα	cinnamaldehyde	154.5	7.65	154.2	7.64	154.4	7.66
Jβ	cinnamaldehyde	126.7	6.81	127.1	6.80	127.4	6.88
J6	cinnamaldehyde	122.4	7.21	123.9	7.22	124.3	7.20
PB2/6	*p*-hydroxybenzoate	-	-	132.1	7.70	-	-
Stα/β	stilbene	-	-	-	-	127.1	6.88

Table S7.5. Average solution NMR chemical shifts of lignin from literature. The reference (Ref.) for each chemical shift is provided. Not detected (-).

	Eucalyptus			Poplar			Spruce		
	^{13}C (ppm)	^{1}H (ppm)	Ref.	^{13}C (ppm)	^{1}H (ppm)	ref	^{13}C (ppm)	^{1}H (ppm)	ref
G2	111.3	6.98	89, 96-99	111.1	6.98	88, 100-103	110.3	6.94	104, 105
G5	115.2	6.72	89, 96-99	115.0	6.80	88, 100-103	115.2	6.85	104, 105
		6.94	89, 96-99	-	-	-	-	-	-
G6	119.3	6.80	89, 96-99	119.1	6.8	88, 100-103	118.9	6.83	104, 105
G'2	112.4	7.50	98	111.4	7.51	100	-	-	98, 100
G'6	124.0	7.00	98	123.3	7.60	100	-	-	100
S2/6	104.2	6.99	89, 96-99	103.8	6.67	88, 100-103	-	-	-
S'2/6	106.5	7.31	89, 96-99	106.1	7.22	88, 100-103	-	-	-
H2/6	-	-	-	127.9	7.19	100	-	-	-
Aα(G)	71.6	4.74	89, 96-99	71.5	4.75	101	71.1	4.75	104, 106
Aα (S)	72.2	4.87	89, 96-99	71.9	4.83	88, 101-103	-	-	104, 106
Aβ(G)	83.9	4.29	89, 96-99	83.7 [a]	4.29	88, 100, 101	84.0	4.29	104, 106
Aβ(S)	86.4	4.09	89, 96-99	85.8	4.13	88, 100-103	-	-	104, 106
Aγ	59.8	3.39	96-99	59.9	3.45	100-103	59.9	3.24	104, 106
		3.69	96-99			100, 101		3.61	104, 106
Bα	85.2	4.66	89, 96-99	85.0	4.64	88, 100-103	85.0	4.65	104, 106
Bβ	53.7	3.08	96-99	53.6	3.07	100-103	53.6	3.07	104, 106
Bγ	71.4	3.82	96-99	70.6	3.81	100, 101	70.9	3.74	104, 106
Cα	87.1	5.46	89, 96-99	87.0	5.47	88, 100-103	87.0	5.47	104, 106
Cβ	53.5	3.45	96-99	53.2	3.48	100-102	53.1	3.46	104, 106
Cγ	62.7	3.67	96-99	62.5	3.74	100, 101	62.8	3.72	104, 106
Dα	83.6	4.83	96	81.4	4.7	101	83.3	4.84	104, 106
Dβ	85.9	3.88	96	-	-	-	85.4	3.88	104, 106
Eα	81.4	5.09	96, 97	81.2	5.07	100	81.7	5.03	106
Eα'	86.6	4.39	96	-	-	-	-	-	-
Eβ	60.0	2.74	96, 97	59.7	2.77	100	59.3	2.77	106
Eβ'	79.5	4.11	96, 97	79.5	4.12	100	-	-	-
E2/2'	112.4	6.26	96, 97	113.3	6.22	100	-	-	96, 97, 100
E6/6'	119.1	6.07	96, 97	118.9	6.07	100	-	-	96, 97, 100
Iα	-	-	-	128.4	6.44	100	-	-	-
Iβ	-	-	-	128.2	6.25	100	-	-	-

Table S7.5 (cont'd)

Iγ	61.7	4.09	[96-98]	61.4	4.10	[100]	61.5	4.08	[104, 106]
Jα	-	-	-	153.5	7.60	[100]	153.7	7.61	[104]
Jβ	-	-	-	126.1	6.76	[100]	126.3	6.77	[104]
J6	-	-	-	-	-	-	123.2	7.22	[104]
PB2/6	-	-	-	131.1	7.66	[88, 100-103]	-	-	[104]
Stα/β	126	7.0	[99]	-	-	-	128.3	7.12	[104]

Table S7.6. Spectral deconvolution of quantitative ^{13}C spectra for lignin compositional analysis. The attributed chemical shifts are labeled. Unidentified (-).

Eucalyptus				
Chemical Shift [ppm]	Amplitude	Width [ppm]	Integral [%]	Attribution
180.6	117	4.2	2.2	-
177.9	170	2.6	1.9	-
173.8	956	3.9	16.2	Ac^{CO}
171.8	211	3.2	3.0	-
169.4	306	1.5	2.0	-
168.4	314	0.4	0.5	-
167.5	385	0.4	0.7	-
157.3	36	0.1	0.02	-
152.2	401	3.5	6.2	S3/5
147.7	171	3.4	2.6	G3
144.8	550	2.6	6.3	G4/S'3
138.4	63	2.6	0.7	$S'^{b}4$
136.4	341	4.4	6.6	$S1/S'^{a}4$
132.2	149	5.5	3.6	-
129.8	79	0.6	0.2	-
128.2	128	1.9	1.1	-
125.0	133	2.0	1.2	-
119.9	70	3.8	1.2	
114.7	195	6.0	5.2	G5
109.4	185	4.9	4.0	S'2/6
107.6	229	0.9	0.9	A1
104.6	2261	2.6	25.5	$i/s/Xn1^{2f}$
102.2	75	0.3	0.1	$Xn1^{3f}$
100.8	444	2.4	4.6	Man1
98.5	281	1.9	2.4	-
96.4	209	0.8	0.7	-
92.6	135	0.8	0.5	-

Poplar				
Chemical Shift [ppm]	Amplitude	Width [ppm]	Integral [%]	Attribution
181.0	109	6.7	3.1	
173.6	1037	3.2	14.1	Ac^{CO}
171.7	173	1.1	0.8	
152.6	744	3.8	12.2	S3/5
148.2	217	3.3	3.0	G3
145.2	151	2.3	1.5	G4
136.9	165	2.9	2.0	$S^{b}4$
134.2	315	3.8	5.1	$S^{a}4$

Table S7.6 (cont'd)

131.2	193	3.9	3.2	
120.0	42	2.4	0.4	
115.5	188	8.4	6.8	G5
108.9	16	1.9	0.1	
104.7	3650	2.6	40.1	i/s/Xn1^{2f}
102.4	60	0.5	0.1	Xn1^{3f}
100.6	632	2.7	7.2	
96.5	62	0.3	0.1	

Spruce

Chemical Shift [ppm]	Amplitude	Width [ppm]	Integral [%]	Attribution
183.0	96	3.3	1.2	
181.7	248	1.4	1.3	
178.0	206	2.5	1.9	
174.1	836	3.1	9.5	Ac^{CO}
171.9	224	2.1	1.7	
169.5	36	1.1	0.1	
157.4	375	0.4	0.6	
153.4	323	5.3	6.3	
148.2	466	5.3	9.2	G^{a}3/4
145.5	289	3.6	3.8	G^{b}3/4
137.1	89	2.8	0.9	G^{b}1
134.6	163	2.6	1.6	G$^{c/d}$1
132.1	322	3.8	4.5	
129.9	389	1.2	1.7	G^{a}1
128.6	266	1.7	1.7	
125.0	52	7.6	1.4	
120.1	57	2.9	0.6	G6
114.7	389	14.2	20.6	G2/5
107.7	623	0.9	2.1	A1
104.8	2103	2.6	18.4	
101.8	275	1.6	1.6	
100.4	907	1.4	4.6	M1
98.8	492	1.5	2.8	
96.5	477	0.7	1.2	β-Glc1
92.6	308	0.7	0.7	α-Glc1

Table S7.7. The Intensities of intermolecular cross peaks of eucalyptus. In sum, 98 intermolecular interactions are identified, including 45 strong, 29 medium, and 24 weak restraints. In gated 1-s PDSD spectra, a peak higher than 4% is categorized as a strong restraint (s, in bold), between 2% - 4% for medium (m) restraint, and below 2% for weak (w) restraint. All the peaks shown in the gated 100-ms PDSD spectrum are categorized as strong restraints. The intensity is a relative ratio of the peak, which is normalized by the integral of the ^{13}C cross-section.

	Atom1 (ω_1)	Atom2 (ω_2)	Gated 100 ms PDSD (%)	Gated 1s PDSD (%)	Type
cellulose-xylan	Ac^{CO}	i4		0.13	w
	Ac^{Me}	i4		1.40	w
	Ac^{CO}	s4		0.16	w
	Ac^{Me}	s4		0.85	w
cellulose-lignin	G3	i4		1.84	w
	G4,S'3/5	i4		0.49	w
	G5	i4		3.29	m
	S1/4^a,S'1^a	i4	**2.11**	**7.81**	**s**
	S/S'4^b	i4		1.28	w
	S3/5	i4		1.27	w
	OMe	i4	**0.67**	**0.99**	**s**
	G3	s4		2.39	m
	G3	s6	**0.94**	**4.05**	**s**
	G4,S'3/5	s4		1.24	w
	G4,S'3/5	s6		3.49	m
	G5	s4		0.83	w
	S1/4^a,S'1^a	s4		2.35	m
	S1/4^a,S'1^a	s6	**1.23**	**5.49**	**s**
	S3/5	s4		1.34	w
	S3/5	s6		4.21	**s**
	S/S'4^b	s4		1.24	w
	S/S'4^b	s6		**5.25**	**s**
	OMe	s4	**0.98**	**1.77**	**s**
	OMe	s6	**1.34**	**3.39**	**s**
lignin-lignin	G3	S1/4^a,S'1^a		3.42	m
	G3	S2/6^b		2.41	m
	G3	S3/5		7.00	**s**
	G3	S/S'4^b		3.37	m
	G4,S'3/5	S2/6^a		1.34	w
	G4,S'3/5	S2/6^b		3.45	m

Table S7.7 (cont'd)

	G4,S'3/5	S3/5		**4.75**	**s**
	G5	S1/4^{a},S'1^{a}	**1.36**	**7.34**	**s**
	G5	S2/6^{a}		3.26	m
	G5	S3/5		5.18	**s**
	G5	S/S'4^{b}		3.91	m
	S1/4^{a},S'1^{a}	G3		3.87	m
	S3/5	G2		2.40	m
	S3/5	G3		1.88	w
	S3/5	G4,S'3/5		1.88	w
	S3/5	G5		0.48	w
	S3/5	S'1^{b}/4^{a}		3.39	m
	S/S'4^{b}	G3		2.58	m
	S/S'4^{b}	G5		2.20	m
xylan-lignin	Ac^{CO}	G4,S'3/5		2.26	m
	Ac^{CO}	G5		0.47	w
	Ac^{CO}	S1/4^{a},S'1^{a}		1.83	w
	Ac^{CO}	S3/5		2.01	m
	Ac^{CO}	OMe		**4.61**	**s**
	G3	Ac^{CO}		3.85	m
	G4,S'3/5	Ac^{CO}	**3.12**	**8.98**	**s**
	G5	Ac^{CO}		**6.87**	**s**
	S1/4^{a},S'1^{a}	Ac^{CO}	**2.36**	**9.32**	**s**
	S3/5	Ac^{CO}		3.58	m
	S/S'4^{b}	Ac^{CO}	**2.65**	**8.08**	**s**
	OMe	Ac^{CO}		**6.01**	**s**
	Ac^{Me}	G4,S'3/5		1.56	w
	Ac^{Me}	S3/5		2.33	m
	Ac^{Me}	S/S'4^{b}		1.95	w
	Ac^{Me}	OMe		2.85	m
	G3	Ac^{Me}		**5.12**	**s**
	G4,S'3/5	Ac^{Me}	**0.99**	**6.31**	**s**
	G5	Ac^{Me}	**1.19**	**6.94**	**s**
	S1/4^{a},S'1^{a}	Ac^{Me}		**4.26**	**s**
	S3/5	Ac^{Me}		3.60	m
	S/S'4^{b}	Ac^{Me}		**4.27**	**s**
	OMe	Ac^{Me}		**4.34**	**s**
	G3	Xn4^{2f}		3.34	m
	G3	Xn5$^{2f, 3f}$		3.34	m
	G4,S'3/5	Xn4^{2f}		1.59	w
	G4,S'3/5	Xn4^{3f}		1.19	w
	G4,S'3/5	Xn5$^{2f, 3f}$	**1.68**	5.03	**s**
	G5	Xn4^{2f}		2.11	m
	G5	Xn4^{3f}		2.81	m
	G5	Xn5$^{2f, 3f}$		1.33	w

Table S7.7 (cont'd)

	S1/4[a],S'1[a]	Xn4^{2f}		**5.29**	**s**
	S1/4[a],S'1[a]	Xn4^{3f}	**0.64**	**2.71**	s
	S1/4[a],S'1[a]	Xn5$^{2f, 3f}$		**5.64**	**s**
	S3/5	Xn4^{2f}		3.42	m
	S3/5	Xn5$^{2f, 3f}$		**4.92**	**s**
	S/S'4[b]	Xn4^{2f}		2.86	m
	S/S'4[b]	Xn4^{3f}		**4.07**	**s**
	S/S'4[b]	Xn5$^{2f, 3f}$		**4.15**	**s**
	OMe	Xn4^{2f}	**0.96**	**3.66**	s
	OMe	Xn4^{3f}	**1.34**	**1.55**	**s**
	OMe	Xn5$^{2f, 3f}$	**1.67**	**3.85**	**s**
mixed sugar-lignin	G3	i/s/Xn1^{2f}	**2.61**	**6.08**	**s**
	G3	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3	**4.36**	**23.70**	**s**
	G4,S'3/5	i/s/Xn1^{2f}		1.23	w
	G4,S'3/5	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3	**3.96**	**18.03**	**s**
	G5	i/s/Xn1^{2f}		**5.84**	**s**
	G5	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3	**7.36**	**22.50**	**s**
	S1/4[a],S'1[a]	i/s/Xn1^{2f}	**2.71**	**8.47**	**s**
	S3/5	i/s/Xn1^{2f}	**2.62**	**7.68**	**s**
	S3/5	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3		26.12	**s**
	S/S'4[b]	i/s/Xn1^{2f}		2.22	m
	S/S'4[b]	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3		**17.08**	**s**
	OMe	i/s/Xn1^{2f}		**8.71**	**s**
	OMe	Xn2/3^{3f},Xn3^{2f},i2/5,s2/3		**14.56**	**s**

Table S7.8. The Intensities of intermolecular cross peaks of poplar wood. In sum, 80 intermolecular interactions are identified, including 22 strong, 33 medium, and 25 weak restraints. In gated 1-s PDSD spectra, a peak higher than 4% is categorized as a strong restraint (s, in bold), between 2% - 4% for medium (m) restraint, and below 2% for weak (w) restraint. All the peaks shown in the gated 100-ms PDSD spectrum are categorized as strong restraints. The intensity is a relative ratio of the peak, which is normalized by the integral of the ^{13}C cross-section.

	Atom1 (ω_1)	Atom2 (ω_2)	Gated 100 ms PDSD (%)	Gated 1s PDSD (%)	Type
cellulose-xylan	Ac^{Me}	i4		2.16	m
	Ac^{CO}	s4		3.37	m
	Ac^{Me}	s4		1.76	w
	Ac^{Me}	s6		2.13	m
cellulose-lignin	G4	i4		1.80	w
	$S1^a$	i4		2.36	m
	$S1^b$	i4		2.51	m
	S3/5	i4		1.11	w
	OMe	i4		0.90	w
	$S1^a$	s4		1.78	w
	$S1^a$	s6		**4.26**	**s**
	$S1^b$	s4		1.55	w
	$S1^b$	s6		**5.32**	**s**
	S3/5	s4		3.21	m
	S3/5	s6		**6.04**	**s**
	G4	s4		2.52	m
	G4	s6		**4.01**	**s**
	OMe	s4		1.89	w
	OMe	s6		3.07	m
lignin-lignin	$S1^a$	$G2^a$		2.15	m
	$S1^a$	G3		2.88	m
	$S1^a$	G4		1.98	w
	$S1^a$	G5		2.62	m
	$S1^b$	$G2^a$		1.86	w
	$S1^b$	$G2^b$		1.27	w
	$S1^b$	G3		1.85	w
	$S1^b$	G4		1.81	w
	$S1^b$	G5		1.31	w
	G4	$S1^a$		**4.36**	**s**

Table S7.8 (cont'd)

	G4	S1[b]		1.67	w
	G4	S2/6[a]		3.67	m
	G4	S3/5	**2.97**	**9.76**	**s**
	S3/5	G3		2.48	m
	S3/5	G4		1.13	w
	S3/5	G5		2.20	m
xylan-lignin	Ac^{CO}	G2[a]		2.02	m
	Ac^{CO}	G3		1.28	w
	Ac^{CO}	G4		1.12	w
	Ac^{CO}	G5		1.38	w
	Ac^{CO}	S1[a]		1.37	w
	Ac^{CO}	S1[b]		0.57	w
	Ac^{CO}	S2/6[b]		2.44	m
	Ac^{CO}	S3/5		2.71	m
	Ac^{CO}	OMe		**5.09**	**s**
	G4	Ac^{CO}		3.91	m
	S1[a]	Ac^{CO}		2.48	m
	S1[b]	Ac^{CO}		**6.49**	**s**
	S3/5	Ac^{CO}		**4.58**	**s**
	OMe	Ac^{CO}		**6.99**	**s**
	Ac^{Me}	G2[b]		0.80	w
	Ac^{Me}	G4		0.65	w
	Ac^{Me}	S1[b]		0.51	w
	Ac^{Me}	S2/6[b]		**4.91**	**s**
	Ac^{Me}	S3/5		3.08	m
	Ac^{Me}	S4		1.39	w
	Ac^{Me}	OMe		3.48	m
	G4	Ac^{Me}		3.48	m
	S1[a]	Ac^{Me}		3.81	m
	S1[b]	Ac^{Me}		**5.00**	**s**
	S3/5	Ac^{Me}	**1.44**	**2.87**	**s**
	OMe	Ac^{Me}	**1.62**	**3.91**	**s**
	G4	$Xn4^{2f}$		2.37	m
	G4	$Xn4^{3f}$		3.93	m
	G4	$Xn5^{2f, 3f}$		2.55	m
	S1[a]	$Xn4^{2f}$		2.43	m
	S1[a]	$Xn4^{3f}$		3.97	m
	S1[a]	$Xn5^{2f, 3f}$	**1.17**	**3.58**	**s**
	S1[b]	$Xn4^{2f}$		1.72	w
	S1[b]	$Xn4^{3f}$		2.28	m
	S1[b]	$Xn5^{2f, 3f}$		**5.16**	**s**
	S3/5	$Xn4^{2f}$		2.15	m
	S3/5	$Xn4^{3f}$		3.36	m
	S3/5	$Xn5^{2f, 3f}$		**4.66**	**s**

Table S7.8 (cont'd)

	OMe	$Xn4^{2f}$	**1.88**	**3.29**	**s**
	OMe	$Xn4^{3f}$	**1.56**	**2.53**	**s**
	OMe	$Xn5^{2f, 3f}$	**2.09**	**4.13**	**s**
mixed sugar-lignin	G4	$i/s/Xn1^{2f}$		2.33	m
	S1[a]	$i/s/Xn1^{2f}$		2.92	m
	S3/5	$i/s/Xn1^{2f}$	**2.47**	**5.04**	**s**
	OMe	$i/s/Xn1^{2f}$	**1.95**	**11.46**	**s**

Table S7.9. The Intensities of intermolecular cross peaks of poplar spruce. In sum, 94 intermolecular interactions are identified, including 51 strong, 9 medium, and 34 weak restraints. In gated 1-s PDSD spectra, a peak higher than 4% is categorized as a strong restraint (s, in bold), between 2% - 4% for medium (m) restraint, and below 2% for weak (w) restraint. All the peaks shown in the gated 100-ms PDSD spectrum are categorized as strong restraints. The intensity is a relative ratio of the peak, which is normalized by the integral of the ^{13}C cross-section.

	Atom1 (ω_1)	Atom2 (ω_2)	Gated 100 ms PDSD (%)	Gated 1s PDSD (%)	Type
cellulose-xylan	Ac^{CO}	i4		0.66	w
	Ac^{CO}	i6		3.06	m
	Ac^{Me}	i4		1.38	w
	Ac^{Me}	i6		2.08	m
	Ac^{CO}	s4	**0.96**	**3.08**	**s**
	Ac^{CO}	s6	**2.06**	**13.45**	**s**
	Ac^{Me}	s4	**1.21**	**3.92**	**s**
	Ac^{Me}	s6	**2.23**	**9.84**	**s**
cellulose-lignin	$G1^{a}$	i4	**1.12**	**2.50**	**s**
	$G1^{a}$	i6	**0.69**	**2.26**	**s**
	$G1^{b,c}$	i4	**0.36**	**1.10**	**s**
	$G1^{d}$	i4	**0.68**	**1.52**	**s**
	$G3^{a}$	i4	**1.03**	**2.93**	**s**
	$G3^{a}$	i6	**0.78**	**1.29**	**s**
	$G3^{c}$	i4		1.47	w
	$G4^{a,b,c}$	i4	**2.66**	**4.48**	**s**
	OMe	i4	**3.98**	**9.61**	**s**
	$G1^{a}$	s4		2.20	m
	$G1^{a}$	s6	**2.36**	**5.98**	**s**
	$G1^{b,c}$	s4	**1.22**	**2.39**	**s**
	$G1^{b,c}$	s6	**2.94**	**8.96**	**s**
	$G1^{d}$	s4		1.82	w
	$G3^{a}$	s4		1.44	w
	$G3^{a}$	s6	**2.36**	**5.05**	**s**
	$G3^{c}$	s4	**1.07**	**2.21**	**s**
	$G3^{c}$	s6	**1.98**	**8.07**	**s**
	$G4^{a,b,c}$	s4	**1.17**	**3.57**	**s**
	$G4^{a,b,c}$	s6	**2.32**	**6.63**	**s**
	OMe	s4	**3.63**	**9.99**	**s**
	OMe	s6	**4.01**	**23.53**	**s**

Table S7.9 (cont'd)

	Ac^{CO}	$G1^{a}$		0.73	w
	Ac^{CO}	$G1^{b,c}$		0.67	w
	Ac^{CO}	$G1^{d}$		0.93	w
	Ac^{CO}	$G2^{a,c}$		1.98	w
	Ac^{CO}	$G2^{d}$		1.95	w
	Ac^{CO}	$G3^{a}$	**1.06**	**4.90**	**s**
	Ac^{CO}	$G3^{b}$		0.11	w
	Ac^{CO}	$G3^{c}$		2.61	m
	Ac^{CO}	$G3^{d}$		0.24	w
	Ac^{CO}	$G4^{a,b,c}$		3.22	m
	Ac^{CO}	$G6^{c}$		1.57	w
	Ac^{CO}	OMe	**2.59**	**7.82**	**s**
	$G1^{a}$	Ac^{CO}	**2.36**	**6.31**	**s**
	$G1^{b,c}$	Ac^{CO}	**1.29**	**3.06**	**s**
	$G1^{d}$	Ac^{CO}		1.92	w
	$G3^{a}$	Ac^{CO}	**2.22**	**5.09**	**s**
	$G3^{c}$	Ac^{CO}	**1.03**	**2.09**	**s**
	$G4^{a,b,c}$	Ac^{CO}	**0.69**	**1.55**	**s**
	OMe	Ac^{CO}	**1.19**	**3.49**	**s**
	Ac^{Me}	$G1^{a}$		1.33	w
	Ac^{Me}	$G1^{b,c}$		0.30	w
	Ac^{Me}	$G2^{a,c}$		1.65	w
xylan-lignin	Ac^{Me}	$G3^{a}$		2.26	m
	Ac^{Me}	$G3^{c}$		3.00	m
	Ac^{Me}	$G3^{d}$		0.67	w
	Ac^{Me}	$G4^{a,b,c}$		1.66	w
	Ac^{Me}	$G4^{d}$		1.67	w
	Ac^{Me}	$G6^{c}$		1.64	w
	Ac^{Me}	OMe	**1.49**	**5.37**	**s**
	$G1^{a}$	Ac^{Me}		3.40	m
	$G1^{b,c}$	Ac^{Me}		1.36	w
	$G1^{d}$	Ac^{Me}		1.20	w
	$G3^{a}$	Ac^{Me}		0.41	w
	$G3^{c}$	Ac^{Me}		0.72	w
	$G4^{a,b,c}$	Ac^{Me}		1.59	w
	OMe	Ac^{Me}	**1.20**	**2.29**	**s**
	$G1^{a}$	$Xn4^{2f}$	**0.86**	**3.88**	**s**
	$G1^{a}$	$Xn4^{3f}$		1.13	w
	$G1^{b,c}$	$Xn4^{2f}$	**0.98**	**2.75**	**s**
	$G1^{b,c}$	$Xn4^{3f}$	**1.12**	**2.27**	**s**
	$G1^{d}$	$Xn4^{2f}$	**1.21**	**2.09**	**s**
	$G1^{d}$	$Xn4^{3f}$	**1.51**	**3.06**	**s**
	$G1^{d}$	$Xn5^{2f, 3f}$	**2.67**	**6.47**	**s**
	$G3^{a}$	$Xn4^{2f}$	**1.36**	**2.82**	**s**

Table S7.9 (cont'd)

	G3[a]	Xn4[3f]	**1.54**	**4.14**	**s**
	G3[a]	Xn5[2f, 3f]	**1.11**	**2.46**	**s**
	G3[c]	Xn4[2f]	**2.01**	**4.43**	**s**
	G3[c]	Xn4[3f]		1.94	w
	G4[a,b,c]	Xn4[2f]	**1.86**	**4.48**	**s**
	G4[a,b,c]	Xn4[3f]	**0.87**	**2.48**	**s**
	G4[a,b,c]	Xn5[2f, 3f]	**0.99**	**2.79**	**s**
	OMe	Xn4[2f]	**4.69**	**13.7**	**s**
	OMe	Xn4[3f]	**3.96**	**9.92**	**s**
mannan-lignin	G3[c]	Man1		1.19	w
	OMe	Man1		2.59	m
mannan-xylan	Ac[CO]	Man1	**1.32**	**2.04**	**s**
mixed sugar-lignin	G1[a]	i/s/Xn1[2f]		0.94	w
	G1[b,c]	i/s/Xn1[2f]		0.60	w
	G1[d]	i/s/Xn1[2f]		0.58	w
	G3[a]	i/s/Xn1[2f]		1.53	w
	G3[c]	i/s/Xn1[2f]		1.43	w
	G4[a,b,c]	i/s/Xn1[2f]	**0.66**	**1.53**	**s**
	OMe	i/s/Xn1[2f]	**2.61**	**5.99**	**s**
	OMe	Xn2/3[3f],Xn3[2f],i2/5,s2/3	**3.39**	**15.46**	**s**

Table S7.10. Intermolecular interactions of polymers in intact plant stems. In the summary of intermolecular cross-peaks between different polymers of the secondary cell wall of woods, a total of 272 cross-peaks have been identified and categorized as 112 strong, 75 medium, and 85 weak interactions.

	Interactions	**s**	**m**	**w**	**Total**
Eucalyptus	cellulose-xylan	0	0	4	**4**
	cellulose-lignin	8	4	8	**20**
	lignin-lignin	4	11	4	**19**
	xylan-lignin	22	13	7	**42**
	mixed sugar-lignin	11	1	1	**13**
Poplar	cellulose-xylan	0	3	1	**4**
	cellulose-lignin	4	5	6	**15**
	lignin-lignin	2	6	8	**16**
	xylan-lignin	14	17	10	**41**
	mixed sugar-lignin	2	2	0	**4**
Spruce	cellulose-xylan	4	2	2	**8**
	cellulose-lignin	18	1	3	**22**
	lignin-lignin	0	0	0	**0**
	xylan-lignin	25	5	23	**53**
	Mannan-lignin	0	1	1	**2**
	Mannan-xylan	1	0	0	**1**
	mixed sugar-lignin	3	0	5	**8**
Total		**112**	**75**	**85**	**272**

Table S7.11. Water-edited intensities of polysaccharide and lignin of Eucalyptus. The intensity ratio is obtained by comparing the water-edited and control 2D ^{13}C-^{13}C correlation spectra. Standard deviations of NMR signal-to-noise ratios are used as error bars. The numbers (bold) in parentheses indicate the average intensities under each type of the polysaccharide or lignin.

Type	cross peaks	Intensity
interior cellulose **(0.24)**	i4-1	0.44±0.03
	i4-3[a]	042±0.05
	i4-3[b,c]	0.42±0.03
	i4-2/5	0.42±0.09
	i4-6	0.40±0.07
	i6-1	0.48±0.19
	i6-4	0.4±0.2
	i6-3[a,b]	0.5±0.1
	i6-3[c]	0.5±0.1
	i6-2/5,2	0.5±0.1
surface cellulose **(0.36)**	s4-1	0.55±0.09
	s4-3/5	0.6±0.1
	s4-2	0.6±0.1
	s4-6	0.5±0.1
xylan^{2f} **(0.31)**	Xn^{2f}4-1	0.52±0.09
	Xn^{2f}4-3	0.55±0.08
	Xn^{2f}4-5	0.47±0.09
	Xn^{2f}4-2	0.7±0.2
xylan^{3f} **(0.44)**	Xn^{3f}2/3-1	0.7±0.2
	Xn^{3f}2/3-4	0.6±0.2
	Xn^{3f}2/3-5	0.6±0.2
	Xn^{3f}5-1	0.5±0.1
	Xn^{3f}5-4	0.5±0.1
lignin **(0.51)**	S3/5-S1[b]	0.50±0.08
	S3/5-S1[a]	0.48±0.03
	S3/5-S2/6[a]	0.61±0.04
	S3/5-S2/6[b]	0.49±0.09
	S3/5-OMe	0.53±0.08
	S’3/5-4	0.61±0.09
	S’3/5-2, G4-2	0.5±0.1
	S’3/5-OMe, G4-OMe	0.5±0.1
	S4-S3/5	0.46±0.09
	S4-S2/6	0.44±0.08
	S4-S1	0.47±0.08
	S4-OMe	0.51±0.09

Table S7.11 (cont'd)

	S2/6-S3/5	0.6±0.2
	S2/6-S1	0.5±0.2
	S2/6-S4	0.5±0.1
	S2/6-OMe	0.51±0.09
lignin OMe **(0.42)**	OMe-S3/5	0.42±0.06
	OMe-G3	0.43±0.07
	OMe-S1	0.40±0.09
	OMe-S4	0.5±0.1
	OMe-S2/6	0.4±0.1

Table S7.12. Water-edited intensities of polymers in poplar. The intensity ratio is obtained by comparing the water-edited and control 2D ^{13}C-^{13}C correlation spectra. Standard deviations of NMR signal-to-noise ratios are used as error bars. The numbers (bold) in parentheses indicate the average intensities under each type of the polysaccharide or lignin.

Type	cross peaks	Absolute Intensity
interior cellulose **(0.11)**	i4-1	0.11±0.04
	i4-3[a]	0.09±0.07
	i4-3[b,c]	0.09±0.07
	i4-2/5	0.10±0.08
	i6-1	0.12±0.05
	i6-4	0.12±0.03
	i6-3[a,b]	0.11±0.03
	i6-3[c]	0.11±0.04
	i6-2/5	0.11±0.02
surface cellulose **(0.11)**	s4-1	0.12±0.06
	s4-3/5	0.10±0.04
	s4-2	0.11±0.04
	s4-6	0.11±0.03
xylan[2f] **(0.30)**	Xn4[2f]-1	0.36±0.06
	Xn4[2f]-3	0.42±0.06
	Xn4[2f]-2	0.20±0.05
	Xn4[2f]-5	0.21±0.06
xylan[3f] **(0.25)**	Xn[3f]2/3-1	0.23±0.05
	Xn[3f]2/3-4	0.39±0.09
	Xn[3f]2/3-5	0.15±0.03
	Xn[3f]5-1	0.31±0.07
	Xn[3f]5-4	0.18±0.09
guaiacyl **(0.19)**	G5-3	0.14±0.04
	G5-1	0.14±0.03
	G5-OMe	0.18±0.02
	G2-3	0.28±0.06
	G2-1	0.16±0.05
	G2-6	0.21±0.04
	G2-OMe	0.21±0.03
syringyl **(0.16)**	S3/5-1[b]	0.16±0.09
	S3/5-1[a]	0.21±0.06
	S3/5-2/6	0.15±0.04
	S3/5-OMe	0.14±0.05
	S1[b]-S3/5	0.14±0.07
	S1[b]-S2/6	0.20±0.06
	S1[b]-OMe	0.15±0.06

Table S7.12 (cont'd)

	S1[a]-S3/5	0.13±0.03
	S1[a]-OMe	0.11±0.03
	S2/6-S3/5	0.17±0.04
	S2/6-S1[b]	0.20±0.06
	S2/6-OMe	0.11±0.03
syringyl/guaicyl OMe **(0.14)**	OMe-S3/5	0.13±0.04
	OMe-G3	0.14±0.03
	OMe-S1	0.14±0.04
	OMe-G5	0.15±0.05
	OMe-S2/6	0.15±0.03

Table S7.13. Water-edited intensities of polysaccharides and lignin of spruce. The intensity ratio is obtained by comparing the water-edited and control 2D ^{13}C-^{13}C correlation spectra. Standard deviations of NMR signal-to-noise ratios are used as error bars. The numbers (bold) in parentheses indicate the average intensities under each type of the polysaccharide or lignin.

Type	cross peaks	Absolute Intensity
interior cellulose **(0.38)**	i4-1	0.44±0.02
	i4-3[a]	0.37±0.04
	i4-3[b,c]	0.37±0.03
	i4-2/5	0.39±0.03
	i6-1	0.42±0.07
	i6-4	0.4±0.1
	i6-3	0.32±0.05
	i6-2/5,2	0.4±0.1
surface cellulose **(0.43)**	s4-1	0.5±0.1
	s4-2/5	0.44±0.03
	s4-3	0.4±0.2
	s4-6	0.46±0.03
xylan^{2f} **(0.57)**	Xn^{2f}4-1	0.7±0.1
	Xn^{2f}4-3	0.43±0.04
	Xn^{2f}4-5	0.57±0.04
	Xn^{2f}4-1	0.67±0.05
	Xn^{2f}4-2	0.56±0.06
xylan ^{3f} **(0.40)**	Xn^{3f}2/3-1	0.4±0.1
	Xn^{3f}2/3-4	0.32±0.07
	Xn^{3f}2/3-5	0.4±0.1
	Xn^{3f}5-4	0.44±0.05
	Xn^{3f}5-1	0.48±0.04
mannan **(0.76)**	Man1-4	0.86±0.06
	Man1-5	0.63±0.06
	Man1-6	0.79±0.05
guaiacyl **(0.47)**	G3-1	0.4±0.1
	G3-2	0.47±0.04
	G3-OMe	0.52±0.09
	G1-3	0.5±0.1
	G6-3	0.40±0.06
	G6-OMe	0.35±0.03
	G2-3[b,c]	0.51±0.06
	G2-3[a]	0.65±0.05
	G2-OMe	0.54±0.15
guaiacyl OMe **(0.42)**	OMe-G1	0.40±0.06
	OMe-G2	0.42±0.07

Table S7.14. ^{13}C-T_1 relaxation times of lignin and polysaccharides in the three woods. The data are fit using single exponential equations $I\ (t) = 1 - 2e^{-t/T}$ for T_1, DP (inversion recovery) and $I\ (t) = e^{-t/T}$ for T_1, CP (Torchia CP). Standard deviations of the fitting parameters are used as error bars. Unidentified (-).

	Eucalyptus				Poplar				Spruce			
	Atom (ppm)	T_1, CP (s)	Atom (ppm)	T_1, DP (s)	Atom (ppm)	T_1, CP (s)	Atom (ppm)	T_1, DP (s)	Atom (ppm)	T_1, CP (s)	Atom (ppm)	T_1, DP (s)
Ac^{CO}	174	2.1±0.1	174	1.6±0.2	174	2.83±0.05	174	2.4±0.2	174	1.4±0.2	174	1.3±0.1
S3/5	-	-	153	3.1±0.5	-	-	153	1.3±0.1	-	-	-	-
G3	148	4.6±0.2	-	-	148	4.2±0.2	-	-	148	3.7±0.1	148	-
G4/ S'3/5	145	4.7±0.3	145	1.8±0.2	145	4.1±0.1	145	2.4±0.2	145	3.6±0.1	145	2.4±0.2
S1	136	4.5±0.2	136	2.2±0.3	-	-	136	2.7±0.2	-	-	-	-
S4	134.5	4.7±0.3	134.5	1.9±0.3	134.5	4.2±0.1	134.5	2.7±0.2	-		-	
G1	-	-	-	-	-	-	-	-	130.8	2.9±0.1	130.8	0.9±0.1
G6	-	-	-	-	-	-	-	-	120.5	3.3±0.1	120.5	2.6±0.3
OMe	57	4.6±0.3	57	2.5±0.2	57	4.4±0.3	56.5	2.8±0.2	57	3.4±0.3	57	2.1±0.2
i/s/Xn^{2f}1	105	4.1±0.1	105	2.3±0.7	105	4.1±0.1	105	3.7±0.2	105	3.45±0.04	105	3.2±0.2
i4	89	4.8±0.2	89	4.4±0.2	89	4.8±0.2	89	4.3±0.2	89	3.97±0.06	89	3.9±0.2
s4	84	3.9±0.1	84	2.3±0.3	84	3.7±0.2	84	3.1±0.2	84	3.12±0.09	84	1.5±0.2
i6	65	3.7±0.3	65	2.9±0.3	65	3.6±0.3	65	3.4±0.3	65	3.0±0.2	65	2.4±0.3
s6	62	2.7±0.3	61	-	62	2.7±0.3	-	-	62	2.0±0.2	-	-
Xn^{3f}1	102.5	3.4±0.2	102.5	2.1±0.3	102.5	3.4±0.2	102.5	2.2±0.2	102.5	2.7±0.1	102.5	1.7±0.2
Man	-	-	-	-	-	-	-	-	100.2	2.3±0.1	100	0.8±0.1
Xn^{2f}4	82	3.7±0.2	82	1.4±0.2	82	3.7±0.2	82	2.1±0.4	81.8	2.8±0.1	82	1.1±0.2
Xn^{3f}4	78	3.6±0.2	78	1.2±0.2	78	3.6±0.2	78	1.4±0.3	78	2.5±0.1	78	0.8±0.1
$Xn^{2f/3f}Ac^{Me}$	21.5	3.9±0.1	21.5	2.2±0.2	21.5	4.7±0.2	21.5	1.5±0.1	21.5	2.1±0.2	21.5	1.1±0.1

Table S7.15. ^{1}H-$T_{1\rho}$ relaxation times of lignin and polysaccharides in three samples. The data are fit using a single exponential equation: $I(t) = e^{-t/T}$. Standard deviations of the fitting parameters are used as error bars. Unidentified (-).

	Eucalyptus		Spruce		Poplar	
	Atom (ppm)	T_1 (ms)	Atom (ppm)	T_1 (ms)	Atom (ppm)	T_1 (ms)
S3/5	153.9	16.8±0.8	-	-	153.2	19.5±0.6
G1[a]	-	-	131	8.3±0.6	-	-
G1[b/c]	-	-	134	11.6±0.5	-	-
G1[d]	-	-	136	10.1±0.7	-	-
G3[a]	148.1	10±1	152.2	13±1	148.1	14±1
G3[c]	-	-	147.6	12.8±0.9	-	-
S'3/5/G4	145.6	11.4±0.8	-	-	145.0	13.8±0.9
S4	135.9	12±1	-	8.0±0.5	-	-
G1	-	-	134.2	8.9±0.7	-	-
S'[b]4	130.1	11.1±0.7	-	-	-	-
G6	-	-	119.0	8.8±0.6	-	-
G5	115.3	8.0±0.7	116.3	8.2±0.7	116.2	9.1±0.8
G2	-	-	113.5	8.3±0.6	-	-
S'1	110.9	10.1±0.9	-	-	-	-
OMe	57	14.2±0.7	57	12.5±0.7	57	15.1±0.8
i/s/Xn1[2f]	105.0	28.1±0.9	105.0	29±2	105.0	36±2
i4	89.0	35±1	89.0	35±1	89.0	43±2
s4	84.2	22 ±1	84.1	20±1	84.5	28±2
i6	65.0	26±2	65.0	26±2	64.8	32±2
s6	61.5	16±1	61.5	15±1	61.5	19±1
Xn1[3f]	102.5	12.8±0.9	102.3	12±1	102.4	13.0±0.8
Man1	-	-	101	10.4±0.9	-	-
Xn4[2f]	82.0	14.6±0.9	82.0	13.7±0.9	81.8	17±1
Xn4[3f]	78.1	11.5±0.9	78.0	10.2±0.9	78.0	12±1
Xn[2f,3f]Ac[Me]	21.0	17.6±0.6	21.0	13±1	21.5	19.5±0.5

Table S7.16. Parameters of 1D solid-state NMR experiments measured on the three wood samples. Recycle delay (d1); number of scans (NS); number of points of time domain for the direct (td2) and indirect (td1) dimensions; the acquisition time of the direct dimension (aq2); the evolution time of indirect dimension (aq1); excitation frequency for proton ($\nu^{1}H$) and carbon ($\nu^{13}C$) channels.

	CP	DP	Quantitative DP	^{13}C-T_1 Torchia	^{13}C-T_1 Inversion-recovery	^{1}H-$T_{1\rho}$	1D Water-edited
Field (T)	14.1	14.1	14.1	9.4	9.4	9.4	9.4
Temp. (K)	294	294	294	300	300	300	278
MAS (kHz)	14	14	14	10	10	10	10
d1(s)	1.7	2	35/40[a]	2	30	2	1.8
NS	128	128	64	256	128	256/128/64[b]	512
td2	2048	2800	2800	1600	1600	1400	1400
aq2 (ms)	14	19.6	19.6	16	16	14	14
$\nu^{1}H$ (kHz)	83.3	-	-	62.5	-	62.5	83.3
$\nu^{13}C$ (kHz)	62.5	62.5	62.5	62.5	62.5	-	62.5
CP ($^{1}H/^{13}C$) (kHz)	62.5	-	-	62.5	-	51[c]	62.5
CP contact time (ms)	1	-	-	1	-	1.0	1.0
^{1}H decoupling (kHz)	83.3	83.3	83.3	62.5	83.3	62.5	83.3
Processing	GM (-7, 0.07)	GM (-7, 0.07)	GM (-7, 0.07)	GM (-7, 0.07)	GM (-7, 0.07)	GM (-5, 0.1)	GM (-5, 0.1)

[a]Recycle delays are 40s (for eucalyptus) and 35s (for the other two). [b]To protect the probe during longer LG-SL time (14 ms and 19 ms), smaller NSs are used (128/64). [c]For ^{1}H-$T_{1\rho}$, 51 kHz for CP matching means $w1_{eff}$ is 62.5 kHz during LG-SL and LG-CP.

Table S7.17. Parameters of 2D NMR experiments measured on the three wood samples. Recycle delay (d1); number of scans (NS); number of points of time domain for the direct (td2) and indirect (td1) dimensions; the acquisition time of the direct dimension (aq2); the evolution time of indirect dimension (aq1); excitation frequency for proton ($\nu^{1}H$) and carbon ($\nu^{13}C$) channels.

	CP-INADQ	DP-INADQ	RFDR	Dipolar-gated PDSD	2D Water-edited	DARR	DIPSHIFT	DP-PDSD	HSQC
Field (T)	14.1	14.1	9.4	14.1	9.4	9.4	9.4	9.4	11.8
Temp. (K)	294	294	298	294	278	278	298	298	-
MAS (kHz)	14	14	10	14	10	10	7.5	12	-
d1(s)	1.6	1.6	2.0	1.48	1.65	1.65	2.0	1.7	1
NS	32	16	128	160	128/256 [a]	128	128	192	224
td2	1400	1800	1600	1400	1400	1400	1600	1600	2048
td1	340	112	280	180	152	152	13	200	256
aq2 (ms)	14	18	16	14	14	14	16	16	12.78
aq1 (ms)	4.5	4.5	7.00	3.42	4.94	4.94	0.14	4.20	3.91
$\nu^{1}H$ (kHz)	83.3	-	71.4	83.3	83.3	83.3	83.3	-	31.3
$\nu^{13}C$ (kHz)	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	20.8
CP ($^{1}H/^{13}C$) (kHz)	62.5	-	62.5	62.5	62.5	62.5	62.5	-	-
CP contact time (ms)	1	-	1	1	1	1	1	-	-
^{1}H decoupling (kHz)	83.3	83.3	71.4	83.3	83.3	83.3	83.3	83.3	31.3
Mixing time (ms)	-	-	1.6	20/100/1000 [b]	50	50	-	100	0.86[c]

Table S7.17 (cont'd)

Processing	QSINE (ssb:2.8)	QSINE (ssb:3)	GM (-10, 0.05) QSINE (ssb:4)[e]	GM (-15, 0.04) QSINE (ssb:3)[e]	GM (-15, 0.04) QSINE (ssb:4.5)[e]	GM (-30, 0.03) QSINE (ssb:4.5)[e]	GM (-5, 0.1)	GM (-30, 0.03) QSINE (ssb:4)[e]	QSINC (ssb:3.5) QSINE (ssb:3.5)[d]

[a] 128 NS for control one and 256 NS for 4ms-SD water edited one. [b] 1000ms/100ms for long and short mixing time and spruce also has 20-ms mixing time spectrum. [c] $1/(8J_{CH})$. [d] Along direct and indirect dimensions respectively. For the later forward linear prediction is applied with 32 coefficients. [e] GM for lignin region and QSINE for carbohydrate region.

Table S7.18. Parameters of ssNMR experiments measured for never-dried spruce and eucalyptus samples. The same experiments with same parameters were measured again for both two wood samples after lyophilization and rehydration. Never-dried spruce was involved in all 1D and 2D experiments, and never-dried eucalyptus was only involved in 1D experiments. Recycle delay (d1); number of scans (NS); number of points of time domain for the direct (td2) and indirect (td1) dimensions; the acquisition time of the direct dimension (aq2); the evolution time of indirect dimension (aq1); excitation frequency for proton ($\nu^{1}H$) and carbon ($\nu^{13}C$) channels.

	CP	DP	Quantitative DP	^{13}C-T_1 Torchia	^{1}H-$T_{1\rho}$	1D Water-edited	2D Water-edited	CP-INADQ	DP-INADQ	53-ms CORD
Field (T)	9.4	9.4	9.4	9.4	9.4	9.4	9.4	9.4	9.4	9.4
Temp. (K)	298	298	298	298	298	280	280	298	298	298
MAS (kHz)	10	10	10	10	10	10	10	10	10	10
d1(s)	2	2	35	2	2	2	1.7	1.8	2	2
NS	1k	1k	512	512	512/256/128[a]	1k/512	96	192	192	128
td2	1600	1600	1600	1600	1600	1400	1600	1400	1600	1600
td1	1	1	1	1	1	1	166	90	90	166
aq2 (ms)	16	16	16	16	16	14	16	14	16	16
aq1 (ms)	-	-	-	-	-	-	4.15	5.31	5.31	4.15
$\nu^{1}H$ (kHz)	83.3	-	-	83.3	62.5	83.3	83.3	83.3	-	83.3
$\nu^{13}C$ (kHz)	62.5	62.5	62.5	62.5	-	62.5	62.5	62.5	62.5	62.5
CP match (^{1}H/^{13}C) (kHz)	62.5	-	-	62.5	51[b]	62.5	62.5	62.5	-	62.5
CP contact time (ms)	1.0	-	-	1.0	1.0	1.0	1.0	1.0	-	1.0

Table S7.18 (cont'd)

^{1}H decoupling (kHz)	83.3	83.3	83.3	83.3	62.5	83.3	83.3	83.3	83.3	83.3
Processing	GM (-5, 0.1)	GM (-10, 0.05)	GM (-10, 0.05)	GM (-5, 0.1)	GM (-5, 0.1)	GM (-5, 0.1)	GM (-20, 0.04) QSINE (ssb:4)[c]	QSINE (ssb:4)	QSINE (ssb:4)	GM (-30, 0.03)

[a] To protect the probe during longer LG-SL time (14 ms and 19 ms), smaller NSs are used (256/128). [b] For ^{1}H-$T_{1\rho}$, 51 kHz for CP matching means $w1_{eff}$ is 62.5 kHz during LG-SL and LG-CP. [c] GM for lignin region and QSINE for carbohydrate region.

Table S7.19. Comparison of ^{13}C chemical shifts with literature-reported values. The comparison results are shown for carbohydrates, lignin aromatic carbons as well as covalent linkers. The references are indicated for each comparison.

Carbohydrate		Eucalyptus	Poplar	Spruce	reference 1	reference 2	reference 3
					Wang et al. 2014[107]	*Kang et al. 2019*[108]	*Gao et al. 2020*[109]
	C1	104.9	104.8	104.8	105.0	105.8	105.2
	C2	72.1	72.1	72.2	72.9	71.5	72.8
i[a]	C3	75.2	75.0	75.3	75.8	75.8	75.1
	C4	88.7	88.9	89.0	89.2	89.1	88.1
	C5	72.3	72.4	72.4	72.9	72.5	72.8
	C6	65.0	64.9	65.0	65.2	64.9	64.8
					Wang et al. 2014	*Kang et al. 2019*	*Gao et al. 2020*
	C1	104.9	104.8	104.8	105.0	105.1	105.2
	C2	72.1	72.1	72.2	72.5	72.9	72.8
i[b]	C3	75.2	75.0	75.3	75.6	74.1	75.8
	C4	88.7	88.9	89.0	88.8	89.1	89.0
	C5	72.3	72.4	72.4	72.5	72.5	72.6
	C6	65.0	64.9	65.0	65.1	65.0	65.4
					Wang et al. 2014	*Kang et al. 2019*	*Gao et al. 2020*
	C1	104.9	104.8	104.8	105.0	104.2	105.2
	C2	72.1	72.1	72.2	72.5	71.1	72.8
i[c]	C3	74.1	74.3	74.3	75.4	-	75.8
	C4	87.9	87.9	88.1	88.8	87.9	89.5
	C5	71.1	70.1	71.0	72.5	-	73.0
	C6	65.7	65.5	65.6	65.1	65.8	65.3
					Wang et al. 2014	*Kang et al. 2019*	*Gao et al. 2020*
	C1	104.9	104.8	104.8	105.0	105.2	105.2
s[f]	C2	72.1	72.1	72.2	72.5	72.7	72.8
	C3	74.6	75.7	75.1	75.6	75.8	75.1
	C4	84.4	84.3	84.5	84.7	84.3	83.8

Table S7.19 (cont'd)

	C5	74.6	75.0	75.0	75.7	75.4	75.1
	C6	62.3	62.3	62.1	62.9	62.5	61.8
s^g					*Wang et al. 2014*	*Kang et al. 2019*	*Gao et al. 2020*
	C1	104.9	104.8	104.8	105.0	106.1	105.2
	C2	72.1	72.1	72.2	72.5	73.0	72.8
	C3	73.6	75.2	74.3	75.6	73.1	75.4
	C4	83.5	83.4	84.3	84.7	83.5	84.1
	C5	73.6	73.7	74.3	75.9	74.6	75.4
	C6	61.4	61.3	61.4	62.0	61.5	62.2
Xn^{2f}					*Simmons et al. 2016*[25]	*Kang et al. 2019*	*Gao et al. 2020*
	C1	104.8	104.8	104.8	105.2	105.1	105.1
	C2	7.2	72.1	72.2	72.3	72.5	72.3
	C3	74.0	73.2	73.3	75.2	73.5	75.3
	C4	82.2	82.0	82.2	82.2	82.3	82.1
	C5	64.0	63.9	63.9	64.3	64.6	64.1
	AC^{CO}	174.0	174.0	174.0	173.6	174.0	174.0
	AC^{Me}	21.0	21.0	21.0	21.6	21.0	21.4
Xn^{3f}					*Simmons et al. 2016*	*Kang et al. 2019*	*Gao et al. 2020*
	C1	102.5	102.2	103.5	102.6	102.5	102.5
	C2	73.4	73.6	73.9	73.7	73.5	73.7
	C3	74.1	74.2	74.2	74.7	74.5	74.7
	C4	77.4	77.9	77.8	77.4	78.0	77.3
	C5	63.3	64.1	65.4	63.9	62.5	63.8
	AC^{CO}	174.0	174.0	174.0	173.6	174.0	174.0
	AC^{Me}	21.0	21.0	21.0	21.6	21.0	21.4
Xn					*Kang et al. 2019*		

Table S7.19 (cont'd)

	C1	-	/		102.5		
	C2	-	/		73.5		
	C3	74.2	/		74.5		
	C4	78.5	/		78.0		
	C5	64.1	/		62.5		
	AC^{CO}				174.0		
	AC^{Me}						
GlcA[a]					*Duan et al. 2021* [111]		
	C1	101.5	101.5	101.5	99		
	C2	71.3	71.3	71.3	72		
	C3	79.8	79.8	79.8			
	C4	69.9	69.9	69.9			
	C5	76.6	76.6	76.6			
GlcA[b]					*Simmons et al. 2016*		
	C1	99.2	99.2	99.2	98.6		
	C2	69.9	69.9	69.9	72.4		
	C3	-	-	-			
	C4	-	-	-			
	C5	-	-	-			
GlcA[c]					*Wang et al. 2014*	*Simmons et al. 2016*	
	C1	98.5	98.5	98.5	98.6	98.6	
	C2	71.6	71.6	71.6	72.4	72.4	
	C3	80.8	80.8	80.8	78.3		
	C4	-	-	-	-		
	C5	-	-	-	72.6		
	C6	-	-	-	177.5		
GalA					*Wang et al. 2014*	*Terrett et al. 2019*[112]	*Phyo et al. 2017*[113]

Table S7.19 (cont'd)

	C1	100.2	100.2	100.2	100.2	101.1	99.9
	C2	68.1	68.1	68.1	69.3	69.8	69.0
	C3	-	-	-	-	-	69.8
	C4	79.6	79.6	79.6	-	-	78.8
	C5	-	-	-	-	-	71.3
	C6	-	-	-			171.6
					Wang et al. 2014	*Terrett et al. 2019*	
M	C1	/	/	101.2	100.2	101.9	
	C2	/	/	71.3	69.3	72.0	
	C3	/	/	75.1	-	-	
	C4	/	/	-	-	80.4	
	C5	/	/	75.1	-	75.8	
	C6	/	/	62.1	-	61.6	
					Terrett et al. 2019		
M^{Ac}	C1	/	/	101.6	100.9		
	C2	/	/	72.1	71.9		
	C3	/	/	75.1	75.9		
	C4	/	/	80.7	80.4		
	C5	/	/	75.1	75.8		
	C6	/	/	62.1	61.6		
					Wang et al. 2014	*Terrett et al. 2019*	*Duan et al. 2021*
A(t-Ara)	C1	/	/	108.1	110.0	108.6	109
	C2	/	/	82.1	82.2	82.1	82
	C3	/	/	-	77.7	78.3	
	C4	/	/	85.6	84.9	86.1	
	C5	/	/	62.1	62.3	62.8	
					Wang et al. 2014	*Phyo et al. 2017*	
A^{a}	C1	107.8	107.8	107.8	107.1	107.2	
	C2	81.8	81.8	81.8	89.8	87.8	

Table S7.19 (cont'd)

	C3	77.2	77.2	77.2	76.8	76.2	
	C4	84.6	84.6	84.6	78.6	82.4	
	C5	61.8	61.8	61.8	70.7	67.3	
A^b					*Wang et al. 2014*	*Phyo et al. 2017*	
	C1	107.8	107.8	107.8	107.3	108.4	
	C2	80.1	80.1	80.1	87.8	85.8	
	C3	-	-	-	72.6	81.1	
	C4	-	-	-	-	81.9	
	C5	-	-	-	-	67.0	
A^c					*Wang et al. 2014*	*Phyo et al. 2017*	
	C1	109.7	109.7	109.7	110.0	108.4	
	C2	81.8	81.8	81.8	82.2	81.8	
	C3	77.4	77.4	77.4	77.7	77.8	
	C4	82.9	82.9	82.9	84.9	83.1	
	C5	67.6	67.6	67.6	62.3	67.8	
A^d					*Wang et al. 2014*	*Phyo et al. 2017*	
	C1	106.9	/	/	109.2	107.7	
	C2	85.3	/	/	82.1	82.2	
	C3	-	/	/	78.2	77.5	
	C4	-	/	/	85.7	84.9	
	C5	-	/	/	62.4	62.1	
Lignin		**Eucalyptus**	**Poplar**	**Spruce**	**reference 1**	**reference 2**	**Other solution NMR studies**
G					*Kang et al. 2019*	*Ralph et al. 2009*	
	C1	-	/	/	131.6	130.5	

Table S7.19 (cont'd)

	C2	111.1	/	/	-	112.2	111.5 [96] 110.8 [97] 111.7 [114] 111.6 [89] 110.9 [99] 110.9 [100] 110.9 [101] 111.4 [102] 111.4 [103] 110.9 [88] 111.0 [100] 110.3 [104]
	C3	148.9	/	/	148.2	150.3	
	C4	144.9	/	/	148.2	150.1	
	C5	115.9	/	/	115.5	112.2	115.2 [96] 115.4 [114] 115.4 [89] 114.7 [99] 114.9 [100] 115.1 [101] 115.0 [102] 115.4 [103] 114.6 [88] 114.9 [105] 115.2 [104]

Table S7.19 (cont'd)

	C6	-	/	/	119.6	120.8	119.5 [96] 118.8 [97] 119.6 [114] 119.5 [89] 119.0 [99] 119.0 [100] 118.8 [101] 119.5 [102] 119.3 [103] 119.1 [88] 118.6 [105] 118.9 [104]
	OMe	57.0	/	/	56.9	56.1	
G^a					*Kang et al. 2019*		
	C1	/	-	130.7	132.1		
	C2	/	112.1	113.4	114.5		
	C3	/	48.2	152.2	146.2		
	C4	/	144.9	145.3	146.2		
	C5	/	116.1	116.3	114.5		
	C6	/	-	-	-		
	OMe	/	56.6	56.7	56.7		
G^b					*Kang et al. 2019*		
	C1	/	-	133.5	134.8		
	C2	/	113.8	110.5	111.5		
	C3	/	148.2	149.7	147.4		
	C4	/	144.9	146.1	147.4		
	C5	/	116.1	-	115.8		
	C6	/	-	120.2	120.4		
	OMe	/	56.6	57.0	57.0		
G^c					*Kang et al. 2019*		
	C1	/	/	134.2	134.8		
	C2	/	/	113.5	111.5		

Table S7.19 (cont'd)

	C3	/	/	147.6	147.4		
	C4	/	/	144.1	147.4		
	C5	/	/	-	115.8		
	C6	/	/	119.2	120.4		
	OMe	/	/	57.0	57.0		
G[d]					*Kang et al. 2019*		
	C1	/	/	136.0	134.8		
	C2	/	/	109.5	111.5		
	C3	/	/	151.6	147.4		
	C4	/	/	142.2	147.4		
	C5	/	/	-	115.8		
	C6	/	/	121.1	120.4		
	OMe	/	/	57.0	57.0		
S[a]					*Kang et al. 2019*		
	C1	134.9	134.2	/	136.2		
	C2	102.8	104.0	/	105.5		
	C3	153.3	153.1	/	152.7		
	C4	133.5	134.6	/	134.5		
	C5	153.3	153.1	/	152.7		
	C6	102.8	104.0	/	105.5		
	OMe	57.0	56.6	/	56.7		
S[b]					*Kang et al. 2019*		
	C1	-	137.1	/	136.2		
	C2	109.7	102.5	/	105.5		
	C3	153.2	152.7	/	152.7		
	C4	137.2	135.8	/	134.5		
	C5	153.2	152.7	/	152.7		
	C6	109.7	102.5	/	105.5		
	OMe	57.0	56.6	/	56.7		
S'[a]					*Rencoret et al. 2011*	*Yuan et al. 2011*	
	C1	134.1	/	/			

Table S7.19 (cont'd)

	C2	109.2	/	/	106.6	106.2	
	C3	144.8	/	/	-	-	
	C4	136.3	/	/	-	-	
	C5	144.8	/	/	-	-	
	C6	109.2	/	/	106.6	106.2	
	OMe	56.5	/	/	-	-	
					Rencoret et al. 2011	*Yuan et al. 2011*	
S'[b]	C1	136.5	/	/			
	C2	112.5	/	/	106.6	106.2	
	C3	145.5	/	/	-	-	
	C4	138.5	/	/	-	-	
	C5	145.5	/	/	-	-	
	C6	112.5	/	/	106.6	106.2	
	OMe	56.5	/	/	-	-	
					Chemical shifts from solution NMR spectra		
Lignin Linkages		**Eucalyptus**	**Poplar**	**Spruce**	**Eucalyptus**	**Poplar**	**Spruce**
A(G) / A(S)	C7(α)	72.3	72.0	72.4	71.6 / 72.5 [96] - / 71.8 [97] 71.4 / 72.6 [114] 71.7 / 72.4 [89] - / 71.8 [99]	71.5 / 71.9 [101] - / 72.0 [102] - / 71.8 [103] - / 71.8 [88]	71.1 [104] 71.0 [106]
	C8(β)	84.5	86.3	84.3	84.1 / 86.5 [96] 83.6 / 85.7 [97] 84.1 / 86.4; 87.3 [114] 84.1 / 86.4 [89] 83.6 / 85.8 [99]	83.9 / 85.9 [100] 83.4 / 85.8 [101] - / 85.8 [103] 83.8 / 85.8 [88]	84.0 [104] 83.9 [106]
	C9(γ)	59.4	60.4	59.9	60.0 [96] 59.4 [97] 60.1 [114] 59.6 [99]	59.6 [100] 59.8 [101] 60.2 [103]	59.9 [104] 59.8 [106]

Table S7.19 (cont'd)

B	C7(α)	84.4	84.8	-	85.4 [96] 84.7 [97] 85.4 [114] 85.5 [89] 84.8 [14]	84.8 [100] 85.2 [101] 85.7 [102] 84.7 [103] 84.8 [88]	84.9 [104] 85.0 [106]
	C8(β)	53.5	52.6	53.9	53.7 [96] 53.3 [97] 54.1 [114] 53.5 [99]	53.5 [100] 53.6 [101] 53.6 [103]	53.5 [104] 53.6 [106]
	C9(γ)	71.2	70.5	70.3	71.7 [96] 71.0 [97] 71.7 [114] 71.3 [99]	71.0 [100] 70.1 [101]	70.8 [104] 70.9 [106]
B'	C7(α)	-	83.9	/			
	C8(β)	53.5	52.9	/			
	C9(γ)	74.4	74.1	/			
E	C7(α)	81.1	80.3	81.4	81.7 [96] 81.0 [97]	81.2 [100]	81.7 [106]
	C8(β)	59.6	59.7	60.0	60.3 [96] 59.7 [97]	59.7 [100]	59.3 [106]
	C9(γ)	/	/	/			
V	C7(α)	75.8	77.2	76.2			75.8 [115]
	C8(β)	78.4	79.0	78.9			78.0 [115]
	C9(γ)	59.8	60.7	60.1			
C	C7(α)	/	/	88.1	87.7 [96] 86.4 [97] 87.5 [114] 87.4 [89] 86.7 [99]	86.8 [100] 87.0 [101] 87.4 [102] 87.1 [103] 86.8 [88]	87.0 [104] 87.0 [106]

Table S7.19 (cont'd)

	C8(β)	/	/	54.2	53.7 [96] 53.1 [97] 53.7 [114] 53.5 [90]	53.3[100] 53.2 [101] 53.2 [102]	53.1 [104] 53.0 [106]
	C9(γ)	/	/	/	63.3 [96] 62.3 [97] 62.5 [99]	62.5[100] 62.5 [101]	62.8 [104] 62.7 [106]
C'	C7(α)	/	/	90.2			
	C8(β)	/	/	49.9			
	C9(γ)	/	/	/			

CHAPTER 8: SOLID-STATE NMR OF UNLABELED PLANT CELL WALLS: HIGH-RESOLUTION STRUCTURAL ANALYSIS WITHOUT ISOTOPIC ENRICHMENT

Research paper reprinted with permission: *Biotechnol. Biofuels* 14:14, (2021)

Authors: Wancheng Zhao, Alex Kirui, Fabien Deligey, Frederic Mentink-Vigier, Yihua Zhou, Baocai Zhang, and Tuo Wang

8.1 Abstract

Multidimensional solid-state nuclear magnetic resonance (ssNMR) spectroscopy has emerged as an indispensable technique for resolving polymer structure and intermolecular packing in primary and secondary plant cell walls. Isotope (^{13}C)-enrichment provides feasible sensitivity for measuring 2D/3D correlation spectra, but this time-consuming procedure and its associated expenses have restricted the application of ssNMR in lignocellulose analysis. Here, we present a method that relies on the sensitivity-enhancing technique Dynamic Nuclear Polarization (DNP) to eliminate the need for ^{13}C-labeling. With a 26-fold sensitivity enhancement, a series of 2D ^{13}C-^{13}C correlation spectra were successfully collected using the unlabeled stems of wild-type *Oryza sativa* (rice). The atomic resolution allows us to observe a large number of intramolecular cross peaks for fully revealing the polymorphic structure of cellulose and xylan. NMR relaxation and dipolar order parameters further suggest a sophisticated change of molecular motions in a *ctl1 ctl2* double mutant: both cellulose and xylan have become more dynamic on the nanosecond and microsecond timescale, but the motional amplitudes are uniformly small for both polysaccharides. By skipping isotopic labeling, the DNP strategy demonstrated here is universally extendable to all lignocellulose materials. This time-efficient method has landed the technical foundation for understanding polysaccharide structure and cell wall assembly in a large variety of plant tissues and species.

8.2 Introduction

The past decade has witnessed the rapid advances in multidimensional solid-state NMR (ssNMR) capabilities that have enabled high-resolution characterization of intact plant cell walls. This spectroscopic method provides a wealth of atomic-level information on the conformational structure of polysaccharides, covalent linkage patterns of matrix polysaccharides, dynamical profile and water contact, as well as cellulose-matrix packing on the subnanometer scale[1]. With a rapidly expanding territory, from eudicotyledons (*Arabidopsis thaliana*) to commelinid monocotyledons (*Zea mays*, *Brachypodium distachyon*, etc.)[2-5], from primary to secondary cell walls[6-10], and from plants to algal and fungal species[11-13], ssNMR has progressively evolved into a vital tool for characterizing carbohydrate-rich biomaterials. The molecular information of cell wall architecture can serve as the structural basis for improving the current technologies of biofuel production using lignocellulosic biomass[14].

Enriching the cell walls with NMR-active isotopes (such as ^{13}C and ^{15}N) is a prerequisite for measuring two- and three-dimensional (2D/3D) correlation experiments, which provides the spectral resolution required for resolving numerous carbon and nitrogen sites in cell wall polymers. Two strategies can be employed: plants can be grown in the dark, using a medium containing ^{13}C-labeled glucose[2]; otherwise, $^{13}CO_2$ can be continuously supplied to the plants grown in a day-night cycle[4, 15]. Depending on the developmental stage and the tissue of interest, labeling can be time-consuming and costly. In-vitro replication procedures also weaken the merit of ssNMR as an analytical technique targeting natural tissues.

The recent development of Magic-Angle Spinning Dynamic Nuclear Polarization (MAS-DNP) methods has presented a unique opportunity for circumventing these drawbacks[16-19]. MAS-DNP enhances NMR sensitivity by tens to hundreds of folds, which allows us to skip isotope-

enrichment and use unlabeled samples to measure 2D ^{13}C-^{13}C/^{15}N correlation spectra for high-resolution structural characterization[20-22]. Regarding the plant biomass, three exploratory studies have been conducted to reveal the restructuring effect of ball-milling on cotton cellulose[23], the alternation of xylan conformations induced by genetic mutations of rice[24], and the compositional changes of lignin in high-S and low-S poplar[25]. With the rapid development of DNP instrumentation[26-29] and radical design[30-33], this is certainly a direction of great potential but not yet explicitly explored for plant materials.

This methodology article aims at establishing a universally applicable toolbox for characterizing polymer structure and assembly in unlabeled plant biomass. This is achieved by combining a series of DNP-enabled experiments that probe the composition and conformational structure of polymers with conventional ssNMR measurements that examine the rate and amplitudes of molecular motions. Implementation of this method will expand the ssNMR capabilities and enable high-resolution investigations of unlabeled cell walls, which, at least in part, provides a replacement and upgrade to the conventional methods that rely on ^{13}C-enriched materials. Most of the structural aspects previously investigated using ^{13}C-enriched samples, such as the composition, conformation, packing, and motion of cell wall polymers, can be studied using unlabeled materials via a blend of ssNMR and MAS-DNP methods (**Table 8.1**). This technical advance will eliminate the threshold that has long been impeding lignocellulose characterization, which will immediately benefit the research communities of plant biology, biomaterials, and bioenergy.

Table 8.1 Capabilities of solid-state NMR and DNP. Technical aspects are categorized as being fully capable (Yes), partially limited by insufficient resolution or sensitivity (In part), and infeasible or unsuitable (No).

Samples and Techniques	Polymer Composition	Structural Polymorph	Intermolecular Packing	Molecular Motion	Water Contact
^{13}C-labeled; ssNMR	Yes	Yes	Yes	Yes	Yes
Unlabeled; ssNMR	In part	No	No	In part	In part
Unlabeled; MAS-DNP	Yes	Yes	In part	No[a]	No[a]

[a]MAS-DNP is conducted at cryogenic temperature (~100 K); therefore, it is unsuitable for investigating molecular motions. It is also better to investigate polymer-water contacts at room temperature.

8.3 Results

8.3.1 Sensitivity Enhancement by MAS-DNP

For decades, one-dimensional (1D) ^{13}C spectra have been conducted using unlabeled materials to analyze the structure and composition of polysaccharides at natural isotope abundance (1.1% for ^{13}C)[34-37]. **Figure 8.1** shows the 1D quantitative spectra of four rice samples including the wild-type (WT) stems, two single mutants, which are *ctl1* (harboring an identical mutation of *brittle culm 15*) and *ctl2*, as well as a double mutant *ctl1 ctl2*. These mutations happen to rice *CHITINASE-LIKE1* (*CTL1*) and its homolog *CTL2*, which have been suggested a role in controlling cellulose biosynthesis thereby affecting wall integrity[38-41], as *ctl1* and *ctl1 ctl2* plants

exhibited obvious brittleness phenotypes. Quantitative detection of all carbons is achieved using a recently developed pulse sequence, MultiCP, which counts on ^{1}H T_1 relaxation to repeatedly restore ^{1}H magnetization between the many CP blocks included in each individual scan of the experiment[42-44]. The rice stems are collected from the field, without any dissolution procedures or chemical digestions; therefore, the cell walls being analyzed are still native and intact. Various peaks of cellulose (interior chains: i; surface chains: s) and xylan (Xn) have been observed (**Figure 8.1a**), indicating the predominance of secondary cell wall components in these mature stems.

Different from the WT stems, the double mutant has a higher amount of methyl ether (-OMe) groups, which is a chemical substitution frequently occurring in lignin. The *ctl1 ctl2* sample shows a higher peak at 56 ppm for methyl ether groups **(Figure 8.1a**) but comparable intensities for aromatic carbons (**Figure S8.1**). This observation indicates a higher fraction of guaiacyl (G) and syringyl (S) units that are heavily substituted by methyl ether groups. The *ctl1 ctl2* sample has shown stronger xylan peaks, for example, the carbon 1 of 3-fold xylan at 102 ppm ($Xn1^{3f}$), the carbon 4 of two-fold xylan at 82 ppm ($Xn4^{2f}$), and the methyl carbon of acetyl group at 21 ppm (Ac^{Me}). Subtraction of the *WT* and *ctl1 ctl2* spectra has generated a difference spectrum, which contains well-resolved signals from 2-fold and 3-fold xylan, revealing a higher amount of xylan in the double mutant. The spectral patterns of the two single mutants are intriguing: the spectrum of *ctl1* resembles that of the double mutant while *ctl2* is similar to the WT sample (**Figure 8.1b**). Combined with the cell wall defects detected in *bc15*[41], the results have suggested a stronger role of the *ctl1* mutation in modulating cell wall composition and structure. While useful information can be obtained by following this conventional 1D analysis, only a few carbon sites can be resolved. This resolution issue can be partially alleviated by spectral subtraction partially, but significant improvement is still needed for characterizing these polysaccharides and cell walls.

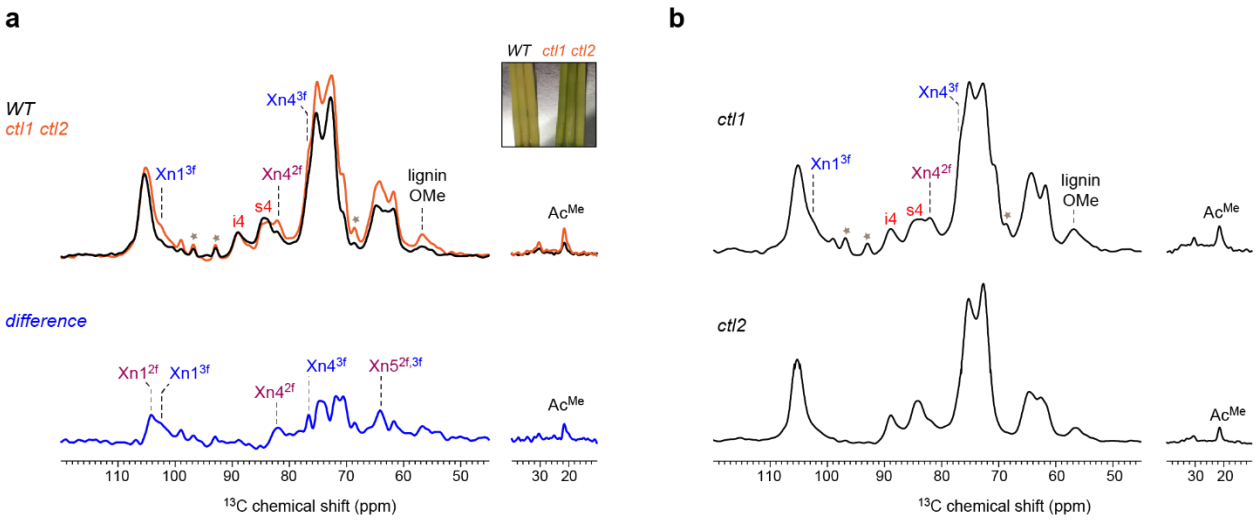

Figure 8.1 1D ^{13}C spectra of unlabeled cell walls of wild-type rice and mutants. **a**, 1D ^{13}C MultiCP spectra of wild-type sample (WT; black) and *ctl1 ctl2* double mutant (orange) yielding quantitative detection of all molecules. Abbreviations of carbohydrate assignments are shown in Figure 8.3, except for xylan acetyl group (Ac). The difference spectrum is obtained by subtraction of the two parent spectra after normalization by i4 peaks. The difference spectrum mainly contains xylan signals, revealing the higher hemicellulose content in *ctl1 ctl2* mutant. The insert picture on the right top side is the representative photo of the two kinds of rice stems. **b**, 1D ^{13}C MultiCP spectra of *ctl1* and *ctl2* single mutants. The spectral pattern of *ctl1* mutant is similar to that of the *ctl1 ctl2* double mutant while *ctl2* is closer to the WT sample.

The rice material is then impregnated in a solvent mixture of ^{13}C-depleted, d_8-glycerol/D_2O/H_2O, which contains 10 mM of biradical AMUPol[30]. ^{13}C-depletion of glycerol is a necessity for investigating unlabeled samples where both biomolecules and glycerol are at natural ^{13}C abundance: it can effectively eliminate the glycerol signals that overlap with the carbohydrate signals. This protocol, when compared with a matrix-free method that was recently applied to cotton[23, 45, 46], better retains polymer hydration in cell walls. The payoff is a moderate decline in the signal-to-noise ratios because the solvent occupies some volume in the MAS rotor, which will

otherwise be filled with more biomolecules.

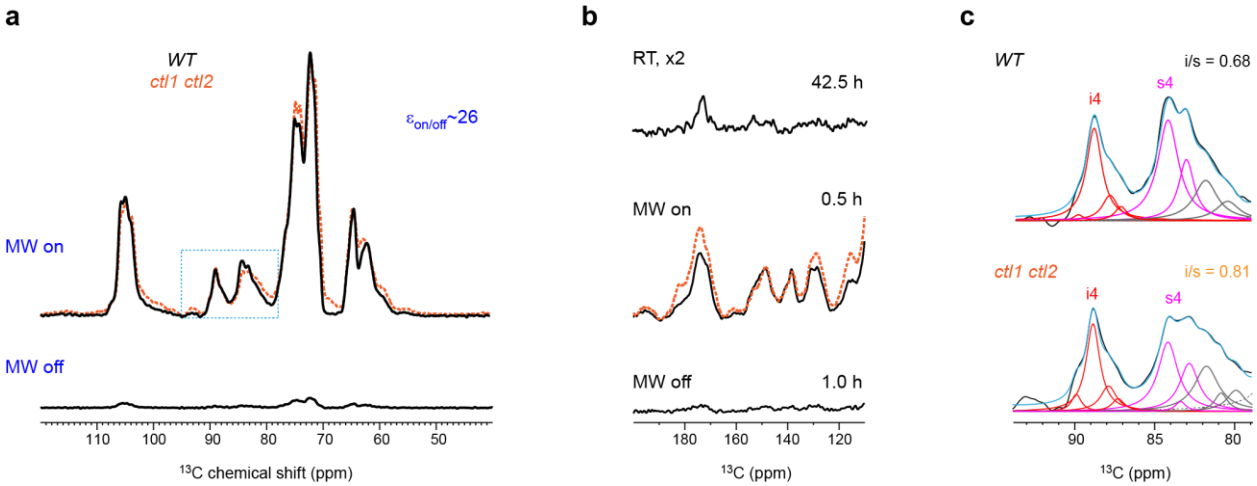

Figure 8.2 Sensitivity enhancement from DNP. **a**, DNP enhances the NMR sensitivity by 26-fold on the wild-type sample. This value is obtained by quantifying the intensity ratio between the spectra collected with microwave on (MW on) and microwave off (MW off). The DNP-enhanced spectrum of *ctl1 ctl2* (orange dashline) is compared with the wild-type (WT) sample. **b**, Lignin region of the room-temperature spectrum (RT; top), DNP-enhanced spectrum (MW on; middle), and non-DNP spectrum (MW off; bottom). The experimental time is labeled for each panel. The "×2" symbol means the intensities of the room-temperature spectrum is manually increased by two folds. DNP significantly saves experimental time and enhances NMR sensitivity. **c**, Spectral deconvolution of the interior cellulose carbon 4 (i4) and surface cellulose carbon 4 (s4) region (dashline square in panel **a**) reveals the crystallinity change in rice cellulose. The experimentally measured spectra (black) is overlaid with the simulated spectra (blue), with the deconvoluted peaks underlying the spectral envelope. The interior-to-surface (i/s) ratio of cellulose is quantified for both samples, which has shown an increase in cellulose crystallinity in the *ctl1 ctl2* mutant. The fit parameters are listed in Tables S8.1 and S8.2.

Under microwave irradiation, polarization of the electrons in the biradical will thereafter

be transferred to ^{1}H spins in unlabeled rice stem, and then to the natural abundance ^{13}C for detection. The sensitivity enhancement ($\varepsilon_{on/off}$) is 26-fold for the wild-type sample and 22-fold for the *ctl1 ctl2* double mutant (**Figure 8.2a, Figure S8.2a**), which respectively shortens the experimental time by factors 676 and 484 times. Unlike the optimal performance of ^{13}C-depleted, d_8-glycerol/D_2O/H_2O mixture, a sample prepared using a different solvent of d_6-DMSO/D_2O/H_2O only shows an enhancement of 5-fold, thus being excluded from the protocol (**Figure S8.2b**). The same gain of sensitivity also occurs to lignin. We have observed strong signals of aromatic carbons (120-160 ppm) within a short time of 0.5 h (**Figure 8.2b**). These signals, however, are invisible in conventional ssNMR even after 42.5 h of measurements at room temperature.

The DNP experiment preferentially detects the molecules with a highly ordered structure, for example, the cellulose microfibrils in plants (**Figure S8.3**) and the highly microcrystalline chitin in fungi[10, 12]. At the cryogenic temperature (~100 K) of MAS-DNP, the conformational distribution of mobile molecules, such as xylan and small molecules, will be fully trapped, resulting in broader lines and lower intensity. Recently, concerns have also been raised regarding the radical distribution in cell walls, which were supposed to induce a biased detection of molecules in spatial proximity to the radicals. This argument is invalid because all molecules within tens of nanometer to a radical can be efficiently polarized via ^{1}H-^{1}H spin diffusion, which is due to the inherently long relaxation times of the ^{1}Hs in cell walls[47]. This mechanism ensures a homogenous polarization of all molecules throughout the cell wall, which has been confirmed by the consistent spectral envelopes with and without microwave irradiation (**Figure S8.4**).

In addition, the crystallinity of cellulose becomes lower in the double mutant. Cellulose crystallinity is evaluated using the intensity ratio between the carbon 4 of interior glucan chains (i4; 89 ppm) and the carbon 4 of surface glucan chains (s4; 84 ppm) in cellulose. When normalized

by the i4 signal, the s4 peak is lower in the double mutant than in the WT sample as consistently observed at room temperature (**Figure 8.1a**) and under the DNP condition (**Figure 8.2a**). Spectral deconvolution allows us to resolve the underlying resonances and quantify the interior-to-surface (i/s) ratios, which are 0.68 for WT and 0.81 for c*tl1 ctl2* (**Figure 8.2c, Figure S8.5, Tables S8.1 and S8.2**). The higher cellulose crystallinity in the double mutant may originate from an increased degree of cellulose bundling in the cell wall. The observation here actually counters previous X-ray analysis of *Arabidopsis*, which has shown reduced content of crystalline cellulose in a *ctl1 ctl2* mutant[40]. The discrepancy can be attributed to the different plant species or the distinct structural aspects being examined in diffraction and spectroscopic techniques for determining the ambiguously defined "cellulose crystallinity"[48]. Our unpublished X-ray diffraction data on the same rice stems have observed a higher relative crystallinity index (RCI) for cellulose in the double mutant., which has confirmed the NMR observation.

8.3.2 2D ^{13}C-^{13}C Correlation Spectra of Unlabeled Cell Walls

Benefiting from the substantial gain of NMR sensitivity, we have successfully collected a 2D ^{13}C-^{13}C J-INADEQUATE spectrum[49, 50] using unlabeled stems of WT rice (**Figure 8.3a**). Because of the low abundance (1.1%) of ^{13}C in nature and the even lower probability (0.01%) for observing a cross peak between two carbons, such a 2D experiment has long been deemed impossible, but can now be finished now within 35 h using the rice stems.

We have observed eight types of glucose units in cellulose: types a-e for the glucan chains embedded in the core of microfibrils and types f-h for those exposed on the surface. Their ^{13}C chemical shifts are consistent with those observed in *Arabidopsis, Zea mays,* and *Brachypodium distachyon*, revealing the polymorphic nature of cellulose structure[51]. Hemicellulose shows weak signals, for example, the carbon 4 and carbon 5 in three-fold xylan (Xn4^{3f} and Xn5^{3f}) as well as

the carbon 3 and carbon 4 in two-fold conformers ($Xn3^{2f}$ and $Xn4^{2f}$). In addition, the α-xylose (x) of xyloglucan also exhibits some weak signals, indicative of a small portion of primary cell walls. Besides polysaccharides, we have also observed some self-correlation cross peak of the CH_2 acyl chains in lipid polymers. Therefore, the remarkable resolution of the DNP-enabled 2D spectrum allows us to unambiguously resolve many carbon sites in cellulose, matrix polysaccharides, and lipids, which is otherwise impossible.

The validity of natural-abundance MAS-DNP is confirmed by the consistent spectral pattern between unlabeled (**Figure 8.3a**) and ^{13}C-labeled (**Figure 8.3b**) rice stems. In addition, a few peaks are omitted in the natural-abundance DNP spectrum. These signals are primarily from the arabinose residues of xylan sidechains, some carbon sites in 2-fold, 3-fold and mixed forms of xylan backbones, as well as the acetyl groups (**Figure 8.3c**). Peak counting has confirmed MAS-DNP's preferential detection of ordered molecules: we can effectively detect all cellulose carbons and part of xylan backbones, but not for arabinose sidechains (**Figure 8.3d**). These mobile molecules, when trapped at 100 K, bear a wide distribution of conformations, which has broadened out their signals.

While the J-INADEQUATE experiment only probes through-bond correlations, we have also performed a CHHC experiment to detect long-range and through-space correlations (**Figure 8.4a**). The CHHC sequence relies on ^{1}H-^{1}H spin diffusion to transfer polarization, which is followed by cross-polarization (CP) from ^{1}H to ^{13}C for site-specific detection[52]. The Full Width at Half Maximum (FWHM) linewidths are around 3 ppm and the representative signal-to-noise ratios are ranging from 4 to 41 (**Figure S8.6**). The off-diagonal cross peaks are mostly intramolecular correlations within cellulose, for example, the i6-3 cross peak at (65, 75 ppm) and the s3-6 cross peak at (75, 62 ppm) in the 1-ms CHHC spectrum. Elongating the mixing time to 2 ms allows us

to extend the reach; therefore, many additional cross peaks now become observable (**Figure 8.2b**). The new signals are primarily from cellulose, such as the i6-1 cross peak at (65, 105 ppm) and i2,5-4 cross peak at (72, 89 ppm). We have also observed some signature signals of xylan, for example, Xn^{2f}2-4 at (73, 82 ppm) and Xn^{2f}3-4 at (78, 82 ppm).

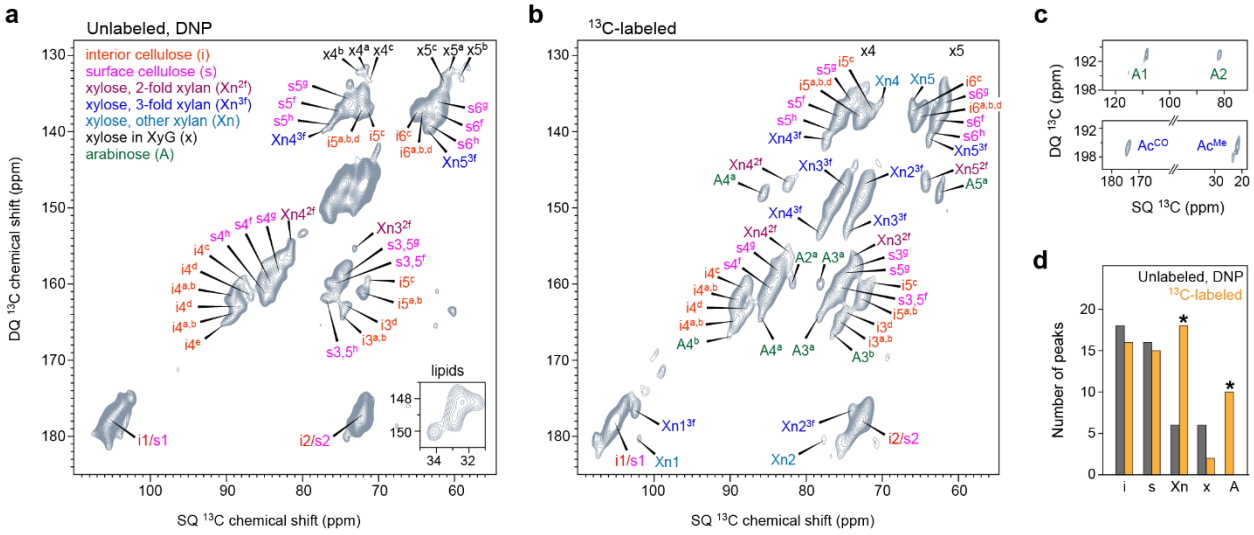

Figure 8.3 DNP enables 2D correlation spectroscopy using unlabeled and native cell walls. **a,** 2D ^{13}C–^{13}C INADEQUATE spectrum of unlabeled WT rice collected on a 600 MHz/395 GHz DNP system. DNP effectively detects and resolves the signals from cellulose and hemicellulose at natural isotope abundance. Superscripts are used to denote the eight types of glucose units in interior and surface cellulose and the 2-fold and 3-fold conformations in xylan. Inset shows the lipid signals that are folded in the indirect dimension. **b**, 2D ^{13}C–^{13}C INADEQUATE spectrum of ^{13}C labeled WT rice stems measured on a 400 MHz NMR spectrometer. The spectral pattern is comparable to the DNP-assisted natural-abundance spectrum in panel a. **c**, Arabinose and acetyl signals are detected using ^{13}C-labeled samples but become invisible in the DNP spectrum of unlabeled materials. **d**, Number of peaks detected using labeled and unlabeled samples, which are tabulated in Table S8.3. Asterisks indicate the components that are poorly detected using DNP.

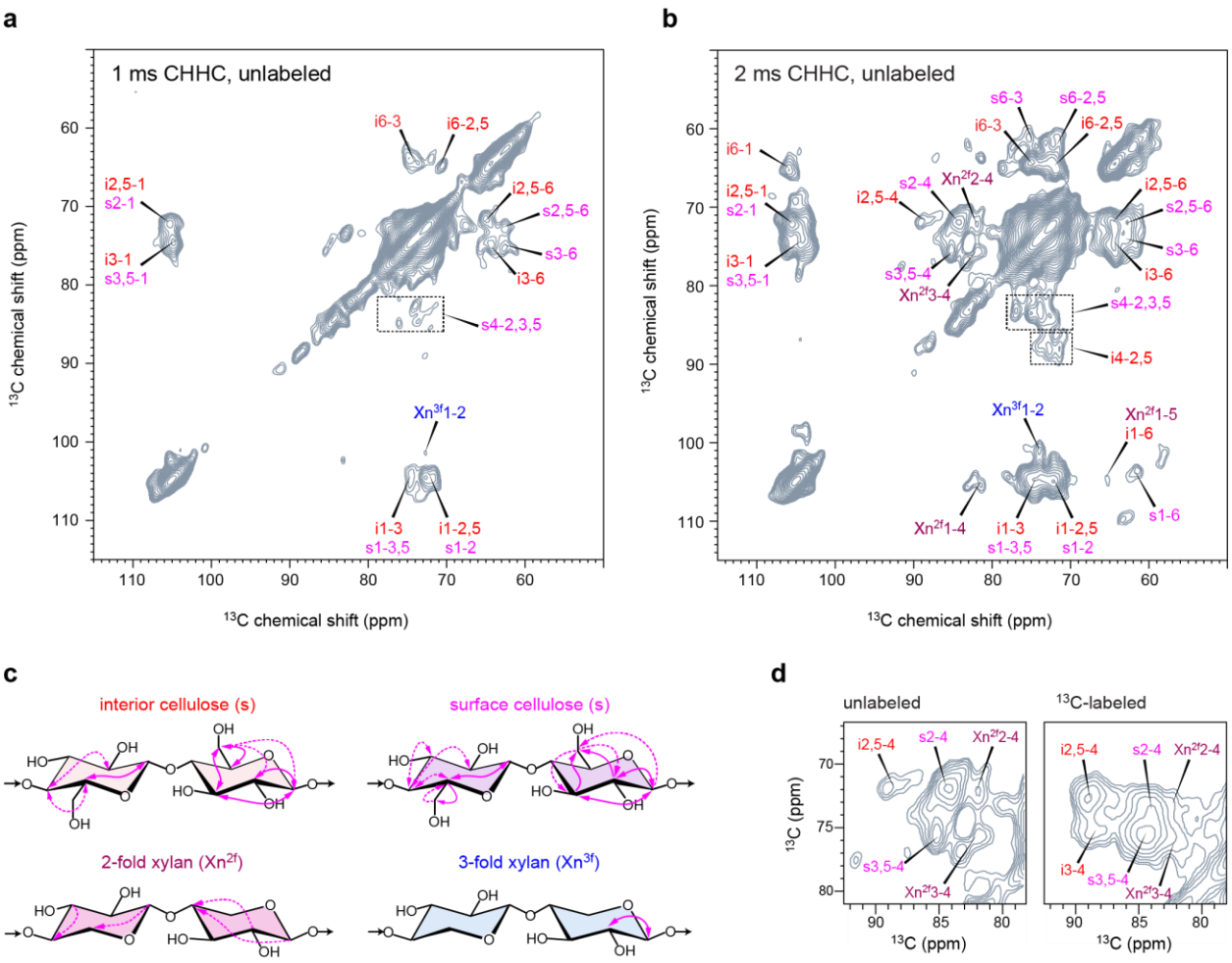

Figure 8.4 DNP-enabled 2D CHHC spectra of unlabeled rice stems. The CHHC spectra of unlabeled WT sample are measured with **a**, 1 ms and **b**, 2 ms ^{1}H-^{1}H mixing times. **c**, Summary of through-space intramolecular cross peaks. Solid lines denote the cross peaks observable in both 1 and 2 ms spectra while dashlines indicate the cross peaks that are observed only in the 2-ms spectrum. Arrows are used to indicate whether the cross peaks are bidirectional (e.g. C1-to-C2 and C2-to-C1 cross peaks) or unidirectional (e.g. only from C1 to C2). **d**, Comparison of polysaccharide signals in the 2-ms CHHC spectrum of the unlabeled sample and a 5-ms PDSD spectrum of ^{13}C-labeled rice.

All the inter-carbon correlations are summarized in **Figure 8.4c**. For cellulose, the large

number of cross peaks allows us to assign the ^{13}C chemical shifts of all carbons in interior and surface glucan chains. The flat-ribbon 2-fold xylan has shown 4 distinguishable signals while the non-flat 3-fold xylan only exhibits a single cross peak from carbon 1 to carbon 2 (Xn^{3f}1-2). The spectral pattern of CHHC resembles those PDSD spectra that measured using ^{13}C-labeled plant biomass (**Figure 8.4d, Figure S8.7**). Therefore, natural-abundance MAS-DNP has the full capability of investigating cellulose structure; it is also partially prepared for investigating matrix polysaccharides.

It should be noted that the CHHC experiment is specifically chosen for natural-abundance MAS-DNP. For most other 2D correlation methods, such as PDSD and DARR, the diagonal peaks will dominate the spectrum of unlabeled materials, because the probability of observing off-diagonal internuclear correlation is very low at natural ^{13}C abundance. The CHHC experiment can effectively avoid this issue due to the ^{13}C-^{1}H-^{1}H-^{13}C pathway used for polarization transfer[53]. Alternatively, dipolar homonuclear recoupling schemes could also be used for characterizing unlabeled biomaterials[22, 54].

8.3.3 Motional Rates and Amplitudes of Cell Wall Polysaccharides

To understand the effect of *ctl1 ctl2* mutation on polymer dynamics, we have measured ^{13}C-T_1 and ^{1}H-$T_{1\rho}$ relaxation times using the unlabeled stems of *WT* and *ctl1 ctl2* samples. As the most rigid component of cell walls, cellulose exhibits very slow ^{13}C-T_1 relaxation (typically on the order of 10-40 s), revealing its lack of motion on the nanosecond timescale (**Figure 8.5a, b**, **Table S8.4** and **Figure S8.8**). In contrast, the ^{13}C-T_1 of 3-fold xylan, the conformer that fills the interfibrillar space, is shorter by 4-5 times: 7.7 s for carbon 1 and 10 s for carbon 4 in the WT sample. The 2-fold xylan, a unique form induced by its packing with cellulose surface, has shown an intermediate ^{13}C-T_1 of 10.5 s for its resolved peak of carbon 4. These ^{13}C-T_1 time constants are

substantially longer than the values reported for uniformly ^{13}C-labeled materials (4-6 s for cellulose and 1-2 s for matrix polysaccharides)[55]. This can be attributed to the lack of spin exchange between rigid and mobile motifs in unlabeled materials.

The ^{1}H-$T_{1\rho}$ data is highly consistent with the ^{13}C-T_1 results. Cellulose remains as the most rigid component on the slower, microsecond timescale, with a representative ^{1}H-$T_{1\rho}$ time of 20-50 ms in WT stems (**Figure 8.5c, d**). The 2-fold and 3-fold conformers of xylan have more pronounced dynamics, showing ^{1}H-$T_{1\rho}$ values of 18 ms and 8-16 ms, respectively. Interestingly, most polysaccharides have considerably faster ^{13}C-T_1 and ^{1}H-$T_{1\rho}$ relaxations in the *ctl1 ctl2* mutant, indicative of a flawed cell wall formed by polysaccharides that are highly mobile on both nanosecond and millisecond timescales.

To examine the motional amplitudes of these polysaccharides, we have conducted the ^{13}C-^{1}H dipolar-chemical shift (DIPSHIFT) correlation experiment (**Figure 8.6**)[56]. The ^{13}C-^{1}H dipolar couplings (in kHz) are measured at representative carbon sites of cellulose and xylan, which are further converted to dipolar order parameters. A near-unity value indicates restricted motional amplitudes for the C-H bonds while a small order parameter suggests large-amplitude motions. Strikingly, all polysaccharides are highly rigid and exhibit large order parameters above 0.90 (**Figure 8.6b**). This is very different from the results collected using *Arabidopsis* and Brachypodium primary cell walls, in which the order parameters ranges from 0.3 to 0.6 for matrix polymers[5, 55]. Cellulose has identical order parameters in WT and *ctl1 ctl2* samples. The hemicellulose, however, has shown subtle changes: the order parameter has slightly decreased for the carbon 1 of 3-fold xylan ($Xn1^{3f}$) but increased for the carbon 4 of 2-fold xylan ($Xn4^{2f}$). These two peaks are the best-resolved signals of xylan, and the discrepancy observed here suggests a small extent of phase separation in the double mutant: the cellulose-packed 2-fold xylan becomes

even more rigid in the mutant while the interfibrillar 3-fold conformer undergoes even larger amplitude motions. The relaxation and DIPSHIFT experiments can systematically assess the motional characteristics of polysaccharides, which complement the 2D MAS-DNP data to provide a complete view of polymer composition, structure, and dynamics in unlabeled biomass.

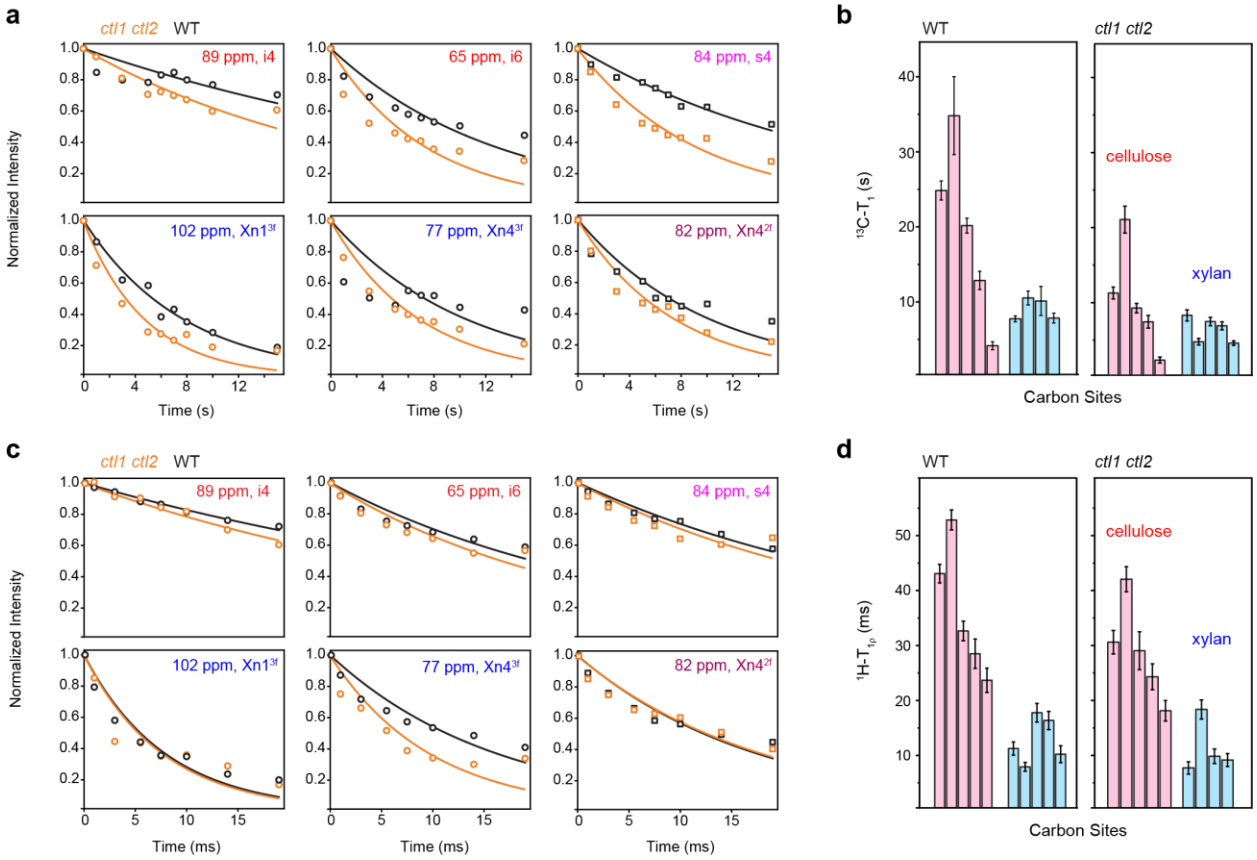

Figure 8.5 Carbohydrate dynamics probed by NMR relaxation using unlabeled rice stems. **a**, Representative ^{13}C-T_1 relaxation curves of cellulose (top row) and xylan (second row). The *ctl1 ctl2* double mutant has shown substantially faster relaxation than the wild-type sample, indicating more pronounced motions on the nanosecond timescale. **b**, Bar diagram summarizing the ^{13}C-T_1 relaxation times of cellulose (pink) and xylan (blue). The x-axis represents different carbon sites, which can be found in **Table S8.4**. **c,** Representative ^{1}H-$T_{1\rho}$ relaxation curves. For most peaks, the double mutant shows faster ^{1}H-$T_{1\rho}$ relaxation than the wild-type sample, indicating enhanced

Figure 8.5 (cont'd)

dynamics on the microsecond timescale. All the ^{1}H-$T_{1\rho}$ relaxation time constants are summarized in panel **d**.

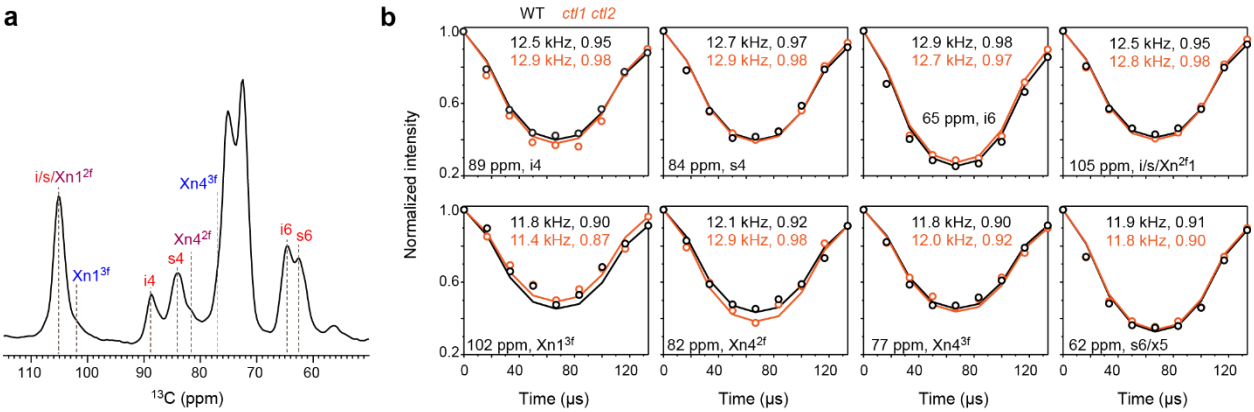

Figure 8.6 Dipolar order parameters of polysaccharides in unlabeled rice cell walls. **a**, The first slice of the 2D DIPSHIFT experiment measured under 7.5 kHz MAS, which shows the signals of cellulose and xylan. **b**, $^{13}C-^{1}H$ dipolar evolution curves extracted from 2D DIPSHIFT experiment. The fit dipolar coupling values (in kHz) and the calculated dipolar order parameters are labeled for each panel. Most carbon sites, except for some xylan peaks, have shown comparable order parameters in the wild-type and double mutant samples.

8.4 Discussion

Using a common agriculture lignocellulosic biomass, rice stems, we have demonstrated the feasibility of employing ssNMR and MAS-DNP methods to investigate the structure and dynamics of cell wall polysaccharides without ^{13}C enrichment. The strategy and the experiments described above can be directly applied to many other biomass samples. The method is time-efficient: the total experimental time is 254 h for the wild-type stems, which includes 132 h of measurements on a 600 MHz/395 GHz DNP instrument and 122 h on a conventional 400 MHz ssNMR spectrometer.

Multiple factors can impact the experimental time, for example, the mixing protocol that determines biradical distribution in cell wall materials, the sample pH that affects the lifetime and stability of radicals, and the choice of DNP solvent. As the beginning step, it is crucial to collect a set of 1D spectra to screen several samples prepared using different protocols and conditions, each of which only takes a few minutes to measure. By monitoring the absolute sensitivity and the DNP gain, the optimal condition will be identified for measuring 2D experiments.

At present, natural-abundance DNP still has major limitations. First, it is impossible to derive a detailed pattern of intermolecular contacts using only unlabeled materials (**Table S8.1**). This is due to the demanding requirement of sensitivity for probing longer-range correlations, which can even be challenging for some ^{13}C-labeled samples. Progress is being made: long-range correlations, typically on the range of 3-5Å, have been observed recently using model cyclic peptides[22]. Once successfully extended to cellular samples, these techniques will fully enable structural determination. Partial information could be obtained if the polymer contact depends on the conformational structure. For example, the flat-ribbon structure of 2-fold xylan is responsible for coating cellulose surface; therefore, we have tracked its amount to estimate the extent of xylan-cellulose packing in unlabeled rice stems[24]. Second, molecules with a high disorder suffer from intensity loss under MAS-DNP conditions. Fortunately, we still have distinguishable carbon sites for different xylan conformers, which can be used for structural and compositional analysis.

Tremendous efforts have been devoted to further improving MAS-DNP efficiency. Multiple biradicals have been developed recently, such as the AsymPolPOK family that shortens the DNP buildup time[33] and the Tinypol and TEMTRIPol-I radicals that have shown improved performance over AMUPol on high-field (18.8 and 21.1 Tesla) DNP instruments[31, 57, 58]. The resolution improvement at high-field MAS-DNP (e.g. 800 MHz/527 GHz) will allow us to better

analyze biomolecular structures, at least for the molecules bearing structural order[59, 60]. The rapidly evolving MAS-DNP technology has a great potential to bring biomass analysis to a new level.

8.5 Materials and Methods

8.5.1 Rice Stem Preparation

The mutant *ctl1* was generated by backcrossing a previously reported mutant *brittle culm 15* (A213L mutation in *OsCTL1*) into *Oryza sativa* cv. Nipponbare background[41]. The *ctl2* is an insertional mutant (NC2596 from rice *Tos17* insertion mutant database) with a *Tos17* insertion at the 1633 bp downstream of the *OsCTL2* (*LOC_Os08g41100*) coding sequence. The double mutant *ctl1 ctl2* was created by crossing *ctl1* and *ctl2* and then screening out from the F_2 progenies via molecular identification. All the WT and mutant rice plants (*Oryza sativa*) were grown in the experimental fields at the Institute of Genetics and Developmental Biology in Beijing (China). At least five mature stems from different plants of each genotype were harvested for the measurement.

8.5.2 Dynamic Nuclear Polarization Sample Preparation

The unlabeled rice stems were sliced using a razor into pieces that are on the dimension of a few millimeters. The materials were mixed with a stock solution, which contains 10 mM AMUPol radical[30] and a solvent mixture of ^{13}C-depleted, d_8-glycerol/D_2O/H_2O (60/30/10 vol%). Around 50 mg of the rice material was impregnated in 100 μL of the AMUPol solution and then ground for 20-30 min, which allows the radicals to penetrate the porous cell walls. Around 33 mg of the plant material was transferred to a 3.2-mm sapphire rotor for DNP experiments. In addition, a different sample has been prepared using DMSO/D_2O/H_2O (60/30/10 vol%) solvent to evaluate the effect of different solvents on DNP efficiency. A video protocol can be found in reference[61].

8.5.3 Dynamic Nuclear Polarization Experiment

All DNP experiments were conducted on a 600 MHz (14.1 Tesla)/395 GHz MAS-DNP

instrument at National High Magnetic Field Laboratory, Tallahassee, Florida[27]. The DNP spectra were collected using a 3.2-mm probe under 8 kHz MAS. All ^{13}C chemical shifts were reported on the TMS scale. The microwave irradiation is around 12 W. The temperature was 104 K with microwave irradiation and decreased to 94 K when the microwave was turned off. The DNP build up time was measured to be 2.80 s for the wild-type sample and 2.63 s for the *ctl1 ctl2* sample. Therefore, the recycle delay was typically 3.6 and 3.2 s for the wild-type and *ctl1 ctl2* samples, respectively. 1D ^{13}C CP spectra were collected using a 1 ms contact time. A total number of 256 and 512 scans were collected on the wild-type, and *ctl1 ctl2* samples, respectively. The total experimental time ranges from 16 to 29 mins for each sample. Spectral deconvolution of the cellulose peaks was performed using DmFit software[62].

Two types of 2D ^{13}C-^{13}C correlation experiments were conducted on the rice stems: a 2D CP J-INADEQUATE experiment that probes through-bond correlations[49] and a 2D CHHC experiment that probes spatial correlations[52]. For the INADEQUATE spectra, a total number of 320 scans were recorded within 35 h for the wild-type sample. The indirect dimension includes 110 points. CHHC spectrum was much more time-consuming and was only collected on the wild-type sample. For 1 ms CHHC, a total of 320 scans were collected over 36 h, with 64 points in the indirect dimension, and for 2 ms CHHC, a total of 608 scans were collected over 60 h, with the indirect dimension varying from 50- 60 points. The CHHC experiment relied on 3 CP blocks to transfer polarization from ^{1}H to ^{13}C, back to ^{1}H to enable ^{1}H spin diffusion, and then back to ^{13}C for site-specific detection. The corresponding contact times of the three CP blocks were 1 ms, 0.5 ms, and 0.5 ms, respectively.

8.5.4 Room-Temperature Solid-State NMR

Multiple experiments were conducted at room-temperature to compare with the DNP data.

Around 65 mg of unlabeled rice stems were packed into a 4-mm ZrO_2 rotor for measurements. All non-DNP experiments were conducted on a 400 MHz (9.4 Tesla) Bruker spectrometer using a 4-mm HCN probe. 1D ^{13}C quantitative spectra were collected using the recently developed MultiCP pulse sequence at room temperature under 10 kHz [43]. This experiment allows for the quantitative detection of all ^{13}C signals with enhanced sensitivity. Within each experiment, 8 CP blocks were used for WT and *ctl1 ctl2* double mutant sample, and 9 blocks were used for *ctl1* and *ctl2* single mutant rice, with a z-filter time of 0.9 s between two CP blocks for repolarization. In total, 20k, 23k, 23k, and 20k scans were collected for the WT, *ctl1*, *ctl2*, and *ctl1 ctl2* samples, respectively. To investigate the molecular motion of polysaccharides in unlabeled rice stems, we have measured ^{13}C-T1 and ^{1}H-$T_{1\rho}$ relaxation. ^{13}C-T_1 was measured using a CP-based experiment[63]. ^{1}H-$T_{1\rho}$ was measured using a spinlock field of 62.5 kHz. For ^{13}C T_1 relaxation, 4k and 3k scans are collected for each time point of WT and *ctl1 ctl2* samples, respectively. Meanwhile, 6k (WT) and 2k (*ctl1 ctl2*) scans are collected for each time point of the ^{1}H-$T_{1\rho}$ relaxation experiment. In addition, we have conducted the dipolar-chemical shift (DIPSHIFT) correlation experiment under 7.5 kHz MAS[56]. Frequency-switched Lee Goldburg (FSLG) sequence was used for ^{1}H homonuclear decoupling, with a transverse field of 83 kHz and an effective field of 102 kHz. The scaling factor is 0.577 as verified using a model tripeptide MLF.

In addition, we have measured a series of 2D PDSD experiments using ^{13}C-labeled rice stems to compare with the DNP-assisted 2D CHHC experiment collected on unlabeled samples. Around 85 mg ^{13}C-labeled rice material was packed into a 4-mm ZrO_2 rotor for the solid-state NMR experiment. The mixing times were chosen to be 1, 2, 3, and 5 ms for four different spectra.

8.6 Conclusion

We have presented a novel strategy that integrates natural-abundance DNP with dynamics

measurement to analyze unlabeled plant cell walls. Because this protocol no longer requires isotopic enrichment, it now becomes possible to rapidly screen a large collection of lignocellulose materials found in nature or engineered in vitro. Consequently, the large-scale and high-resolution biomass characterization will provide the structural foundation for improving the biotechnologies of biofuel production.

8.7 Acknowledgments

The DNP analysis of plant cell walls is supported by the U.S. Department of Energy (grant no. DE-SC0021210). Room-temperature experiments and spectral deconvolution were conducted by FD, who is supported as part of the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award number DE-SC0001090. YZ and BZ were supported by the National Nature Science Foundation of China (31571247). The National High Magnetic Field Laboratory is supported by National Science Foundation through NSF/DMR-1644779 and the State of Florida. The MAS-DNP instrument at NHMFL is funded in part by NIH S10 OD018519, NIH P41-GM122698-01, and NSF CHE-1229170.

REFERENCES

1. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: A critical review. *Solid State Nucl. Magn. Reson.* **2020,** *107*, 101660.

2. Dick-Perez, M.; Zhang, Y. A.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry-Us* **2011,** *50* (6), 989-1000.

3. Wang, T.; Zabotina, O.; Hong, M., Pectin-cellulose interactions in the *Arabidopsis* primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry-Us* **2012,** *51* (49), 9846-9856.

4. Wang, T.; Chen, Y. N.; Tabuchi, A.; Cosgrove, D. J.; Hong, M., The Target of beta-Expansin EXPB1 in Maize Cell Walls from Binding and Solid-State NMR Studies. *Plant Physiol.* **2016,** *172* (4), 2107-2119.

5. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and dynamics of *Brachypodium* primary cell wall polysaccharides from two-dimensional ^{13}C solid-state nuclear magnetic resonance spectroscopy. *Biochemistry-Us* **2014,** *53* (17), 2840-2854.

6. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 4978.

7. Grantham, N. J.; Wurman-Rodrich, J.; Terrett, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **2017,** *3* (11), 859-865.

8. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

9. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry-Us* **2015,** *54* (14), 2335-2345.

10. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 347.

11. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. *J. Biomol. NMR* **2018,** *70* (2), 123-131.

12. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.;

Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

13. Chrissian, C.; Camacho, E.; Kelly, J. E.; Wang, H.; Casadevall, A.; Stark, R. E., Solid-state NMR spectroscopy identifies three classes of lipids in C. neoformans melanized cell walls and whole fungal cells. *J Biol Chem* **2020**.

14. Smith, P. J.; Wang, H. T.; York, W. S.; Pena, M. J.; Urbanowicz, B. R., Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnol. Biofuels* **2017,** *10*.

15. Gao, Y.; Mortimer, J. C., Unlocking the architecture of native plant cell walls via solid-state nuclear magnetic resonance. *Methods Cell Biol.* **2020,** *160*, 121-143.

16. Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Coperet, C.; Emsley, L., Dynamic Nuclear Polarization Surface Enhanced NMR Spectroscopy. *Acc. Chem. Res.* **2013,** *46* (9), 1942-1951.

17. Ni, Q. Z.; Daviso, E.; Can, T. V.; Markhasin, E.; Jawla, S. K.; Swager, T. M.; Temkin, R. J.; Herzfeld, J.; Griffin, R. G., High Frequency Dynamic Nuclear Polarization. *Acc. Chem. Res.* **2013,** *46*, 1933-1941.

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

19. Lee, D.; Hediger, S.; De Paepe, G., Is solid-state NMR enhanced by dynamic nuclear polarization? *Solid State Nucl. Magn. Reson.* **2015,** *66-67*, 6-20.

20. Smith, A. N.; Marker, K.; Hediger, S.; De Paepe, G., Natural Isotopic Abundance ^{13}C and ^{15}N Multidimensional Solid-State NMR Enabled by Dynamic Nuclear Polarization. *The journal of physical chemistry letters* **2019,** *10* (16), 4652-4662.

21. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paepe, G., Rapid Natural-Abundance 2D 13C-13C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew. Chem. Int. Edit.* **2012,** *51* (47), 11766-11769.

22. Marker, K.; Paul, S.; Fernandez-de-Alba, C.; Lee, D.; Mouesca, J. M.; Hediger, S.; De Paepe, G., Welcoming natural isotopic abundance in solid-state NMR: probing pi-stacking and supramolecular structure of organic nanoassemblies using DNP. *Chem. Sci.* **2017,** *8* (2), 974-987.

23. Kirui, A.; Ling, Z.; Kang, X.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic Resolution of Cotton Cellulose Structure Enabled by Dynamic Nuclear Polarization Solid-State NMR. *Cellulose* **2019,** *26*, 329-339.

24. Zhang, L.; Gao, C.; Mentink-Vigier, F.; Tang, L.; Zhang, D.; Wang, S.; Cao, S.; Xu, Z.; Liu, X.; Wang, T.; Zhou, Y.; Zhang, B., Arabinosyl Deacetylase Modulates the Arabinoxylan Acetylation Profile and Secondary Wall Formation. *Plant Cell* **2019**.

25. Perras, F. A.; Luo, H.; Zhang, X.; Mosier, N. S.; Pruski, M.; Abu-Omar, M. M., Atomic-Level Structure Characterization of Biomass Pre- and Post-Lignin Treatment by Dynamic Nuclear Polarization-Enhanced Solid-State NMR. *J. Phys. Chem. A* **2017,** *121* (3), 623-630.

26. Sergeyev, I. V.; Aussenac, F.; Purea, A.; Reiter, C.; Bryerton, E.; Retzloff, S.; Hesler, J.; Tometich, L.; Rosay, M., Efficient 263 GHz magic angle spinning DNP at 100 K using solid-state diode sources. *Solid State Nucl. Magn. Reson.* **2019,** *100*, 63-69.

27. Dubroca, T.; Smith, A. N.; Pike, K. J.; Froud, S.; Wylde, R.; Trociewitz, B.; Mckay, J.; Mentink-Vigier, F.; van Tol, J.; Wi, S.; Brey, W.; Long, J. R.; Frydman, L.; Hill, S., A quasi-optical and corrugated waveguide microwave transmission system for simultaneous dynamic nuclear polarization NMR on two separate 14.1 T spectrometers. *J Magn Reson* **2018,** *289*, 35-44.

28. Matsuki, Y.; Idehara, T.; Fukazawa, J.; Fujiwara, T., Advanced instrumentation for DNP-enhanced MAS NMR for higher magnetic fields and lower temperatures. *J Magn Reson* **2016,** *264*, 107-115.

29. Chaudhari, S. R.; Berruyer, P.; Gajan, D.; Reiter, C.; Engelke, F.; Silverio, D. L.; Coperet, C.; Lelli, M.; Lesage, A.; Emsley, L., Dynamic nuclear polarization at 40 kHz magic angle spinning. *Phys Chem Chem Phys* **2016,** *18* (15), 10616-22.

30. Sauvee, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem. Int. Edit.* **2013,** *52* (41), 10858-10861.

31. Lund, A.; Casano, G.; Menzildjian, G.; Kaushik, M.; Stevanato, G.; Yulikov, M.; Jabbour, R.; Wisser, D.; Renom-Carrasco, M.; Thieuleux, C.; Bernada, F.; Karoui, H.; Siri, D.; Rosay, M.; Sergeyev, I. V.; Gajan, D.; Lelli, M.; Emsley, L.; Ouari, O.; Lesage, A., TinyPols: a family of water-soluble binitroxidestailored for dynamic nuclear polarization enhancedNMR spectroscopy at 18.8 and 21.1 T. *Chem. Sci.* **2020,** *11* 2810.

32. Stevanato, G.; Casano, G.; Kubicki, D. J.; Rao, Y.; Esteban Hofer, L.; Menzildjian, G.; Karoui, H.; Siri, D.; Cordova, M.; Yulikov, M.; Jeschke, G.; Lelli, M.; Lesage, A.; Ouari, O.; Emsley, L., Open and Closed Radicals: Local Geometry around Unpaired Electrons Governs Magic-Angle Spinning Dynamic Nuclear Polarization Performance. *J. Am. Chem. Soc.* **2020**.

33. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paepe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

34. Atalla, R. H.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. *Science* **1984,** *223* (4633), 283-285.

35. Ha, M. A.; Apperley, D. C.; Evans, B. W.; Huxham, M.; Jardine, W. G.; Vietor, R. J.; Reis, D.; Vian, B.; Jarvis, M. C., Fine structure in cellulose microfibrils: NMR evidence from onion and quince. *The Plant journal : for cell and molecular biology* **1998,** *16* (2), 183-190.

36. Du, X.; Gellerstedt, G.; Li, J., Universal fractionation of lignin–carbohydrate complexes (LCCs) from lignocellulosic biomass: an example using spruce wood. *The Plant journal : for cell and molecular biology* **2013,** *74* (2), 328-338.

37. Bardet, M.; Gerbaud, G.; Giffard, M.; Doan, C.; Sabine, H.; Le Page, L., ^{13}C high-resolution solid-sate NMR for structural elucidation of archaeological woods. *Prog. Nucl. Magn. Reson. Sp.* **2009,** *55*, 199-224.

38. Hermans, C.; Porco, S.; Vandenbussche, F.; Gille, S.; De Pessemier, J.; Van Der Straeten, D.; Verbruggen, N.; Bush, D. R., Dissecting the role of CHITINASE-LIKE1 in nitrate-dependent changes in root architecture. *Plant Physiol.* **2011,** *157* (3), 1313-26.

39. Jiao, S.; Hazebroek, J. P.; Chamberlin, M. A.; Perkins, M.; Sandhu, A. S.; Gupta, R.; Simcox, K. D.; Yinghong, L.; Prall, A.; Heetland, L.; Meeley, R. B.; Multani, D. S., Chitinase-like1 Plays a Role in Stalk Tensile Strength in Maize. *Plant Physiol.* **2019,** *181* (3), 1127-1147.

40. Sanchez-Rodriguez, C.; Bauer, S.; Hematy, K.; Saxe, F.; Ibanez, A. B.; Vodermaier, V.; Konlechner, C.; Sampathkumar, A.; Ruggeberg, M.; Aichinger, E.; Neumetzler, L.; Burgert, I.; Somerville, C.; Hauser, M. T.; Persson, S., Chitinase-like1/pom-pom1 and its homolog CTL2 are glucan-interacting proteins important for cellulose biosynthesis in Arabidopsis. *Plant Cell* **2012,** *24* (2), 589-607.

41. Wu, B.; Zhang, B.; Dai, Y.; Zhang, L.; Shang-Guan, K.; Peng, Y.; Zhou, Y.; Zhu, Z., Brittle culm15 encodes a membrane-associated chitinase-like protein required for cellulose biosynthesis in rice. *Plant Physiol.* **2012,** *159* (4), 1440-52.

42. Duan, P.; Schmidt-Rohr, K., Composite-pulse and partially dipolar dephased multiCP for improved quantitative solid-state (13)C NMR. *J Magn Reson* **2017,** *285*, 68-78.

43. Johnson, R. L.; Schmidt-Rohr, K., Quantitative solid-state ^{13}C NMR with signal enhancement by multiple cross polarization. *J. Magn. Reson.* **2014,** *239*, 44-49.

44. Bernardinelli, O. D.; Lima, M. A.; Rezende, C. A.; Polikarpov, I.; deAzevedo, E. R., Quantitative C-13 MultiCP solid-state NMR as a tool for evaluation of cellulose crystallinity index measured directly inside sugarcane biomass. *Biotechnol. Biofuels* **2015,** *8*.

45. Takahashi, H.; Hediger, S.; De Paepe, G., Matrix-free dynamic nuclear polarization enables solid-state NMR C-13-C-13 correlation spectroscopy of proteins at natural isotopic abundance. *Chem. Commun.* **2013,** *49* (82), 9479-9481.

46. Fernandez-de-Alba, C.; Takahashi, H.; Richard, A.; Chenavier, Y.; Dubois, L.; Maurel, V.; Lee, D.; Hediger, S.; De Paepe, G., Matrix-Free DNP-Enhanced NMR Spectroscopy of Liposomes Using a Lipid-Anchored Biradical. *Chem. Eur. J.* **2015,** *21* (12), 4512-4517.

47. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of Pretreated Wood Fibers Using Dynamic Nuclear Polarization. *J. Phys. Chem. C.* **2019,** *123* (50), 30407-30415.

48. Ling, Z.; Wang, T.; Makarem, M.; Cintron, M. S.; Cheng, H. N.; Kang, X.; Bacher, M.; Potthast, A.; Rosenau, T.; King, H.; Delhom, C. D.; Nam, S.; Edwards, J. V.; Kim, S. H.; Xu, F.; French, A. D., Effects of ball milling on the structure of cotton cellulose. *Cellulose* **2019,** *26* (1), 305-328.

49. Lesage, A.; Auger, C.; Caldarelli, S.; Emsley, L., Determination of through-bond carbon-carbon connectivities in solid-state NMR using the INADEQUATE experiment. *J. Am. Chem. Soc.* **1997,** *119* (33), 7867-7868.

50. Lesage, A.; Bardet, M.; Emsley, L., Through-Bond Carbon−Carbon Connectivities in Disordered Solids by NMR. *J. Am. Chem. Soc.* **1999,** *121* (47), 10987-10993.

51. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

52. Aluas, M.; Tripon, C.; Griffin, J. M.; Filip, X.; Ladizhansky, V.; Griffin, R. G.; Brown, S. P.; Filip, C., CHHC and ^{1}H-^{1}H magnetization exchange: Analysis by experimental solid-state NMR and 11-spin density-matrix simulations. *J Magn Reson* **2009,** *199* (2), 173-187.

53. Kobayashi, T.; Slowing, I. I.; Pruski, M., Measuring Long-Range ^{13}C–^{13}C Correlations on a Surface under Natural Abundance Using Dynamic Nuclear Polarization-Enhanced Solid-State Nuclear Magnetic Resonance. *J. Phys. Chem. C* **2017,** *121,* 24687-24691.

54. Teymoori, G.; Pahari, B.; Stevensson, B.; Eden, M., Low-power broadband homonuclear dipolar recoupling without decoupling: Double-quantum C-13 NMR correlations at very fast magic-angle spinning. *Chem Phys Lett* **2012,** *547*, 103-109.

55. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis thaliana Primary Cell Walls: Evidence from Solid-State NMR. *Plant Physiol.* **2015,** *168* (3), 871-884.

56. Hong, M.; Gross, J. D.; Rienstra, C. M.; Griffin, R. G.; Kumashiro, K. K.; Schmidt-Rohr, K., Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides. *J Magn Reson* **1997,** *129* (1), 85-92.

57. Mentink-Vigier, F.; Mathies, G.; Liu, Y.; Barra, A. L.; Caporini, M.; Lee, D.; Hediger, S.; Griffin, R. G.; De Paepe, G., Efficient cross-effect dynamic nuclear polarization without depolarization in high-resolution MAS NMR. *Chem. Sci.* **2017,** *8*, 8150.

58. Zhai, W.; Lucini Paioni, A.; Cai, X.; Narasimhan, S.; Medeiros-Silva, J.; Zhang, W.; Rockenbauer, A.; Weingarth, M.; Song, Y.; Baldus, M.; Liu, Y., Postmodification via Thiol-click Chemistry Yields Hydrophilic Trityl-nitroxide Biradicals for Biomolecular High-Field Dynamic Nuclear Polarization. *The journal of physical chemistry. B* **2020**.

59. Jaudzems, K.; Bertarello, A.; Chaudhari, S. R.; Pica, A.; Cala-De Paepe, D.; Barbet-Massin, E.; Pell, A. J.; Akopjana, I.; Kotelovica, S.; Gajan, D.; Ouari, O.; Tars, K.; Pintacuda, G.; Lesage, A., Dynamic Nuclear Polarization-Enhanced Biomolecular NMR Spectroscopy at High

Magnetic Field with Fast Magic-Angle Spinning. *Angew. Chem. Int. Edit.* **2018,** *57* (25), 7458-7462.

60. Berruyer, P.; Bjorgvinsdottir, S.; Bertarello, A.; Stevanato, G.; Rao, Y.; Karthikeyan, G.; Casano, G.; Quari, O.; Lelli, M.; Reiter, C.; Engelke, F.; Emsley, L., Dynamic Nuclear Polarization Enhancement of 200 at 21.15 T Enabled by 65 kHz Magic Angle Spinning. *The journal of physical chemistry letters* **2020,** *11*, 8386-8391.

61. Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Wang, P.; Kang, X.; Wang, T., Preparation of Fungal and Plant Materials for Structural Elucidation Using Dynamic Nuclear Polarization Solid-State NMR. *J. Vis. Exp.* **2019,** *144*, e59152.

62. Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calve, S.; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z. H.; Hoatson, G., Modelling one- and two-dimensional solid-state NMR spectra. *Magn. Reson. Chem.* **2002,** *40* (1), 70-76.

63. Torchia, D. A., Measurement of Proton-Enhanced ^{13}C-T_1 Values by a Method Which Suppresses Artifacts. *J Magn Reson* **1978,** *30* (3), 613-616.

APPENDIX

Table S8.1 Fit parameters of ^{13}C CP spectrum of wild-type (WT) sample. The peaks are classified in three groups according to their influence on the C4 region, as by integrating, interior and surface cellulose contribution is shown to respectively be 40.5% and 59.5% of the cellulose content.

δ [ppm]	**Assignment**	**Amplitude**	**Width [ppm]**	**Integral [%]**
Cellulose C4				
90.0		0.4	0.8	0.1
89.0	i4	6.2	1.2	3.3
88.0		1.7	1.2	0.9
87.3		1.0	1.0	0.4
84.3	s4	6.8	1.6	4.7
83.1		4.1	1.2	2.2
Other peaks close to cellulose C4				
81.9		2.7	1.8	2.1
80.5		1.3	1.8	1.0
76.1		13.4	2.6	10.3
Other peaks far from cellulose C4				
105.8		13.4	1.1	6.4
104.9		11.9	1.1	5.7
103.9		10.3	1.2	5.4
74.9		20.8	1.2	7.5
74.0		17.0	1.0	4.8

Table S8.1 (cont'd)

72.1		40.4	2.3	27.4
71.3		3.9	0.4	0.4
64.9		11.0	1.6	5.2
64.4		7.0	0.7	1.5
63.1		6.5	1.4	2.8
61.9		9.5	1.7	4.9
59.6		3.1	2.4	3.1

Table S8.2 Fit parameters of ^{13}C CP spectrum of *ctl1 ctl2* double mutant. The peaks are classified in three groups according to their influence on C4 region, as by integrating, interior and surface cellulose contribution is shown to respectively be 44.6% and 55.4% of the cellulose content.

δ [ppm]	Assignment	Amplitude	Width [ppm]	Integral [%]
Cellulose C4				
90.0	i4	1.2	0.8	0.4
89.0		6.2	1.0	2.5
88.0		1.8	1.2	0.9
87.4	s4	0.9	1.0	0.3
84.3		4.9	1.5	2.9
83.5		0.7	0.8	0.2
83.0		3.4	1.5	1.9
Other peaks close to cellulose C4				
81.9		3.2	1.8	2.3
81.0		1.3	1.0	0.5
80.0		1.5	1.4	0.8
76.1		13.4	2.6	11.3
Other peaks far from cellulose C4				
105.8		9.5	0.8	3.2
104.9		13.0	1.5	7.3
103.8		10.3	1.3	5.3
100.8		1.5	4.4	2.4
74.9		23.8	1.2	7.7

Table S8.2 (cont'd)

74.0		19.0	1.0	4.9
72.1		39.1	2.3	23.9
71.3		7.1	0.5	1.0
70.5		6.0	0.7	1.1
64.9		13.1	1.7	5.9
64.4		3.3	1.0	0.9
63.2		8.3	1.6	3.5
61.9		6.9	1.7	3.2
59.6		4.3	3.5	5.6

Table S8.3 Peak numbers of INADEQUATE spectra shown in Figure 8.3.

	interior cellulose (i)	surface cellulose (s)	xylose in xylan (Xn)	xylose in xyloglucan (x)	arabinose (A)
WT	18	16	6	6	0
ctl1 *ctl2*	16	15	18	2	10

Table S8.4 ^{13}C-T_1 and ^{1}H-T_{1r} relaxation times of cellulose and xylan in WT and *ctl1 ctl2* samples**.** The data is fit using single exponential equation $I\ (t) = e^{-t/T}$, where T could be T_1 or T_{1r}. Error bars are standard deviations of the fitting parameters. CS: ^{13}C chemical shift. Unidentified (-).

Assignments	**Rice WT**				**Rice *ctl1 ctl2* double mutant**			
	CS (ppm)	T_1(CP) (s)	CS (ppm)	$T_{1\rho}$ (ms)	CS (ppm)	T_1(CP) (s)	CS (ppm)	$T_{1\rho}$ (ms)
i/s/Xn^{2f}1	105.6	25±1	105.5	43±2	105.5	11.3±0.8	105.3	31±2
i4	89.3	35±5	89.3	53±2	89.3	21±2	89.1	42±2
s4	84.2	20±1	84.1	33±2	84.1	9.3±0.6	84.0	29±3
i6	65	13±1	64.7	28±3	65.2	7.4±0.8	65	24±2
s6	62.9	4.2±0.5	62.8	24±2	63.1	2.3±0.4	62.8	18±2
Xn-Ac^{CO}	173.7	-	173.7	11±1	174.3	8.3±0.7	174.3	-
$Xn1^{3f}$	102.0	7.7±0.4	102.5	7.9±0.8	102.7	4.8±0.4	101.8	8±1
$Xn4^{2f}$	82.2	10.5±0.9	82.2	18±2	82.2	7.5±0.6	82.2	18±2
$Xn4^{3f}$	77.7	10±2	78.0	16±2	77.5	6.9±0.5	78.0	10±1
Xn-Ac^{Me}	21.6	7.8±0.7	21.7	10±2	21.7	4.6±0.3	21.7	9±1

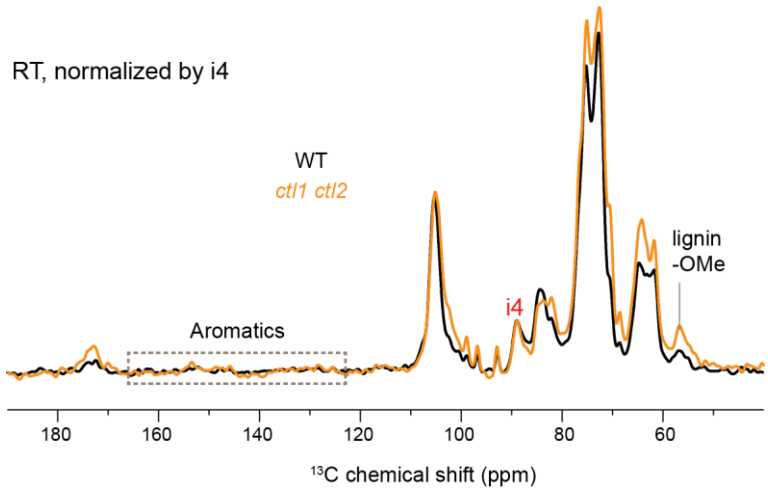

Figure S8.1 Lignin has increased methyl ether substitution in the double mutant. The spectra of wild-type (black) and *ctl1 ctl2* double mutant (yellow) are normalized by the interior cellulose carbon 4 (i4) peak. The lignin methyl ethers (lignin -OMe) has a doubled intensity in *ctl1 ctl2* but the lignin aromatics have a comparable intensity in both samples.

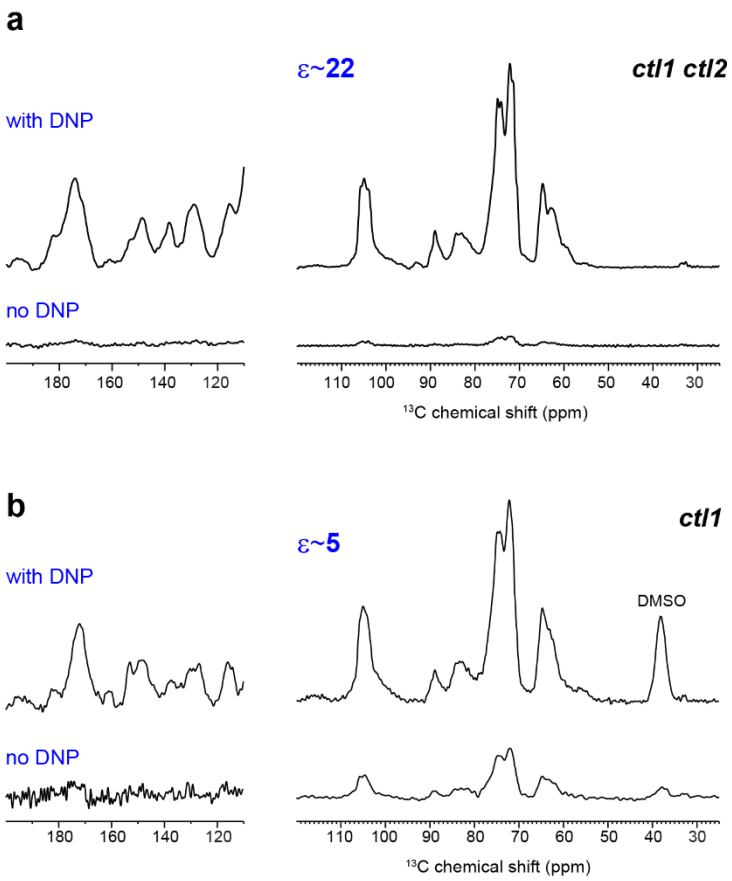

Figure S8.2 Effect of solvents on DNP enhancement. **a,** *ctl1 ctl2* sample in the solvent of ^{13}C-depleted, d_8-glycerol/D_2O/H_2O (60/30/10 vol%) has shown a 22-fold enhancement of sensitivity. **b**, *ctl1* mutant in d_6-DMSO/D_2O/H_2O (60/30/10 vol%) matrix only exhibits a 5-fold of sensitivity gain.

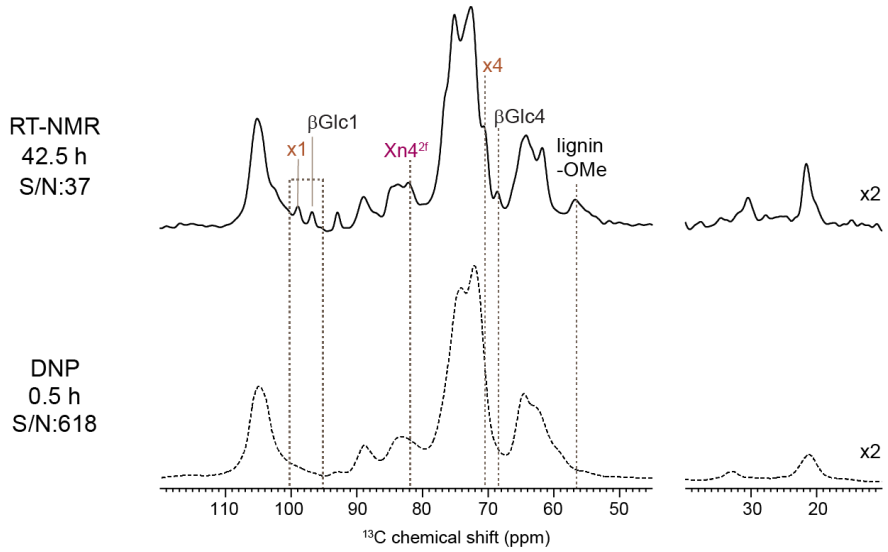

Figure S8.3 Timesaving by DNP on carbohydrate signals in unlabeled rice stems. Top row: The room-temperature spectrum collected on a 400 MHz NMR gives a signal-to-noise (S/N) ratio of 37 for the highest peak after 43 h of measurement. Bottom row: 600 MHz/395 GHz DNP provides a S/N ratio of 618 after only 0.5 h of measurement. Cellulose peaks are well reserved, but intensity suppression has been observed for xylan signals, lignin methyl ether (-OMe) and small molecules (Glc; glucose).

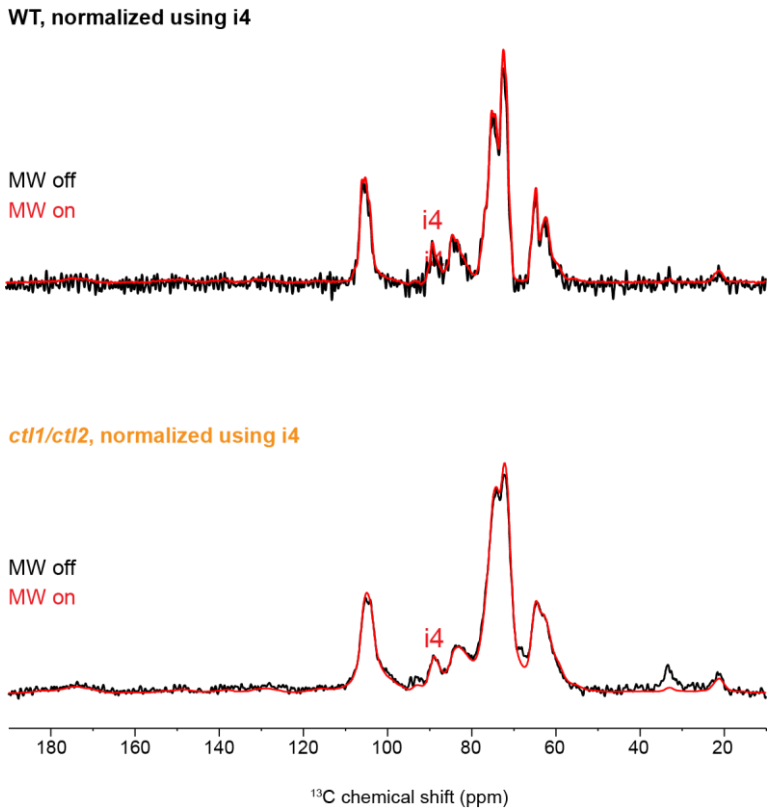

Figure S8.4 DNP polarization is uniform across the cell wall. The microwave-on (MW on) and microwave-off (MW off) spectra are normalized by the interior cellulose carbon 4 peaks (i4) to compare the spectral pattern. The consistent spectral envelope clearly demonstrate that the polarization is uniform across the whole cell wall.

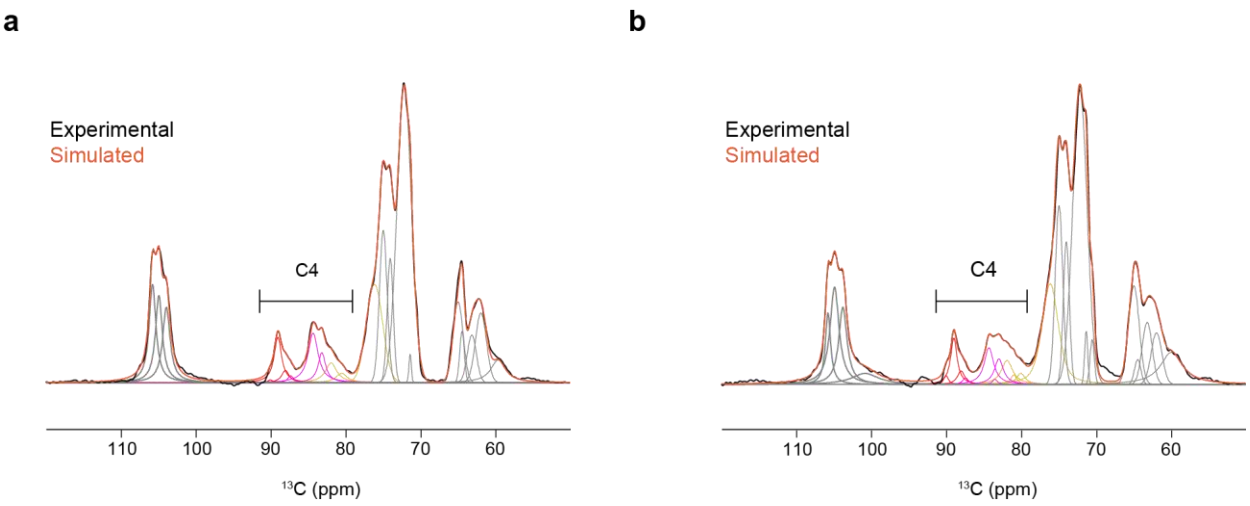

Figure S8.5 The experimental and simulated spectra have a good match. The 120 to 50 ppm regions of **a,** wild-type sample (left) and **b,** *ctl1 ctl2* double mutant (right) are shown. All numerical parameters used to obtain the fits are summarized in Tables S8.1 and S8.2. Color code follows peak classification in these tables: i4 cellulose in red, s4 in magenta, close peaks in dark yellow, and others in grey.

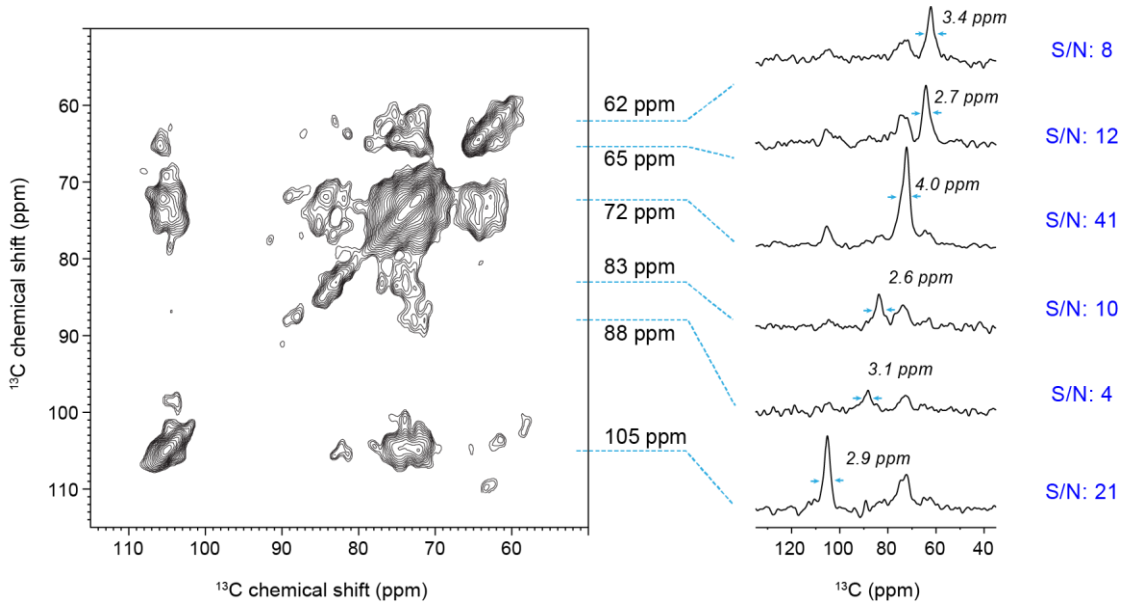

Figure S8.6 1D cross sections of DNP-enabled 2D CHHC spectrum. Representative slices were extracted from the 2 ms CHHC spectrum of unlabeled wild-type rice stem. The ^{13}C FWHM linewidths and signal-to-noise (S/N) ratios are shown for the major peaks.

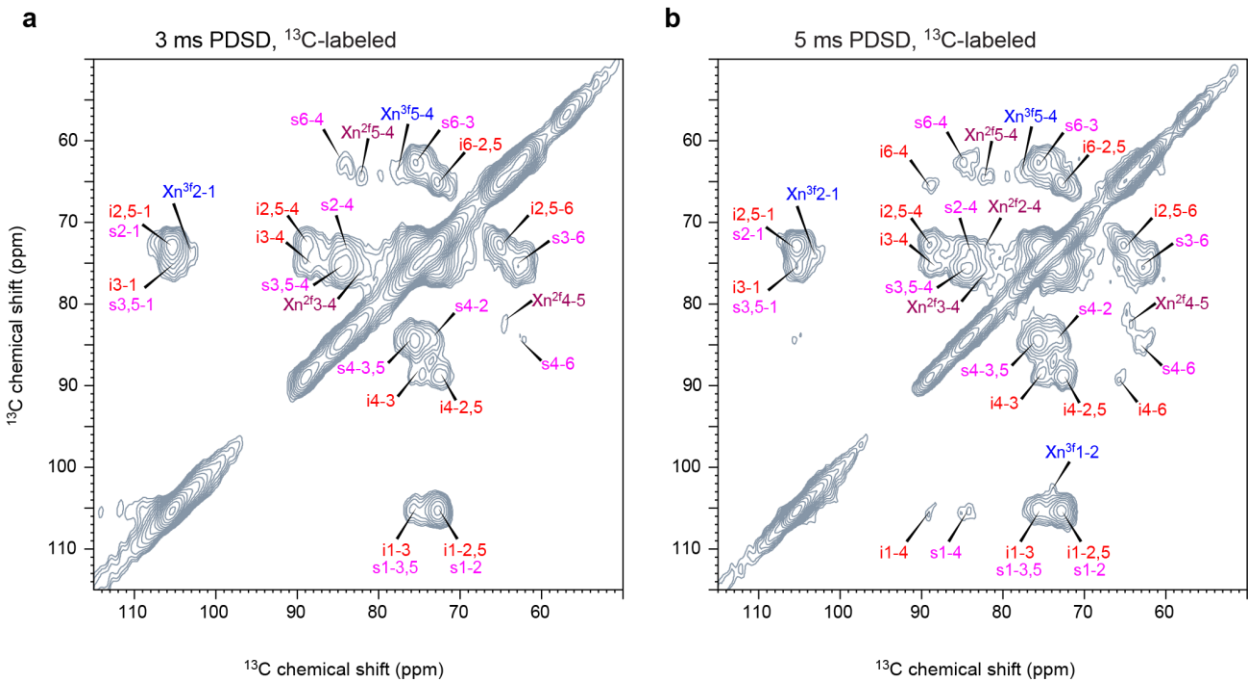

Figure S8.7 2D PDSD spectra of ^{13}C-labeled rice stems. The spectra are collected using **a**, 3 ms and **b**, 5 ms mixing times. Assignments of cellulose and xylan peaks are annotated on the spectra.

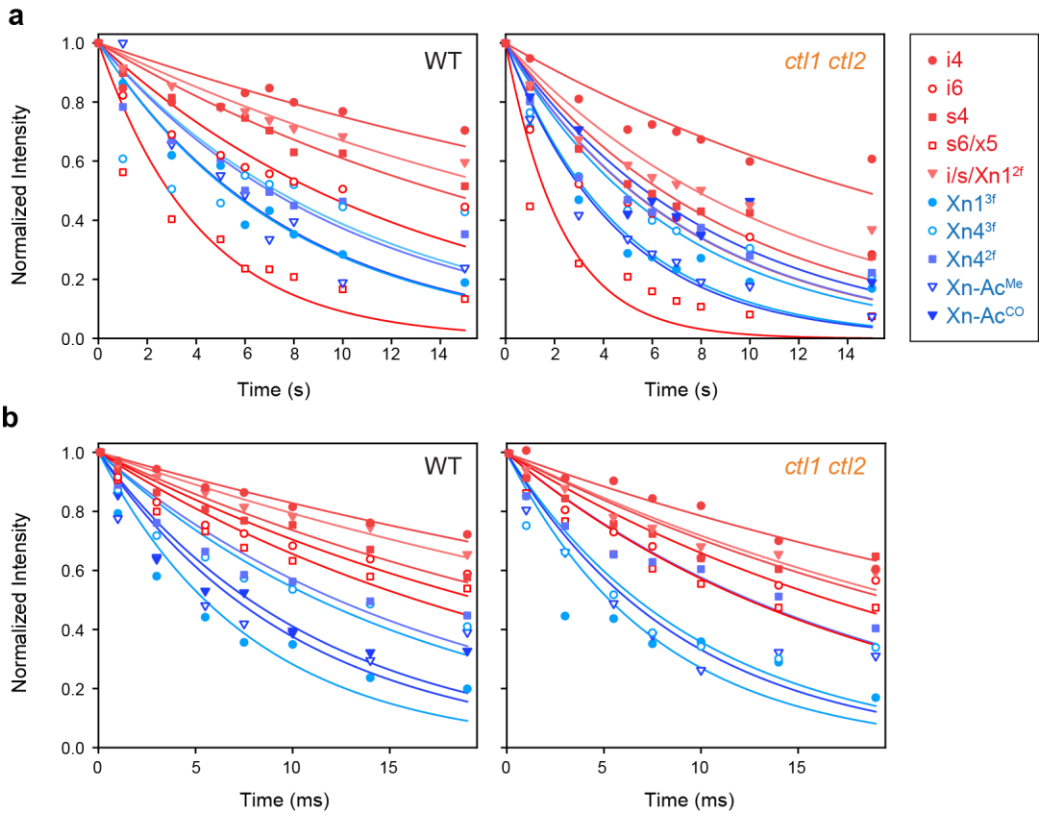

Figure S8.8 NMR relaxation curves of polysaccharides in unlabeled rice stems. The **a**, ^{13}C-T_1 and **b**, ^{1}H-$T_{1\rho}$ data are plotted separately for wild-type and *ctl1 ctl2* samples. Cellulose signals (red) generally exhibit faster relaxation than xylan peaks (blue). The exceptions in panel a only occur to the 62 ppm s6/x5 peak, which has mixed contribution from both cellulose and matrix polymers.

CHAPTER 9: CHRONOLOGICAL MOLECULAR FINGERPRINT OF WETLAND SOIL BY SENSITIVITY-ENHANCED SOLID-STATE NMR

Research paper submitted to *Nature Geoscience*.

Preprinted in *ChemRxiv* (https://doi.org/10.26434/chemrxiv-2023-p4rvg).

Authors: Wancheng Zhao, Elizabeth C Thomas, Debkumar Debnath, Faith Scott, Frederic Mentink-Vigier, John R. White, Robert L. Cook, and Tuo Wang

9.1 Abstract

Soil organic matter (SOM) plays a major role in mitigating greenhouse gas emission and thereby regulating Earth's climate, carbon and water cycles, and biodiversity. Wetland soils contain the highest stores of soil carbon on the planet on an areal basis, accounting for one-third of all the SOM, yet our understanding of carbon sequestration within wetlands lags behind that of upland soils. Here we show the molecular-level fingerprints of wetland soils spanning eleven centuries using advanced solid-state nuclear magnetic resonance (ssNMR) spectroscopy. Remarkably, combining dynamic nuclear polarization (DNP) with SOM enrichment allowed up to an 8,000-fold time-saving over conventional NMR approaches. This innovative approach for SOM characterization revealed that the parent herbaceous plant core molecular structures are preserved, with the aromatic and carbohydrate motifs becoming tightly packed, even after a millennium. Such preserved cores occur alongside molecules from the decomposition of loosely packed parent biopolymers and biogeochemical processing driven by geological global and anthropogenic changes, adding to the chemical diversity of SOM. These findings reveal that particulate organic matter (POM) should be a major focus for wetlands, and other soils with high organic matter content, especially when considering the fate of coastal wetland SOM when exposed to oxygenated water due to erosion.

9.2 Introduction

Soil organic matter (SOM) and associated soil organic carbon (SOC), critically represents one of the major reservoirs of carbon on the planet and is critical to ecosystem services, from the molecular to the global scale[1, 2]. Just the top 1 m of the world's soil contains 1500 petagrams (Pg, billion tons) of carbon, approximately twice the carbon pool contained in the entire atmosphere[3] and more than the atmosphere and vegetation combined[4]. Soil has been a sink to about 210 ± 45 Pg of carbon between the years 1850 and 2021, mitigating around 100 ppm of atmospheric CO_2 levels[5]. Wetland soils contain approximately one-third of this SOC despite occupying only 5-6% of the Earth's land surface area[6]. Coastal wetlands, known as the blue carbon ecosystems, occupy only 0.07-0.22% of the Earth's surface but sequester 0.08-0.22 Pg carbon each year, accounting for more than half of the carbon buried in the oceans annually[7-9], and store the vast majority of this carbon for hundreds to thousands of years, with minimal methane emissions due to a poised redox potential[10].

The anaerobic conditions of wetland soil cause a reduced metabolic efficiency of microbes in acquiring energy during decomposition of complex organic carbon compounds, which leads to increased preservation of carbon in soil profile unless hydrologic conditions change[11]. Subsidence, sea level rise and wave energy promote wetland soil collapse into the surrounding shallow, estuarine aerobic water column, resulting in the oxidation of SOM and releasing the stored carbon back to the atmosphere as primarily carbon dioxide[12-14]. To preserve and manage wetlands, minimize greenhouse gas emissions, and maximize carbon sequestration potential, it is crucial to understand, model, and accurately predict the dynamics of SOM, and hence SOC, in a scalable and applicable manner[15]. Thus, detailed molecular information, including functionality, isomerism, and conformations of the sequestered carbon pool is a fundamental requirement. Currently, there

is limited work on this advanced level of carbon sequestration within coastal wetlands, as most work has been focused on upland soils.

The persistence of SOC was considered to occur through the formation of recalcitrant macromolecules by polymerization during a process called humification[16]. Presently, the dynamic stability concept prevails, whereby organic carbon is preserved as an organic entity via molecular preservation for periods of time, in concert with dynamically changing organic carbon speciation[17]. This new view also breaks SOM down into two major pools: particulate organic matter (POM), which includes occluded SOM and plant materials, and mineral associated organic matter (MAOM), which contains mostly small organic molecules and biological metabolites[18, 19]. Over time, these two more recalcitrant pools may become accessible, and therefore can be viewed as shorter-term stable carbon deposits in the context of geologic time[20] (Swift and Hayes, 2020). Most research on this new paradigm was focused on upland or mineral soils, with the MAOM fraction receiving the bulk of the attention, although the focus is beginning to shift[19, 21, 22]. Most coastal wetlands are POM-dominated systems with a high organic matter to mineral matter ratio. Globally, per unit area, coastal wetlands are considered the most carbon-rich dynamic reservoirs of organic carbon, warranting a closer examination of this critical and most vulnerable to loss carbon pool, as sea level continues to rise.

Here, we examine coastal wetland soil samples from the Mississippi River Delta, USA that have been deposited and preserved over the past 11 centuries, but whose soil C stores are being lost at an accelerated rate due to the high relative sea level rise rate. Sensitivity-enhanced ssNMR is applied, for the first time, to whole soil samples to enable rapid acquisition of 2D $^{13}C/^{1}H$-^{13}C correlation spectra, providing atomic-level structural information of the carbon fraction of SOC. This method allows for a detailed characterization of soil on the fine granular and molecular level,

beyond the moiety characterization level (**Supplementary Discussion**), which reveals that: 1) in the initial stages of SOC sequestration, core structures of the parent herbaceous plant materials are preserved but the aromatic and carbohydrates motifs become tightly packed, with non-carbohydrate components being concentrated in the soil, 2) some structural cores of herbaceous plant biopolymers survive anaerobic microbial degradation, with their original structure and physical packing preserved up to 1000 years since deposition, 3) molecules from the decomposition of loosely packed parent biopolymers and biogeochemical processing are present and add to the diversity of SOC chemical nature, and 4) changes in the carbon speciation during the sequestration processes are driven by both natural geological changes (e.g., delta lobe switching) as well as anthropogenic changes (e.g., river levees). Besides these multifaceted conceptual advances, this high-resolution and rapid technological platform also opens a new research avenue for SOC analysis in undisturbed soil materials.

9.3 Results

9.3.1 Molecular Fingerprinting of Carbohydrate and Aromatics in Wetland Soil

We examined seven soil samples (down to almost 2 m) collected from a salt marsh *Spartina alterniflora*-vegetated island in the Barataria Bay, Louisiana, located along the Gulf of Mexico coastline (**Figure S9.1a-d**). Each year, Louisiana experiences a loss of over 65 km^2 of coastal wetlands, leading to an annual release of over 1 million metric tons of stored carbon from Barataria Basin alone[23, 24]. This island underwent shoreline erosion rates of over 1.5 m y^{-1} and became entirely eroded in 2021 (**Figure S9.1e, f**). A two-step protocol was developed to improve the NMR sensitivity by a factor of 90, reducing experimental duration 8,100 times, essentially reducing a 22-year-long experiment down to one day. This was achieved by employing hydrofluoric acid (HF) treatment, which depleted the mineral component, concentrated SOM, and enhanced their

signals by a factor of 5 (**Figure 9.1a**) without chemically perturbing its structure as described in **Supplementary Methods**[25-27], and DNP, which relies on microwave irradiation to transfer electron polarization to NMR-active nuclei in the soil, resulting in an additional 18 enhancement (**Figure 9.1b**). Together, this allowed for carbon connectivity to be directly mapped by a 2D ^{13}C-^{13}C correlation spectrum on unlabeled soil (**Figure 9.1c**), a task previously impossible by conventional techniques, but now achievable within 16 hours of analytical time.

The carbohydrate signals of the top 10-cm layer of soil (sample 1) were predominantly from cellulose (**Figure 9.1c**), including the glucan chains residing on both the surface and internal domains of the microfibrils (**Figure 9.1d**). Multiple forms of glucan chains are identifiable within cellulose. First, we observed two distinct sets of signals from surface chains (s^f and s^g in **Figure 9.1c**), which have been proposed to contribute to distinctly hydrated exteriors of the microfibrils, namely the concept of hydrophobic and hydrophilic surfaces[28]. Second, in addition to the dominant interior conformers (i^a and i^b), a third form of i^c was also identified in a 2D ^{1}H-^{13}C correlation spectrum, collected within an hour (**Figure 9.1e**), with a unique C4 chemical shift of 87.5 ppm. Type-c glucan has been found in the native cellulose across many different grass and woody plant species and have been attributed to the deeply embedded and inaccessible core of the larger bundle formed by multiple microfibrils (**Figure 9.1d**)[29-31]. The identification of these five glucan types demonstrates that the native cellulosic material has been preserved in surface soil.

Further structural preservation is indicated by the unexpected identification of 2-fold and 3-fold xylan (Xn^{2f} and Xn^{3f}) in the soil (inset of **Figure 9.1c**). The 2-fold and 3-fold refers to the helical screw conformation and indicate the number of sugar residues needed for finishing a 360° helical rotation along the chain (**Figure 9.1f**)[32]. Recent studies of the lignocellulosic plant biomass have revealed the distinct functions of these two xylan conformers, with the flat-ribbon 2-fold

xylan coating the smooth surface of cellulose microfibrils[33, 34] and the zigzag 3-fold xylan preferentially packing with disordered aromatics, namely the lignin domains in plants[30]. Evidently, at least a portion of the structural core of plant lignocellulosic biomass is preserved, intact, in the wetland soil.

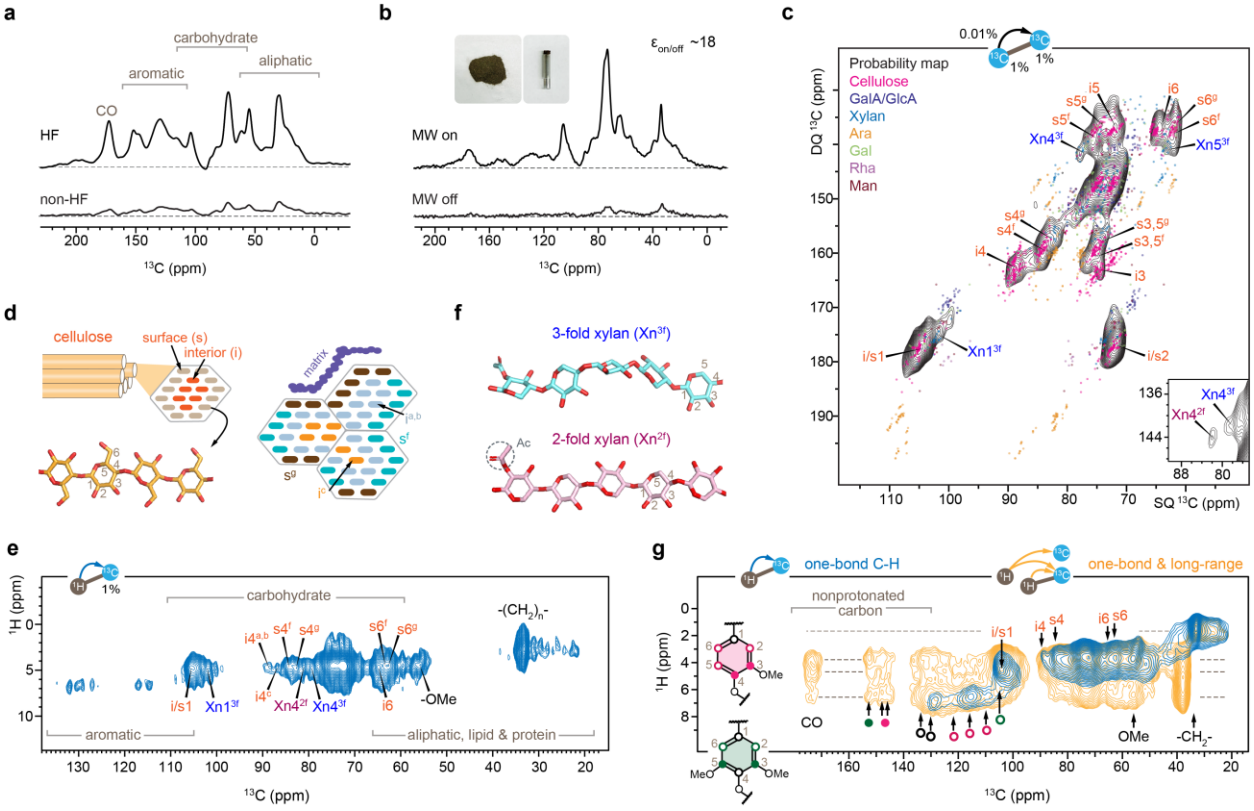

Figure 9.1 Molecular fingerprint and spatial organization of biopolymers in wetland soil. **a,** Comparison of quantitative ^{13}C spectra of the HF and non-HF treated soil sample 1 (surface layer) at room temperature. HF treatment concentrates the organic phase and gives a 5-fold sensitivity boost. **b,** Comparison of 1D ^{13}C spectra with and without microwave (MW) irradiation. The sensitivity enhancement ($\varepsilon_{on/off}$) provided by DNP is 18-fold. Inset shows pictures of the soil sample and a sapphire rotor containing the soil sample. **c**, DNP-enabled 2D ^{13}C-^{13}C correlation spectrum (refocused J-INADEQUATE) of soil sample 1 measured in 16 h, resolving the carbon connectivity of carbohydrate components. Overlay of the measured spectrum (black) with the

Figure 9.1 (cont'd)

probability map constructed using 412 carbohydrate units from Complex Carbohydrate Magnetic Resonance Database[36] indicates the best match with cellulose (magenta dots). Assignments of the interior (i) and surface (s) glucan chains of cellulose microfibrils and the xylose units of 2- and 3-fold xylans (Xn^{2f} and Xn^{3f}) are labeled. Inset shows the xylose carbon 4 region processed with large line-broadening to show xylan signals. **d**, Representative cellulose structure with each microfibril containing eighteen β-1,4-glucan chains on the surface and interior domains. Multiple microfibrils aggregate to form larger bundles that accommodate different forms of glucan chains ($i^{a,b}$, i^{c}, s^{f}, and s^{g}), which are further wrapped by matrix non-cellulosic polymers. **e**, Single-hour DNP 2D ^{1}H-^{13}C correlation spectrum resolves the signals from aliphatic carbons, carbohydrates, and aromatics. Key signals of polymethylene ($-CH_2-)_n$, methoxyl ($-OCH_3$), cellulose, and xylan are labeled. **f**, Structure of 2- and 3-fold xylan conformers. **g**, Overlay of 2D ^{1}H-^{13}C correlation spectrum measured with short (0.1 ms; blue) and long (1.0 ms; yellow) CP contact times. The symbolic representations correspond to the carbons in lignin monolignol units. Dashed lines show the key ^{1}H positions.

Many non-carbohydrate molecules were also identified, including the aromatics and methoxy substitutions in lignin, the acyl chains (or polymethylene[35]) in lipid polymers, and other aliphatic motifs (**Figure 9.1e**). This characterization was achieved using a single-hour 2D experiment that relies on a short CP contact time (0.1 ms) to emphasize one-bond ^{1}H-^{13}C correlations. Protonated carbons in lignin exhibited signals in the ^{13}C chemical shift range of 110-125 ppm (blue spectrum in **Figure 9.1g**). The 105 ppm ^{13}C signals have dual contributions from both carbohydrates and aromatics, including the carbon 1 of cellulose and 2-fold xylan, as well as the carbons 2 and 6 of the S-unit, thus showing one-bond correlations with both carbohydrate and

aromatic protons.

A range of nonprotonated carbons were also observed with a longer CP contact (1.0 ms) that extended the reach of ^{1}H-^{13}C correlation (yellow spectrum of **Figure 9.1g**; **Figure S9.2**). The spotted signals included carbonyl groups (CO) and monolignols, such as the carbon 3 and 5 of the syringyl (S) unit at 154 ppm and the carbon 3 and 4 of the guaiacyl (G) unit at 145-149 ppm. Their chemical nature was confirmed by their strong peaks in dipolar-dephasing spectra that removed all protonated carbon signals (**Figure S9.3**). The NMR observations of SOC unveiled a complex composition in terms of plant polysaccharides, lignin, and lipid polymers.

9.3.2 Domain Distribution of Polymers in Soil

It is noticeable that the aromatic carbons (^{13}C chemical shifts of 100-140 ppm) not only show cross peaks with the aromatic protons at 6-7 ppm, but also cross-talked with carbohydrate and aliphatic protons that resonate at 3-5 ppm, revealing the co-localization of aromatics, aliphatics, and carbohydrates on the nanoscale, consistent with previous work[37]. This concept of molecular mixing is also supported by the cross peaks between carbohydrate carbon sites (^{13}C chemical shifts of 70-110 ppm) with aromatic protons (^{1}H chemical shifts of 6-7 ppm). The only exception was observed in polymethylene (-CH_2-), which failed to show correlations with other carbons or protons, providing a clear indication of domain separation for lipid polymers. This finding corroborates earlier ssNMR results[35, 37] where poly-methylene were found to form large aggregates to resist further degradation, which is a characteristic commonly shared by diverse soil materials in nature.

9.3.3 Preserved Structural Core in Plant Material and Surface Soil

The composition of plant detritus inputs to the soil as well as the redox status of the soil are among the key external factors that affect the rate of carbon sequestration[38]. The plant tissues

gathered at the soil collection site retain highly similar carbohydrate signals when compared to the surface soil (**Figure 9.2a**). Cellulose crystallinity is unchanged, as evidenced by both soil and plant samples showing comparable intensity ratios between the interior cellulose C4 at 89 ppm and the surface cellulose C4 at 84 ppm. With a 24-fold DNP enhancement (**Figure S9.4**), we unambiguously detected the varied signals from multiple cellulose forms and xylan conformers in 2D ^{13}C-^{13}C correlation spectrum of these plants (**Figure 9.2b**), which demonstrated a similar pattern to the soil spectrum. The soil exhibits elevated levels of carbonyls, methoxyls, aromatics, and aliphatics, as revealed by the difference of two parental spectra (**Figure 9.2a**). These components might have accumulated due to their slower decomposition rate when compared to carbohydrates.

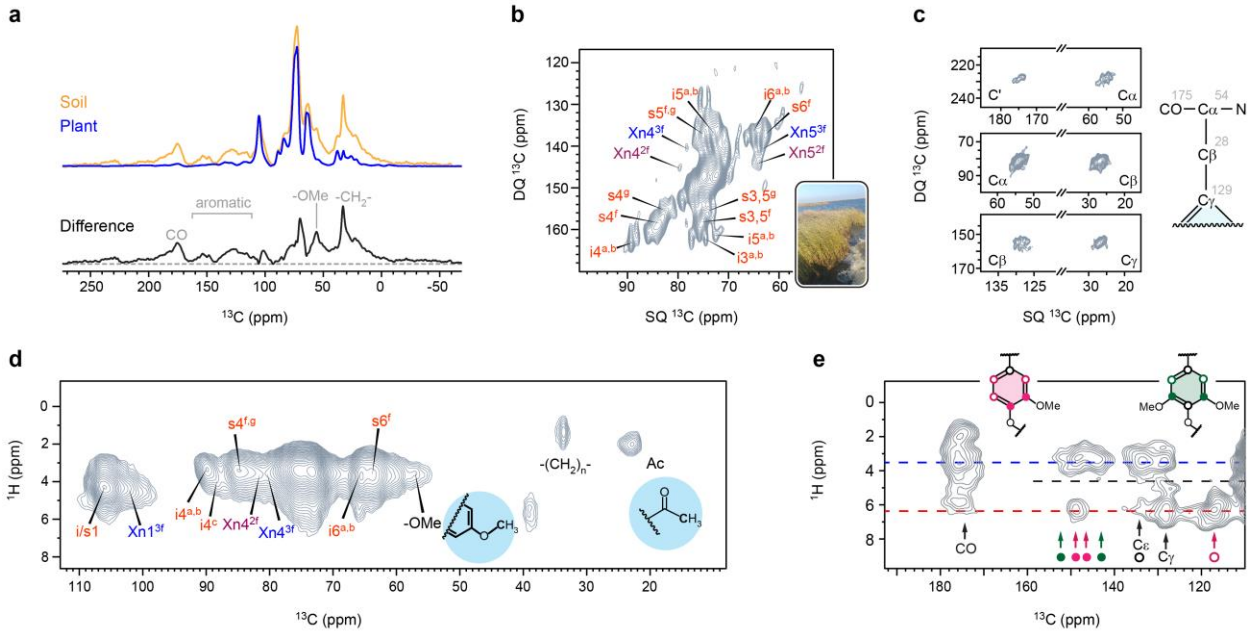

Figure 9.2 Conserved structural core of plant material and wetland soil. **a,** DNP-enhanced ^{13}C spectra of surface soil (yellow; sample 1) and the plant on top of the soil (blue). The bottom panel shows the difference of the two spectra, revealing a signature pattern of lignin, aliphatic and polymethylene, as well as matrix polysaccharides. **b,** Carbohydrate region of DNP enhanced

Figure 9.2 (cont'd)

^{13}C-^{13}C refocused J-INADEQUATE spectrum showing signals from cellulose and xylan. Inset picture shows the original plant material characterized here. **c,** Aromatic amino acids resolved from the plant sample, with chemical shifts labeled on the structure. **d**, 2D ^{1}H-^{13}C correlation spectrum measured using 0.1 ms CP resolving key signals of cellulose, xylan, lipid polymers. Lignin methoxyls (-OCH_3 or -OMe) and xylan acetyls (Ac) are also observed, with structures presented. **e**, 2D ^{1}H-^{13}C correlation spectrum measured using 1 ms CP resolving signals of aromatic and carbonyl carbons from proteins and lignin. Dash lines in blue and red annotate the key positions of methoxyl and aromatic protons, respectively. Black dashed line represents the anticipated correlations between lignin aromatic carbons and carbohydrate anomeric protons, which are notably less numerous than the signals observed in the soil.

Strong signals of aromatic amino acid residues have been spotted, which align with the chemical shifts of histidine or tryptophan (**Figure 9.2c**). These molecules are uniquely abundant in the plant material and are not present in the soil underneath it, likely due to being vulnerable to rapid microbial degradation. Though the carbohydrates signals are highly consistent with the soil, the aliphatic region shows a dramatically simplified pattern (**Figure 9.2d**), with only 3 peaks from the methoxyl group of lignin, the CH_2 groups likely from the acyl chain of lipids in the membrane or from the cutan or suberin, and the acetyl group that serves as an important modification of matrix polysaccharides, such as xylan. The aromatic region is simpler (**Figure 9.2e**). The key carbon sites of both S and G units are still detectable, with these carbons mainly correlating with the methoxyl and aromatic protons and lacking the cross peak with the anomeric protons of carbohydrates at 4.3-4.5 ppm, suggesting substantially reduced interactions between carbohydrates and aromatic polymers.

Despite the conserved structures of individual carbohydrate and lignin components, soil has two unique structural features that are absent in plant materials. First, non-carbohydrate components are highly concentrated, which is likely caused by the faster degradation rate of polysaccharides compared with lignin and polymethylene polymers. Phenols serve as antioxidant during degradation reactions, while polymethylene polymers contain crystalline domains of aliphatic chains, conferring these polymers with high stability[39, 40]. This trend was also confirmed by the analysis of another plant sample collected 30 m inland on the same island (**Figure S9.5**). Secondly, the aromatics and carbohydrates are more tightly packed in the surface soil than in plants. This finding might originate from the faster decay of primary cell walls that only contain cellulose and soft matrix polysaccharides but does not contain lignin, and/or the removal of intra- and extracellular components leading to tighter packing of residual lignocellulosic components.

9.3.4 Mapping Carbon Composition and Packing Along Depth

The molecular composition of the soil matter changes with the depth. In the quantitative 1D ^{13}C spectra[41], the content of polymethylene carbons, marked by the intensities of two adjacent peaks at 33 and 31 ppm, increased substantially from sample 1 (0-10 cm interval) to sample 2 (40-50 cm interval), as shown in **Figure 9.3a**. These two peaks in polymethylene, also observed in many other soil organic matter samples, correspond to crystalline $(CH_2)_n$ chains in all-*trans* conformation (type-a; 33 ppm) and amorphous regions accommodating both *trans* and *gauche* conformations without long-range order (type-b, 31 ppm)[35]. The self-aggregated nature and limited accessibility of polymethylene might have prevented it from being degraded after deposition in the soil. However, when normalized to the aromatic signal, the polymethylene, together with carbohydrates and carbonyls decrease sequentially as one moves deeper in the soil profile (samples 2-7; age range from 1963AD to 945AD). These spectral observations were quantified through a

deconvolution protocol applied to 1D quantitative spectra (**Figure S9.6**), using the carbon sites resolved from high-resolution 2D dataset. Carbohydrates exhibit a substantial reduction in their proportion, decreasing from 29% in sample 1 to 20% in sample 7, while the content of aromatics are being enriched from 29% to 36% when moving deeper from the surface (**Figure 9.3b**).

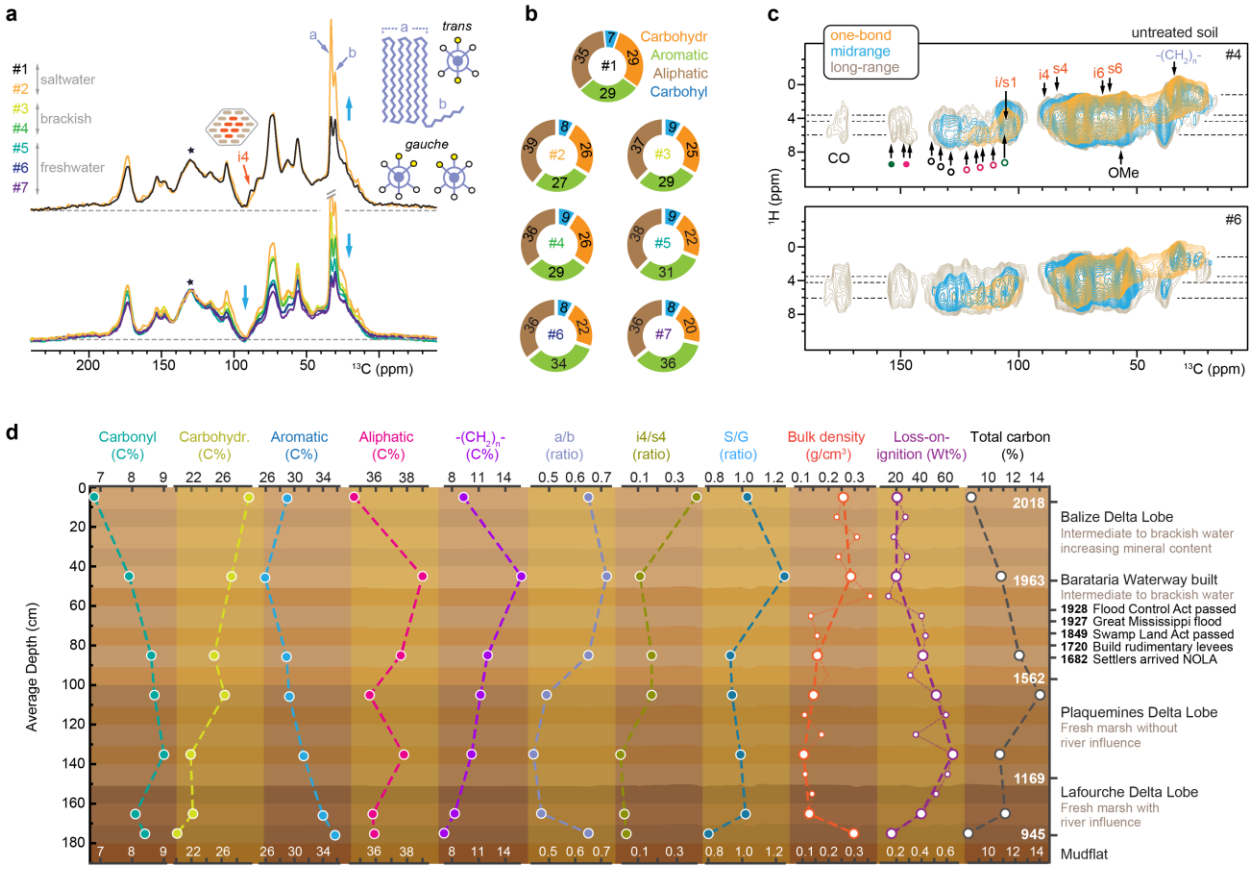

Figure 9.3 Structural changes of SOM in relation to depth. **a**, Overlay of 1D quantitative ^{13}C spectra of seven soil samples with normalization by the major aromatic peak (asterisk). The top panel compares samples 1 and 2, and the bottom panel includes samples 2-7. The two key peaks (types a and b) of polymethylenes are marked, with an illustration of the crystalline all-*trans* domain and the amorphous domain that include both trans and gauche conformations. Three conformers, *gauche*(+), *gauche*(-), and *trans*, are illustrated, with methylene groups represented as yellow circles and hydrogen atoms as open circles. The wetland condition of each soil sample

Figure 9.3 (cont'd)

(saltwater, brackish, or freshwater) are also labeled. **b**, Quantification of the four major carbon types. The details of spectral deconvolution are documented in Figure S9.6, ^{1}H-^{13}C correlation spectra of two untreated soil samples measured with 0.1 ms (yellow), 0.5 ms (light blue), and 1 ms (grey) CP contact times. **d**, molecular and physical evolution of wetland soil over time. The figure depicts a geological timeline through soil depth based on ^{14}C dating results and historic events in the area. Molecular profiling of 7 soil samples reveals the content of major carbon types. Detailed information differentiates different types of polymethylene carbons, cellulose chains, and monolignol units. Bulk density and loss-on-ignition measurements were also taken on 18 soil samples immediately following collection. A higher bulk density indicates a higher mineral content, while a lower loss-on-ignition % suggests a lower organic matter content.

Regarding the ratio between carbohydrate and aromatic moieties, there is a vertical distribution based on salinity as the wetland transitioned from freshwater, to brackish, to a current day salt marsh environment (**Figure 9.3a**)[42]. A similar trend can also be seen regarding the polymethylene-to-aromatics ratio moving from the deeper soil profile to the present-day surface. This finding suggests that, as sea level rises, salinity and resulting shifts in plant species is likely affecting how the sequestered carbon is stored and ultimately degraded in the soil profile.

Native, untreated samples were also investigated using DNP-enhanced 2D ^{1}H-^{13}C correlation experiments (**Figure 9.3c**). It is intriguing that the spectral pattern was consistently maintained in the untreated samples 4 and 6 (**Figure 9.3c**), and in both HF-treated and untreated materials of sample 1 (**Figure S9.7** and **Figure 9.1g**); therefore, HF did not perturb the native structure of the SOC core. This similarity is also an indication that at least a significant portion of the lignocellulosic cores were preserved, supported by the observation of internal cellulose, whose

carbon 4 (i4) shows a major decline starting from sample 2 (**Figure 9.3a**) but still exhibits some weak signals in samples 4 and 6 (**Figure 9.3c**). While most cellulose was decomposed rapidly in the surface layers, a fraction of these crystalline cores were preserved for centuries in this soil.

Moreover, the close spatial proximities between aromatics and carbohydrates observed in samples 4 and 6 (**Figure 9.3c**) resemble those identified in the surface soil (**Figure 9.1g**). The polymethylene also shows the same self-aggregation features in these deeper samples. Therefore, the decomposition of SOM did not happen homogeneously. Some biopolymeric structural cores, like self-aggregated polymethylene and densely packed lignin-polysaccharide domains, have efficiently withstood microbial degradation and maintained their original structure and physical packing after approximately 500 years (at 100 cm depth) and even up to 1000 years (at 180 cm depth). This preservation has been maintained despite the presence of microbial extracellular enzymes present throughout the profile[43].

9.3.5 Natural and Anthropogenic Influences on Carbon Speciation Over a Millennium

The observed nondirectional changes of molecular composition and structure (**Figure 9.3d**) are not expected based on conventional soil aging and humification or the activity of microbial communities at depth. As tracked by the ^{14}C dating, the soil material collected across the ~2-m depth covers a geological timeline of 11 centuries[43, 44]; therefore, the interplays of the delta lobes and water salinity in the Barataria Bay should also play a key role. The Mississippi River watershed is the dominant surface hydrologic feature in North America, which collects runoff from 40% of the continental US between the Rocky Mountain region to the Appalachian Mountains[42]. River deltas are dynamic systems with continually shifting lobe formation and abandonment over time[42]. Therefore, these environmental shifts, over time, can influence the physical and chemical characteristics of the accreted carbon pool based on hydrodynamics and

salinity. One anthropogenic driver has been the construction of river levees over the past century, essentially separating the coastal basins from the river, preventing the historical freshwater and sediment-subsidies from occurring[45].

The continuing decrease of loss-on-ignition (LOI; see Online Methods) and total carbon (TC) and the gradual increase of bulk density (BD) from 1 m depth to the surface of the soil is due to marsh fragmentation (**Figure 9.3d**; **Supplementary Methods**). As the continuous marsh platform begins to erode from all edges, the interior of the marsh becomes closer to the shoreline. Consequently, fine-grained sediments present in the bay are transported into the marsh during storm events.[46, 47]. The soil at 40 cm depth formed under hypersaline condition also shows the most unique chemical characteristics that violate the trends on the molecular level. It shows the highest content of aliphatics and polymethylene, and the polymethylene has a unique structure that is rich in the carbon site resonating at 34 ppm (named type-a polymethylene; the crystalline domain in *all-trans* conformation). The lignin amount is low but contains a high level of S-monolignols with a high degree of methoxy substitutions. Cellulose crystallinity also becomes low: only 10% of glucan chains are now in a crystalline interior environment, while the remaining majority are disordered.

9.3.6 Paradigm of Wetland POM Preservation and Carbon Sequestration

It is imperative to understand the connection between the chemical stability of wetland SOC and its carbon structure, given wetlands' crucial role in global carbon stocks and their ability to sequester more carbon per unit area compared to other soil types[1, 6]. It is also crucial to differentiate between molecular and carbon speciation, microbial transformation, and preservation[17, 18] when considering SOC persistence as sea level continues to rise globally[8, 12]. In this pursuit, ^{13}C ssNMR spectroscopy[48-50], along with the sensitivity enhancement yielded by

DNP[51, 52], has been introduced to minimize the biases introduced by the extraction, solubilization, and relaxation present in liquid state NMR (**Supplementary Discussion**), while leveraging the detailed molecular view NMR spectroscopy allows, especially multidimensional techniques. This technique has allowed new insights into SOC sequestration and the importance of preservation for organic rich and POM-dominated blue carbon systems.

This study reveals robust preservation of the polymeric assembly in the top 10 cm of soil echoing the core structure of plant parent materials, evident in the comparable interior and surface cellulose signals, the preservation of multiple forms of cellulose, xylan conformers, and both the S and G monolignols in lignin (**Figure 9.1c, e, f**). This result can be attributed to the tight packing of some lignin and carbohydrate components, which could be induced by the decay of bulky cell wall cellulose and soft matrix polysaccharides, as evidenced by lower content of these moieties within the whole soil (**Figure 9.2a**). Further evidence for decay taking place in parallel with preservation is the absence of aromatic amino acids within the SOM (**Figure 9.2c**). This new insight allows for a refocusing on the concept of molecular preservation in the form of the conservation of the structural core of parent biopolymers, i.e., recalcitrance, as an important component of carbon storage, especially for high SOM systems, such as we find in wetlands, and counters the concept that free POM, as a whole, is a less than stable form of carbon[53, 54].

The preservation of recalcitrant lignocellulosic domains in soil POM involves maintaining both the molecular structure and supramolecular assembly of participating biopolymers (**Figure 9.3c**). This concurs with the biomolecular transformation of more-accessible molecules on the millennium timescale. The rapid decay of carbohydrates can be explained by their natural preferential utilization by microbes over other molecules, such as aromatic compounds, as both an energy and a nutrient source under the anaerobic soil conditions[53]. The better preservation of

aromatics over polymethylene is likely related to the reduced soil conditions. While soil microbes can produce extracellular compounds such as phenol oxidase, these metalloenzyme require oxygen to oxidize phenol compounds. Hence, in the anaerobic wetland soil profile, these compounds are stable[11]. Additionally, it has been found that high phenolic compounds strongly inhibit hydrolases further muting microbial decomposition of SOM[55], thus we posit that to a limited extent additional aromatic moieties are synthesized by biotic/abiotic processing of the loosely associated lignocellulose[56-58].

The dynamics of deltaic systems on C sequestration were revealed through quantitative ^{13}C NMR data, with major transformations found to occur within the SOC pool over a millennium (**Figure 9.3a, d**). The first important takeaway is that the conditions of the wetland under which the SOC is initially preserved play a major role in the decomposition of the SOC pool. This is evidenced by an increased preservation of the carbohydrate when the wetland transitioned from freshwater, to brackish, to a saltwater-dominated wetland (**Figure 9.3a**); the same general trend can also be seen regarding polymethylene. The relative proportion of preservation only changed with time with the change in the depositional environment, highlighting the environmental controls on plant species as the dominant factor. The initial transition from a freshwater to brackish water wetland was due to geological influences in the form of lobe transition (Lafourche to Plaquemine lobe and Plaquemine to Balize lobe, respectively), while the transition from a brackish water to a salt marsh was influenced by both a geological lobe relocation and levee construction, starving the wetlands of freshwater and sediment inputs[42].

9.4 Methods and Materials

9.4.1 Collection of Soil Material

Soil cores (2 m in length) were extracted with a polycarbonate core tube from a brackish

Spartina alterniflora-dominated island in Barataria Bay, Louisiana, USA (GPS coordinates: 29.44358, -89.899722). Two cores were extracted at different distances (1 m and 2 m, respectively) from the shoreline of the island (**Figure S9.1d**). The extracted materials were divided into 10 cm sections based on depth. The samples were stored on ice during transportation and then kept at 4 $^{\circ}$C for storage until analyzed.

9.4.2 Hydrofluoric Acid Treatment

Visible plant matter was removed from the dried soil samples. Each sample was ground with a mortar and pestle set until the material can pass through a 125 μm sieve. Around 600 mg of ground material was transferred into a 15 mL centrifuge tube, and 10 mL of 2% HF solution was added. The tube was capped and turned end-over-end in a rotary mixer throughout for 9 different time intervals in the following sequence: five 1 h intervals then a 16 h interval followed by two 24 intervals and a finally 72 h interval. In between these intervals, the tubes were placed into a benchtop centrifuge and spun at 2000 rpm for 20 min at room temperature. After centrifugation, the 2% HF solution was decanted and replaced with freshly prepared 2% HF solution. The soil samples were then vacuum filtered with 18 MΩ water for three times to remove the excess HF and freeze-dried for 24 h. This protocol was modified from a previously reported method[26]. Comparison of soil materials before and after HF treatment is shown in **Figures 9.1a** and **S9.8a**.

9.4.3 Solid-State NMR Spectroscopy

For each soil sample, 95-105 mg of HF-treated material were packed into a 4-mm zirconium rotor and measured on a Bruker Avance 400 MHz (9.4 Tesla) NMR spectrometer. All experiments were conducted using a 4 mm probe under 14 kHz MAS at 298 K. 1D quantitative spectra were measured using the MultiCP pulse sequence[41], with 11 CP blocks applied. Each CP block used 1.1 ms contact time, with a delay of 0.6 s between blocks. The acquisition time was set

to 25 ms, and the recycle delay was 1 s. For each sample, 16,384 scans were recorded within 35 h. The field strengths of the radiofrequency pulses were 71.4 kHz for both ^{13}C and ^{1}H hard pulses, and 62.5 kHz for ^{1}H decoupling. The ^{13}C chemical shifts were externally referenced to the tetramethylsilane (TMS) scale by calibrating the adamantane CH_2 peak to 38.48 ppm. In this work, all ssNMR and DNP spectra were collected using the software Topspin 4.0 and analyzed in Topspin 4.1 version. Graphs were plotted using Origin Pro 2019b software and Adobe Illustrator CC Cs6 V16.0.0.

To analyze the content of different carbon pools, deconvolution was performed on the 1D quantitative MultiCP ^{13}C spectra using DMfit[59] (20200306 version) following the positions of the peaks resolved from 2D DNP spectra, as detailed in **Figure S9.6** and **Table S9.1**. This allowed us to convert peak intensities into carbon percentages for different structural motifs including carbohydrates, aromatic, carbonyl, and aliphatic components, as well as the ratios of different carbon sites within each category (**Table S9.2**).

1D rotor-synchronized non-quaternary suppression (NQS) spectra were collected under 14 kHz to identify quaternary carbons[60]. Signals from the protonated carbons were dephased using two delays (30 μs × 2) without heteronuclear decoupling. The CP contact time was 2 ms. The acquisition time and the recycle delay was set to 41 ms and 2 s, respectively. In addition, 1D conventional ^{13}C CP spectra were collected to compare with NQS spectra, with identical experimental parameters. The NQS and CP spectra are shown in **Figure S9.3.**

9.4.4 Preparation of Soil and Plant Samples for MAS-DNP

A stock solution, which is often referred as the DNP juice was prepared using a mixture of D_2O and H_2O (90:10 Vol%) and 10 mM AsymPolPOK biradical (Catalogue# C015P01, CortecNet)[61]. Another two stock solutions were also prepared with the same radical concentration

but using different solvents of d_6-DMSO/D_2O/H_2O (10/80/10 Vol%) and d_6-DMSO/H_2O (90/10 Vol%). The D_2O (Catalogue# DLM-4DR-PK) and d_6-DMSO (Catalogue# DLM-10TC-PK) were from Cambridge Isotope Laboratories. The details parameters of DNP juice composition used for each sample and the setup parameters of all experiments were listed in **Table S9.3**.

The stock solutions were mixed with three types of materials, including HF-treated and non-treated soil as well as plant materials. Around 50 mg of HF-treated soil material was impregnated in 150 μL of the stock solution and vortexed briefly. The mixture was ground mildly for 20 min using a mortar and pestle to allow the radicals to penetrate the porous components of the soil. 30 mg of the final material were then packed into a 3.2 mm sapphire rotor for measurement. For comparison, the two plant samples (*Spartina alterniflora*) collected from the edge of the island (on top of the soil extraction site) and 30 m inland were also processed for MAS-DNP measurement. Around 30 mg of each plant sample was subjected to the same protocol described above to mix with 10 mM AsymPolPOK. For nontreated soil samples, the protocol was modified regarding the concentration of the bi-radical, which has increased to 30 mM to gain more enhancement. The DNP enhancements and Electron paramagnetic resonance (EPR) spectra (EMX Nano benchtop EPR) measured on the plant and soil samples were shown in **Figures S9.4** and **S9.8**. The EPR spectra were plotted by MATLAB R2020a with a toolbox Easyspin (6.0.0). The evaluation of the inhomogeneity was explained in **Supplementary Methods**.

9.4.5 2D $^{13}C/^{1}H$-^{13}C Correlation Experiments Enabled by MAS-DNP

In unlabeled samples, the natural-abundance of ^{13}C isotope is very low (1.1%), and the probability of observing connectivity between two carbon-13 nuclei in a 2D ^{13}C-^{13}C correlation spectrum is inhibitory (0.01%). To obtain sufficient sensitivity for measuring 2D correlation experiments[51], the soil and plant samples were measured on a Bruker 600 MHz/395 GHz MAS-

DNP system at National High Magnetic Field Laboratory, with the microwave irradiation power set to 12 W. The sample temperature was 104 K and 100 K when the microwave was on and off, respectively. The DNP buildup time was 1.3-4.5 s for all the MAS-DNP samples, including the HF-treated and untreated soil samples as well as the plant materials collected 30 m inland and at the edge of the island. Recycle delays were typically set to be 1.3-fold of the DNP buildup time constant for each sample. 1D ^{13}C CP experiments were measured with and without microwave irradiation under 8 kHz for soil and 10.5 kHz for plant samples, with the CP contact time set to 1 ms. The experimental parameters for all 1D and 2D NMR and MAS-DNP experiments are documented in **Table S9.3**.

2D ^{1}H-^{13}C HETCOR experiments were carried out under 8 kHz or 10.5 kHz MAS frequencies. ^{1}H-^{1}H homonuclear decoupling was achieved using either Phase-Modulated Lee–Goldburg (PMLG)[62] or Frequency-Switched Lee-Goldberg (FSLG) sequence[63] with a ^{1}H transverse field strength of 100 kHz, corresponding to an effective field strength of 122 kHz. To vary the range of detection between the proton and carbon sites, ^{1}H magnetization was transferred to ^{13}C using a Hartmann-Hahn (HH) CP block with a variable length, with 0.1 ms for primarily one-bond correlations, 0.5 ms for intermediate range of correlations, and 1.0 ms for long range correlations.

2D ^{13}C-^{13}C correlation experiments were carried out using the refocused INADEQUATE scheme[64]. The experiment was dipolar-based, using the broadband dipolar recoupling SPC5 sequence[65]. The MAS frequencies were 10.5 kHz for the HF-treated soil sample 1 and inland plants and changed to 8 kHz for the plant samples collected at the island edge. For the direct dimension (ω2), the acquisition time was 17 ms for all soil and plant samples. The acquisition time of the indirect dimension (ω1) was 2.7 ms and 1.7 ms for soil and plants, respectively. The indirect

dimensions of the spectra were set to 200 ppm (50 - 250 ppm) to effectively cover the double-quantum chemical shifts of carbohydrate and aromatic polymers. For each sample, 100 increments were collected for the indirect dimension. 320 scans were collected for the soil sample in 16 h, and 160 scans were collected for each of the two plant samples, with experimental time of 13 h and 23 h for the plants on the edge and inland, respectively. To rapidly identify the key carbohydrate components in soil, a probability map was built by extracting 412 datasets of plant carbohydrates from the Complex Carbohydrate Magnetic Resonance Database[36] following a recently reported protocol[66]. All ^{13}C and ^{1}H chemical shifts of identified polymers are documented in **Table S9.4**.

9.4.6 ^{14}C Dating

Prior to ^{14}C dating, the soil sample was pretreated with an acid/alkali/acid solution to avoid potential effect of the secondary carbon components (roots, bacteria) on the determined age of the sample[43, 44]. The decayed plants in the soil were used for ^{14}C dating, which was calibrated to radiocarbon age (years Before Present, yBP) and calendar years (cal AD). The analysis was performed using BetaCal 3.21, INTCAL13 database, and high probability density range method (HPD). The dataset is summarized in **Table S9.5**.

9.4.7 Bulk Density (BD) and Loss-On-Ignition (LOI)

The BD was determined by drying the soil at 60 °C for 24 h in a muffle furnace and then calculated as oven-dry wight per unit volume at field moisture capacity.[9] To determine the LOI, the dried material was ground with a mortar and pestle, and placed into a muffle furnace at 550 °C for 4 h. The mass difference before and after the combustion was divided by the original dry mass to get the percentage value of LOI ratio[13], which represents the relative fraction of organic matter in the sample. The results of these bulk property measurements were detailed in **Supplementary Methods**.

9.4.8 Total Carbon Percentage

The dried sample was ground using a mortar and pestle and sieved with 125 μm sieve to ensure equal particle size. 10 mg of soil were weighed into ceramic crucibles, which were placed into a total organic carbon analyzer (Shimadzu TOC SSM-5000A) to analyze the content of total carbon. Information on the physiochemical property is documented in **Table S9.5**.

9.5 Conclusions and Perspectives

On aggregate, SOC sequestration in the studied coastal wetland is a combination of molecular—including biomolecular—preservation, recalcitrant carbon, and carbon stabilization through dynamic carbon speciation (biological carbon turnover). The hydrogeomorphic setting changed over the 1000 years during which this sequestration has taken place, transitioning from an active freshwater delta to an abandoned freshwater delta lobe, then to a brackish and, eventually, a salt marsh system, as sea level has continued to rise. Despite these drastic surface changes, preservation, via tighter packing, of parent biopolymers has been consistent over time. A new framework of terminology can be derived, in which preservation can be viewed as molecular preservation and sequestration can be viewed as carbon storage regardless of speciation, with preservation being a subcategory of sequestration. This study provides strong evidence for giving equal weight to POM, just as what has been done for MAOM in regard to global SOC management[19], with POM being the major focus for organic soils, such as wetlands which contain ~1/3 of the planet's SOC. Recalcitrance should also be a major part of the focus within the preservation of high organic soil, especially as this preserved SOC becomes quickly processed and converted to greenhouse gases when exposed to highly oxygenated water due to erosion[12, 13, 67]. Therefore, POM and molecular recalcitrance, including biopolymeric structures, are important and may become the main drivers in SOC sequestration for about 1/3 or more of the planet's soil

organic carbon pool.

9.6 Acknowledgments

This research was supported by the U.S. Department of Energy under the grant no. DE-SC0021210 to T.W. and the National Science Foundation Chemical Oceanography grants OCE-1636052 and OCE-2054935 to J.R.W. and R.L.C. The National High Magnetic Field Laboratory is supported by the National Science Foundation through NSF/DMR-1644779 and DMR-2128556 and the State of Florida. The MAS-DNP system at NHMFL is funded in part by NIH P41 GM122698 and NIH RM1-GM148766. The authors thank Dr. Alex Kirui for initial data analysis, and M.P. Hayes and Benjamin Haywood for assistance with sample collection and bulk property measurements.

REFERENCES

1. Nahlik, A. M.; Fennessy, M. S., Carbon Storage in US Wetlands. *Nat. Commun.* **2016,** *7*, 13835.

2. Beillouin, D.; Corbeels, M.; Demenois, J.; Berre, D.; Boyer, A.; Fallot, A.; Feder, F.; Cardinael, R., A global meta-analysis of soil organic carbon in the Anthropocene. *Nat. Commun.* **2023,** *14*, 3700.

3. Lal, R., Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. *Science* **2004,** *304*, 1623-1627.

4. Georgiou, K.; Jackson, R. B.; VVinduskova, O.; Abramoff, R. Z.; Ahlstrom, A.; Feng, W.; Harden, J. W.; Pellegrini, A. F.; Polley, H. W.; Soong, J. L.; Riley, W. J.; Tom, M. S., Global stocks and capacity of mineral-associated soil organic carbon. *Nat. Commun.* **2022,** *13*, 3797.

5. Ruehr, S.; Keenan, T. F.; Williams, C.; Zhou, Y.; Lu, X.; Bastos, A.; Canadell, J. G.; Prentice, I. C.; Sitch, S.; Terrer, C., Evidence and attribution of the enhanced land carbon sink. *Nat. Rev. Earth Environ.* **2023,** *4*, 518-534.

6. Murray, B.; Pendleton, L.; Jenkins, W.; Sifleet, S., Green Payments for Blue Carbon: Economic Incentives for Protecting Threatened Coastal Habitats Nicholas Institute for Environmental Policy Solutions, Duke University, Durham, 2011.

7. Spivak, A. C.; Sanderman, J.; Bowen, J. L.; Canuel, E. A.; Hopkinson, C. S., Global-change controls on soil-carbon accumulation and loss in coastal vegetated ecosystems. *Nat. Geosci.* **2019,** *12*, 685-692.

8. Wang, F.; Lu, X.; Sanders, C.; Tang, J., Tidal wetland resilience to sea level rise increases their carbon sequestration capacity in United States. *Nat. Commun.* **2019,** *10*, 5434.

9. Duarte, C. M.; Losada, I. J.; Hendriks, I. E.; Mazarrasa, I.; Marbà, N., The role of coastal plant communities for climate change mitigation and adaptation. *Nat. Clim. Chang.* **2013,** *3*, 961–968.

10. Poffenbarger, H. J.; Needelman, B. A.; Megonigal, J. P., Salinity Influence on Methane Emissions from Tidal Marshes. *Wetlands* **2011,** *31*, 831–842.

11. Reddy, K. R.; DeLaune, R. D.; Inglett, P. W., *Biogeochemistry of Wetlands: Science and Application*. CRC Press: Boca Raton, 2008.

12. Haywood, B. J.; White, J. R.; Cook, R. L., Investigation of an early season river flood pulse: Carbon cycling in a subtropical estuary. *Sci. Total Environ.* **2018,** *635*, 867-877.

13. Haywood, B. J.; Hayes, M. P.; White, J. R.; Cook, R. L., Potential fate of wetland soil carbon in a deltaic coastal wetland subjected to high relative sea level rise. *Sci. Total Environ.* **2020,** *711*, 135185.

14. Pendleton, L.; Donato, D. C.; Murray, B. C.; Crooks, S.; Jenkins, W. A.; Sifleet, S.; Craft,

C.; Fourqurean, J. W.; Kauffman, J. B.; Marba, N.; Megonigal, P.; Pidgeon, E.; Herr, D.; Gordon, D.; Baldera, A., Estimating Global "Blue Carbon" Emissions from Conversion and Degradation of Vegetated Coastal Ecosystems. *PLOS One* **2012,** *7*, e43542.

15. Moomaw, W. R.; Chmura, G. L.; Davies, G. T.; Finlayson, C. M.; Middleton, B. A.; Natali, S. M.; Perry, J. E.; Roulet, N.; Sutton-Grier, A. E., Wetlands In a Changing Climate: Science, Policy and Management. *Wetlands* **2018,** *38*, 183-205.

16. Stevenson, F. J., *Humus Chemistry: Genesis, Composition, Reactions*. John Wiley & Sons: New York, NY, 1994.

17. Schmidt, M. W. I.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I. A.; Kleber, M.; Kogel-Knabner, I. K.; Lehmann, J.; Manning, D. A. C.; Nannipieri, P.; Rasse, D. P.; Weiner, S.; Trumbore, S. E., Persistence of soil organic matter as an ecosystem property. *Nature* **2011,** *478*, 49-56.

18. Lehmann, J.; Kleber, M., The Contentious Nature of Soil Organic Matter. *Nature* **2015,** *528*, 60-68.

19. Angst, G.; Mueller, K. E.; Castellano, M. J.; Vogel, C.; Wiesmeier, M.; Mueller, C. W., Unlocking complex soil systems as carbon sinks: multi-pool management as the key. *Nat. Commun.* **2023,** *14*, 2967.

20. Lehmann, J.; Hansel, C. M.; Kaiser, C.; Kleber, M.; Maher, K.; Manzoni, S.; Nunan, N.; Reichstein, M.; Schimel, J. P.; Torn, M. S.; Wieder, W. R.; Kogel-Knabner, I. K., Persistence of soil organic carbon caused by functional complexity. *Nat. Geosci.* **2020,** *13*, 529-534.

21. Kleber, M.; Bourg, I. C.; Coward, E. K.; Hansel, C. M.; Myneni, S. C. B.; Nunan, N., Dynamic interactions at the mineral–organic matter interface. *Nat. Rev. Earth Environ.* **2021,** *2*, 402-421.

22. Hall, S. J.; Ye, C.; Weintraub, S. R.; Hockaday, W. C., Molecular trade-offs in soil organic carbon composition at continental scale. *Nat. Geosci.* **2020,** *13*, 687-692.

23. Boesch, D. F.; Jocelyn, M. N.; Mehta, A. J.; Morris, J. T.; Nuttle, W. K.; Simenstad, C. A.; Swift, D. J. P., Scientific Assessment of Coastal Wetland Loss, Restoration and Management in Louisiana. *J. Coast. Res.* **1994,** *20*, 1–103.

24. Jankowski, K. L.; Törnqvist, T. E.; Fernandes, A. M., Vulnerability of Louisiana's Coastal Wetlands to Present-Day Rates of Relative Sea-Level Rise. *Nat. Commun.* **2017,** *8*, 14792.

25. Schmidt, M. W. I.; Knicker, H.; Hatcher, P. G.; Kogel-Knabner, I. K., Improvement of ^{13}C and ^{15}N CPMAS NMR Spectroscopy of Bulk Soils, Particle Size Fractions and Organic Material by Treatment with 10% Hydrofluoric Acid. *Eur. J. Soil Sci.* **1997,** *48*, 319-328.

26. Skjemstad, J.; Clarke, P.; Taylor, J.; Oades, J.; Newman, R., The Removal of Magnetic Materials from Surface Soils - A Solid State ^{13}C CP/MAS NMR Study. *Soil Res.* **1994,** *32*, 1215-1229.

27. Salati, S.; Adani, F.; Cosentino, C.; Torri, G., Studying Soil Organic Matter Using 13C CP-MAS NMR: The Effect of Soil Chemical Pre-Treatments on Spectra Quality and Representativity. *Chemosphere* **2008,** *70*, 2092-2098.

28. Phyo, P.; Wang, T.; Yang, Y.; O'Neill, H.; Hong, M., Direct determination of hydroxymethyl conformations of plant cell wall cellulose using ^{1}H polarization transfer solid-state NMR. *Biomacromolecules* **2018,** *19* (5), 1485-1497.

29. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules* **2016,** *17* (6), 2210-2222.

30. Kang, X.; Kirui, A.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 347.

31. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

32. Busse-Wicher, M.; Gomes, T. C. F.; Tryfona, T.; Nikolovski, N.; Stott, K.; Grantham, N. J.; Bolam, D. N.; Skaf, M. S.; Dupree, P., The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of Arabidopsis thaliana *Plant. J.* **2014,** *79*, 492-506.

33. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

34. Grantham, N. J.; Wurman-Rodrich, J.; Terret, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **2017**, 859-865.

35. Hu, W.-G.; Mao, J.; Xing, B.; Schmidt-Rohr, K., Poly(Methylene) Crystallites in Humic Substances Detected by Nuclear Magnetic Resonance. *Environ. Sci. Technol.* **1999,** *34*, 530-534.

36. Kang, X.; Zhao, W.; Widanage, M. C. D.; Kirui, A.; Ozdenvar, U.; Wang, T., CCMRD: a solid-state NMR database for complex carbohydrates. *J. Biomol. NMR* **2020,** *74* (4), 239-245.

37. Lattao, C.; Birdwell, J.; Wang, J. J.; Cook, R. L., Studying Organic Matter Molecular Assemblage within a Whole Organic Soil by Nuclear Magnetic Resonance. *J. Environ. Qual.* **2008,** *37*, 1501-1509.

38. Yan, J.; Wang, L.; Hu, Y.; Tsang, Y. F.; Zhang, Y.; Wu, J.; Fu, X.; Sun, Y., Composition Selects Different Soil Microbial Structures and in Turn Drives Different Litter Decomposition Pattern and Soil Carbon Sequestration Capability. *Geoderma* **2018,** *319*, 194-203.

39. Lovley, D. R.; Coates, J. D.; Blunt-Harris, E. L.; Phillips, E. J.; Woodward, J. C., Humic Substances as Electron Acceptors for Microbial Respiration. *Nature* **1996,** *382*, 445-448.

40. Kögel-Knabner, I.; de Leeuw, J. W.; Hatcher, P. G., Nature and Distribution of Alkyl Carbon in Forest Soil Profiles: Implications for the Origin and Humification of Aliphatic Biomacromolecules. *Sci. Total Environ.* **1992,** *117-118*, 175-185.

41. Johnson, R. L.; Schmidt-Rohr, K., Quantitative solid-state ^{13}C NMR with signal enhancement by multiple cross polarization. *J. Magn. Reson.* **2014,** *239*, 44-49.

42. Coleman, J. M.; Roberts, H. H.; Stone, G. W., Mississippi River Delta: an Overview. *J. Coast. Res.* **1998,** *14*, 698-716.

43. Sapkota, Y.; White, J. R., Long-term fate of rapidly eroding carbon stock soil profiles in coastal wetlands. *Sci. Total Environ.* **2021,** *753*, 141913.

44. Sapkota, Y.; White, J. R., Marsh edge erosion and associated carbon dynamics in coastal Louisiana: A proxy for future wetland-dominated coastlines world-wide. *Estuar. Coast. Shelf Sci.* **2019,** *226*, 106289.

45. Alexander, J. S.; Wilson, R. C.; Green, W. R. *A brief history and summary of the effects of river engineering and dams on the Mississippi River system and delta: U.S. Geological Survey Circular 1375*; 2012; p 43 p.

46. DeLaune, R. D.; Nyman, J. A.; Patrick, W. H., Peat Collapse, Ponding and Wetland Loss in a Rapidly Submerging Coastal Marsh. *J. Coast. Res.* **1994,** *10*, 1021-1030.

47. Odum, W. E., Comparative Ecology of Tidal Freshwater and Salt Marshes. *Annu. Rev. Ecol. Syst.* **1988,** *19*, 147-176.

48. Simpson, A. J.; McNally, D. J.; Simpson, M. J., NMR spectroscopy in environmental research: From molecular interactions to global processes. *Prog. Nucl. Magn. Reson. Spectr.* **2011,** *58*, 97-175.

49. Mao, J.; Cao, X.; Olk, D. C.; Chu, W.; Schmidt-Rohr, K., Advanced solid-state NMR organic matter. *Prog. Nucl. Magn. Reson. Spectr.* **2017,** *100*, 17-51.

50. Cook, R. L., Coupling NMR to NOM. *Anal. Bioanal. Chem.* **2004,** *378*, 1484-1503.

51. Marker, K.; Paul, S.; Fernandez-de-Alba, C.; Lee, D.; Mouesca, J. M.; Hediger, S.; De Paëpe, G., Welcoming natural isotopic abundance in solid-state NMR: probing π-stacking and supramolecular structure of organic nanoassemblies using DNP. *Chem. Sci.* **2017,** *8*, 974-987.

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

53. Angst, G.; Mueller, K. E.; Nierop, K. G. J.; Simpson, M. J., Plant- or microbial-derived? A

review on the molecular composition of stabilized soil organic matter. *Soil Biol. Biochem.* **2021,** *156*, 108189.

54. Lavallee, J. M.; Soong, J. L.; Cotrufo, M. F., Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. *Glob. Change Biol.* **2020,** *26*, 261-273.

55. Freeman, C.; Ostle, N.; Kang, H., An enzymic 'latch' on a global carbon store. *Nature* **2001,** *409*, 149.

56. Gerke, J., Concepts and misconceptions of humic substances as the stable part of soil organic matter: A review. *Agronomy* **2018,** *2018*, 18258.

57. Cao, X.; Schmidt-Rohr, K., Abundant Nonprotonated Aromatic and Oxygen-Bonded Carbons Make Humic Substances Distinct from Biopolymers. *Environ. Sci. Technol. Lett.* **2018,** *5*, 476-480.

58. De Nobili, M.; Bravo, C.; Chen, Y., The spontaneous secondary synthesis of soil organic matter components: A critical examination of the soil continuum model theory. *App. Soil Ecol.* **2020,** *154*, 103655.

59. Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calvé, S.; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z.; Hoatson, G., Modelling one and two-dimensional solid-state NMR spectra. *Magn. Reson. Chem.* **2002,** *40*, 70-76.

60. Opella, S. J.; Frey, M. H., Selection of nonprotonated carbon resonances in solid-state nuclear magnetic resonance. *J. Am. Chem. Soc.* **1979,** *101*, 5854-5856.

61. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

62. Vinogradov, E.; Madhu, P. K.; Vega, S., High-resolution proton solid-state NMR spectroscopy by phase-modulated Lee–Goldburg experiment. *Chem. Phys. Lett.* **1999,** *314*, 443-450.

63. Bielecki, A.; Kolbert, A. C.; De Groot, H.; Griffin, R. G.; Levitt, M. H., Frequency-Switched Lee—Goldburg Sequences in Solids. *Adv. Magn. Reson.* **1990,** *14*, 111-124.

64. Lesage, A.; Bardet, M.; Emsley, L., Through-bond carbon-carbon connectivities in disordered solids by NMR. *J. Am. Chem. Soc.* **1999,** *188*, 10987-10993.

65. Hohwy, M.; Jakobsen, H. J.; Eden, M.; Levitt, M. H.; Nielsen, N. C., Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *J. Chem. Phys.* **1998,** *108*, 2686-2694.

66. Zhao, W.; Debnath, D.; Gautam, I.; Fernando, L. D.; Wang, T., Charting the Solid-State

NMR Signals of Polysaccharides: A Database-Driven Roadmap. *Magn. Reson. Chem.* **2023**, in press.

67. Steinmuller, H. E.; Hayes, M. P.; Hurst, N. R.; Sapkota, Y.; Cook, R. L.; White, J. R.; Xue, Z.; Chambers, L. G., Does edge erosion alter coastal wetland soil properties? A multi-method biogeochemical study. *Catena* **202,** *187*, 104373.

68. Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced NMR: The AsymPol Family. *J. Am. Chem. Soc.* **2018,** *140* (35), 11013-11019.

69. Viger-Gravel, J.; Lan, W.; Pinon, A.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of Pretreated Wood Fibers Using Dynamic Nuclear Polarization. *J. Phys. Chem. C.* **2019,** *123*, 30407-30415.

70. Jepsen, R.; Roberts, J.; Lick, W., Effects of Bulk Density on Sediment Erosion Rates. *Wat. Air Soil Pollut.* **1997,** *99*, 21-31.

71. Craft, C., Freshwater Input Structures Soil Properties, Vertical Accretion, and Nutrient Accumulation of Georgia and U.S Tidal Marshes. *Limnol. Oceanogr.* **2007,** *52*, 1220-1230.

72. DeLaune, R. D.; Nyman, J. A.; Patrick, W. H., Peat Collapse, Ponding and Wetland Loss in a Rapidly Submerging Coastal Marsh. *J. Coast. Res.* **1994,** *10*, 1021-1030.

73. Qu, W.; Li, J.; Han, G.; Wu, H.; Song, W.; Zhang, X., Effect of Salinity on the Decomposition of Soil Organic Carbon in a Tidal Wetland. *J. Soils Sediments* **2018,** *19*, 487-496.

74. Remusat, L.; Hatton, P. J.; Nico, P. S.; Zeller, B.; Kleber, M.; Derrien, D., NanoSIMS Study of Organic Matter Associated with Soil Aggregates: Advantages, Limitations, and Combination with STXM. *Environ. Sci. Technol.* **2012,** *46*, 3943-3949.

75. Lehmann, J.; Solomon, D.; Kinyangi, J.; Dathe, L.; Wirick, S.; Jacobsen, C., Spatial complexity of soil organic matter forms at nanometre scales. *Nat. Geosci.* **2008,** *1*, 238-242.

76. Lattao, C.; Birdwell, J.; Wang, J. J.; Cook, R. L., Studying Organic Matter Molecular Assemblage within a Whole Organic Soil by Nuclear Magnetic Resonance. *J. Environ. Qual.* **2008,** *37*, 1501-1509.

77. Masoom, H.; Courtier-Murias, D.; Farooq, H.; Soong, R.; Kelleher, B. P.; Zhang, C.; Maas, W. E.; Fey, M.; Kumar, R.; Monette, M.; Stronks, H. J.; Simpson, M. J.; Simpson, A. J., Soil Organic Matter in Its Native State: Unravelling the Most Complex Biomaterial on Earth. *Environ. Sci. Technol.* **2016,** *50*, 1670-1680.

78. Kelleher, B. P.; Simpson, A. J., Humic Substances in Soils: Are They Really Chemically Distinct? *Environ. Sci. Technol.* **2006,** *40*, 4605-4611.

79. Schmidt, M. W. I.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I. A.;

Kleber, M.; Kogel-Knabner, I. K.; Lehmann, J.; Manning, D. A. C.; Nannipieri, P.; Rasse, D. P.; Weiner, S.; Trumbore, S. E., Persistence of soil organic matter as an ecosystem property. *Nature* **2011,** *478*, 49-56.

80. Simon, C.; Duhrkop, K.; Petras, D.; Roth, V. N.; Bocker, S.; Dorrestein, P. C.; Gleixner, G., Mass Difference Matching Unfolds Hidden Molecular Structures of Dissolved Organic Matter. *Environ. Sci. Technol.* **2022,** *56*, 11027-11040.

81. Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M., Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008,** *80*, 8908-8919.

APPENDIX

S9.1 Supplementary Methods

Effect of HF treatment on soil. The application of 2% HF solution efficiently depleted iron and mineral component in soil, allowing for a substantially higher concentration of SOM within a MAS rotor. Comparative analysis of ssNMR spectra showed no noticeable difference between HF-treated and untreated samples, whether at room temperature (**Figure 9.1a)** or at the cryogenic temperature used for DNP measurements (**Figure S9.8a**). Discrepancies were observed between MultiCP and DNP-enhanced CP spectra (**Figure 9.1a, b**), which were not caused by HF treatment but resulted from the application of higher magnetic fields and MAS frequencies, reducing the signals of aromatic motifs with significant chemical shift anisotropy (CSA).

Inhomogeneity of DNP in soil samples. For DNP measurements, soil materials were doped with 10 mM of a stable biradical named AsymPolPok[68]. This biradical is water-soluble and has a small dimension of 1-2 nm, enabling penetration in porous materials. When the electron polarization is transferred to ^{1}H in the DNP matrix, it relies on ^{1}H-^{1}H relayed transfer, ensuring efficient and homogenous polarization across a range of at least a hundred nanometers, as recently reported[69]. Typically, it is straightforward to achieve homogeneous polarization throughout the sample. However, within each soil sample, polarization levels varied among molecules and structural motifs. In the case of HF-treated sample 1, all carbohydrate signals exhibited a uniform 18-fold enhancement, whereas aromatic and aliphatic carbons experienced a comparatively lower enhancement of 7-9-fold (**Figure S9.8b**). Hence, the predominant portions of carbohydrates appear to be more accessible, which may partially explain their accelerated degradation rate through microbial processes in soil. Alternatively, carbohydrates may exhibit enhanced retention of solvent and biradicals, probably due to the rich polar functionalities present in carbohydrates.

Similar patterns were observed in non-treated soil sample 1 (**Figure S9.8c**) and non-treated samples #4 and #6 (**Figure S9.8e, f**), where carbohydrates consistently exhibited the highest enhancement, while other components displayed slightly lower enhancements. Remarkably, the most effective DNP performance was noted in non-treated soil sample 4, demonstrating a 33-fold enhancement for carbohydrates and a 23-fold enhancement for all other molecules (**Figure S9.8e**). Additionally, it is noteworthy that even in plant samples, polymethylene peaks exhibited minimal DNP enhancement (**Figure S9.3b, d**), indicating the presence of lipid polymers in self-aggregated domains that are challenging to penetrate, not only in soil but also in plants. This could also be due to faster ^{1}H relaxation, leading to inadequate preservation of hyperpolarization.

Bulk properties of wetland soil samples. During the Lafourche delta lobe formation, the Mississippi River ran through Barataria Bay and deposited silt and mineral matter into the surrounding wetland soils. The soil at 1.5-1.8 m depth formed during that period, exhibiting very low loss-on-ignition (LOI) ratios of 1.16-1.40 and total carbon (TC) percentage of 8.1-11.2%, but a high bulk density (BD) of 0.14-0.31 g/cm^3 (**Table S9.5**). High BD indicates a high concentrations of large mineral particles, with the soil expected to erode more slowly[70], while a low LOI ratio represent a low content of organic matter. When it came to the Plaquemines delta lobe, the Barataria Bay became a fresh marsh and expected to contain more organic matter[71-73]. Consistently, the soil at the depth of 1-1.5 m showed a high LOI ratio of 1.51-1.65, a high TC percentage of 10.7-14.0%, and a low BD of 0.11-0.18 g/cm^3. The geological and historical timestamps explain the observed changes in the chemical composition of soil organic matter. Across all geological stages listed here, the soil formed during the Lafourche delta lobe is featured with a high content of aromatic carbons and a low amount of carbohydrates. This trend has been flipped in the soil associated with the Balize delta lobe.

S9.2 Supplementary Discussion

S9.2.1 An Incomplete Molecular Characterization of SOM

Advances in analytical instrumentation, in terms of resolution, sensitivity, and less harsh sample pre-treatment, in concert with some highly insightful experimental design, has allowed for new data and insights to emerge and a rethinking of SOM preservation. While there is no argument that these new techniques have allowed a major leap forward in our understanding of SOM and the associated SOC fraction, care must be taken when extrapolating the data interpretation to SOM as a whole. In addition to the bias on upland/mineral soil, as described in the Introduction, two guiding principles are also useful: 1) granularity (speciation versus molecular) and 2) bias on extraction and small molecule and the need for complementary characterization techniques.

S9.2.2 Granularity

Methods that interrogate whole soils, to date, provide data at an elemental to moiety level, e.g., aliphatic (alkyl), carbohydrate (O-alkyl), aromatic (aryl), aromatic with oxygen functional groups (O-aryl), phenolic, carbonyl group, *etc*., but not at the molecular level, which identifies exact molecular structures. This challenge is a result of the complexity of SOM as well as the limited resolution of each individual technique applied to whole soils to date. Some of the key techniques include: Nano Secondary Ion Mass Spectrometry (NanoSIMS), near-edge X-ray spectromicroscopy, and NMR, including 2D methods[74-78]. Consequently, unlike in other research fields, while we are able to interrogate a system and gain a general moiety (structural categories) level understanding, we are unable to ask direct questions regarding the more informative and scalable molecular level.

S9.3.3 Need for Complementary Technique for Soil Characterization

Two techniques that have been able to overcome the granularity limit mentioned above are FT-ICR-MS and multidimensional liquid state NMR, each with their own limitations brought about by three biases. The *extraction bias* arises from the SOM being extracted and, since all too often, such extraction is incomplete[79], thus these techniques are blind to the non-extracted fraction of SOM. The *solubility bias* stems from the fact that not all SOM, even extracted SOM, is equally soluble, with the more soluble fraction being more highly weighted. The third bias is the *small molecule bias*. For liquid state NMR, this bias is a result of molecular rotation, which leads to differences in relaxation behavior, especially T_2 relaxation, with smaller molecules yielding sharper peaks and hence more intense signals compared to larger molecules that yield broader spectral peaks and tend to be lost due to the broad peak convolution for samples as complex as SOM extracts. On the FT-ICR-MS front[80, 81], this small molecule bias, all other things being equal, results in a small molecule ionization preference. This means that the high-resolution FT-ICR and liquid state NMR data are biased towards smaller molecules, with an understanding that the absence of evidence of larger molecules does not mean they are not there. In this context, DNP-enabled high-resolution ssNMR can serve as a complementary technique to provide information on the large macromolecules in soil materials.

Table S9.1 Spectral deconvolution parameters of quantitative ^{13}C spectra for HF-treated soil samples. The attributed chemical shifts (CSs) are labeled. The integral value (I), linewidth (L), and relative amplitude (A) are listed. Uncertain (U).

	Assign.	**Soil 1**	**Soil 2**	**Soil 3**	**Soil 4**	**Soil 5**	**Soil 6**	**Soil 7**
		CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]	CSs(ppm); I(%) [L(ppm), A(%)]
Carbonyl	CO	182.3; 0.28 [6.5, 0.2]	183.8; 0.46 [4.1, 0.5]	178.7; 1.89 [8.7, 1.2]	181.0; 1.17 [9.6, 0.7]	180.2; 0.92 [9.0, 0.6]	179.4; 1.63 [9.2, 1.0]	183.2; 0.66 [8.0, 0.5]
	CO	176.1; 3.06 [5.7, 2.8]	176.2; 3.75 [6.2, 3.0]	175.8; 2.78 [4.9, 3.0]	175.7; 3.42 [5.6, 3.6]	176.1; 2.70 [5.3, 3.0]	175.7; 2.57 [4.9, 3.0]	176.2; 3.00 [6.3, 3.0]
	CO	173.4; 3.29 [4.4, 3.9]	173.6; 3.43 [4.3, 3.9]	173.5; 3.40 [4.2, 4.3]	173.1; 3.78 [4.6, 4.5]	173.3; 4.92 [5.1, 5.7]	173.1; 3.60 [4.8, 4.4]	173.3; 4.21 [5.3, 5.1]
Aromatic	H4	161.5; 0.16 [2.4, 0.4]	162.8; 0.4 [2.9, 0.8]	161.6; 0.07 [1.5, 0.2]	161.5; 0.08 [2.3, 0.2]	161.1; 0.07 [1.5, 0.3]	162.7; 0.06 [1.1, 0.3]	162.3; 0.12 [1.6, 0.5]
	S3/5[b]	158.4; 0.23 [3.1, 0.4]	158.1; 0.69 [4.5, 0.7]	158.4; 0.36 [0.7, 2.9]	157.8; 0.61 [3.9, 0.9]	157.8; 0.66 [3.5, 1.1]	159.0; 0.43 [4.7, 0.5]	158.5; 0.43 [4.7, 0.6]
	S3/5[a]	154.1; 3.07 [5.2, 3.1]	154.2; 2.24 [4.4, 2.5]	154.2; 2.58 [4.8, 2.9]	153.6; 2.35 [5.0, 2.6]	153.9; 2.56 [5.2, 2.9]	154.4; 3.11 [5.5, 3.3]	154.4; 2.95 [5.5, 3.4]
	G3	148.7; 3.02 [6.5, 2.4]	148.9; 1.84 [4.6, 1.9]	149.4; 2.05 [4.7, 0.3]	148.4; 2.51 [5.8, 2.4]	148.5; 2.42 [6.0, 2.4]	148.8; 3.08 [6.4, 2.8]	148.3; 3.34 [7.2, 2.9]
	G4	144.4; 0.17 [2.0, 0.4]	146.0; 0.50 [4.0, 0.6]	146.6; 1.12 [5.3, 1.1]	144.5; 0.64 [6.8, 0.5]	144.6; 0.82 [5.6, 0.9]	145.0; 0.40 [4.2, 0.5]	141.8; 0.87 [9.1, 0.6]

Table S9.1 (cont'd)

	S1	136.6; 4.28 [7.9, 2.8]	137.4; 2.99 [7.5, 2.0]	136.2; 5.06 [9.2, 2.9]	136.1; 4.87 [9.0,3.0]	137.5; 3.04 [7.6, 2.4]	137.1; 4.24 [7.6, 3.2]	137.2; 3.48 [6.4, 3.4]
	G1/H1	131.7; 4.63 [6.8, 3.6]	132.6; 3.93 [7.5, 2.6]	130.7; 5.86 [7.9, 3.9]	130.9; 5.69 [7.6, 4.2]	133.4; 3.56 [7.0, 3.0]	133.0; 3.71 [6.3, 3.4]	132.6; 4.70 [6.6, 4.5]
	H2/6	127.8; 3.52 [6.7, 2.8]	128.9; 3.21 [6.5, 2.4]	125.8; 2.28 [6.7, 1.8]	126.5; 3.34 [7.4, 2.5]	129.6; 6.25 [7.4, 5.0]	129.4; 5.88 [7.1, 4.7]	128.7; 6.38 [7.4, 5.4]
	G6	123.4; 2.51 [6.4, 2.0]	123.9; 3.47 [7.9, 2.6]	121.4; 2.27 [6.8, 1.8]	122.3; 2.39 [6.7, 2.0]	123.8; 4.36 [8.4, 3.1]	123.8; 4.48 [7.8, 3.3]	123.1; 4.98 [8.2, 3.9]
	G5, H3/5	117.0; 4.52 [7.4, 3.4]	117.1; 4.01 [7.9, 2.6]	116.6; 3.52 [6.5, 3.0]	116.7; 3.98 [7.3, 3.2]	117.0; 4.42 [7.6, 3.6]	116.9; 5.91 [8.1, 4.4]	116.6; 4.16 [7.3, 3.8]
	G2	111.4; 1.95 [6.1, 1.7]	110.8; 1.90 [6.3, 1.5]	111.4; 1.93 [5.6, 1.8]	111.2; 1.84 [6.2, 1.6]	111.3; 1.58 [6.6, 1.4]	111.0; 1.6 [5.9, 1.6]	111.0; 2.17 [6.5, 2.1]
Carbohydrate	i/s1	105.9; 3.59 [3.9, 4.8]	105.7; 3.47 [4.3, 4.0]	106.0; 2.36 [3.7, 3.4]	105.3; 2.56 [4.4, 3.2]	105.5; 1.95 [4.4, 2.6]	106.1; 2.20 [4.0, 3.2]	105.2; 2.52 [4.7, 3.4]
	$Xn^{3f}1$	102.7; 0.86 [3.0, 1.5]	101.9; 1.17 [3.5, 1.6]	103.1; 0.90 [3.1, 1.6]	101.9; 0.71 [3.2, 1.2]	102.3; 0.83 [3.2, 1.5]	103.1; 0.71 [2.7, 1.5]	101.0; 0.08 [1.2, 0.4]
	i4	88.8; 0.46 [2.3, 1.3]	88.2; 0.07 [1.4, 0.3]	88.5; 0.12 [1.8, 0.4]	88.4; 0.1 [1.6, 0.4]	88.6; 0.01 [0.5, 0.1]	88.6; 0.03 [0.9, 0.2]	88.5; 0.03 [0.8, 0.3]
	s4	84.9; 1.14 [3.0, 2.0]	84.9; 0.62 [2.6, 1.2]	85.1; 0.67 [2.7, 1.3]	84.7; 0.61 [2.8, 1.2]	84.0; 0.82 [3.4, 1.4]	85.0; 0.81 [3.0, 1.6]	85.0; 0.73 [3.1, 1.5]
	$Xn^{2f}4$	82.4; 1.22 [4.0, 1.6]	82.6; 0.99 [3.7, 1.3]	82.4; 0.98 [3.7, 1.4]	82.2; 0.92 [3.6, 1.4]	80.2; 0.97 [4.2, 1.3]	81.7; 1.03 [4.2, 1.4]	81.8; 0.82 [4.3, 1.2]

Table S9.1 (cont'd)

	C2,3,5	75.9; 7.10 [5.0, 7.4]	76.0; 5.97 [5.4, 5.4]	76.0; 5.42 [5.2, 5.5]	75.9; 5.87 [5.7, 5.7]	75.8; 4.12 [5.3, 4.5]	76.2; 4.72 [5.4, 5.0]	76.2; 3.29 [5.3, 3.9]
	C2,3,5	73.0; 7.90 [5.3, 7.8]	72.9; 9.62 [6.7, 7.0]	72.9; 8.10 [6.2, 6.9]	72.4; 9.84 [6.9, 7.9]	72.1; 7.13 [6.8, 7.1]	72.7; 7.48 [6.6, 6.5]	72.8; 7.97 [7.4, 6.8]
	C6	63.8; 6.32 [7.4, 4.5]	63.5; 4.62 [6.7, 3.3]	63.2; 4.90 [6.7, 3.8]	62.6; 4.84 [7.2, 3.7]	62.5; 7.13 [7.1, 3.2]	63.1; 4.49 [7.3, 3.6]	62.7; 3.48 [6.9, 3.2]
Aliphatic	-OCH_3[b] or protein Cα	56.8; 6.43 [4.7, 7.1]	56.8; 5.34 [4.5, 5.7]	56.8; 6.55 [4.8, 7.2]	56.4; 5.97 [5.0, 6.6]	56.4; 5.61 [5.0, 6.6]	56.6; 7.08 [5.1, 8.0]	56.6; 5.95 [5.2, 7.2]
	-OCH_3[a] or protein Cα	51.1; 3.15 [8.2, 2.0]	51.4; 3.29 [8.3, 2.0]	50.8; 4.27 [8.5, 2.7]	50.7; 3.7 [7.6, 2.7]	51.0; 4.80 [8.4, 3.4]	50.7; 3.90 [8.4, 2.7]	51.0; 3.08 [7.3, 2.7]
	R_3CH & R_4C	43.1; 3.0 [7.6, 2.1]	43.1; 2.34 [6.4, 1.8]	43.1; 2.30 [5.6, 2.2]	43.2; 2.43 [6.2, 2.2]	42.8; 3.30 [7.2, 2.7]	43.3; 2.78 [6.6, 2.4]	43.9; 2.27 [7.8, 1.9]
	-CCH	38.6; 1.46 [4.8, 1.6]	39.3; 0.61 [3.2, 1.0]	39.2; 1.35 [4.2, 1.7]	38.6; 2.25 [5.9, 2.1]	38.2; 2.76 [6.5, 2.5]	38.9; 2.84 [6.8, 2.4]	39.1; 4.28 [8.9, 3.0]
	-$(CH_2)_n$-[a]	33.9; 3.52 [2.2, 8.5]	33.9; 6.45 [1.9, 16.8]	33.8; 4.45 [2.0, 12.0]	33.5; 3.57 [2.1, 9.3]	33.5; 2.97 [2.2, 7.7]	33.9; 2.57 [2.3, 6.5]	33.6; 2.61 [2.6, 6.4]
	-$(CH_2)_n$-[b]	30.9; 5.45 [3.4, 8.4]	31.0; 8.96 [3.7, 11.6]	31.0; 6.81 [3.6, 9.8]	30.6; 7.25 [3.7, 10.9]	30.6; 6.70 [4.0, 9.8]	31.0; 5.49 [3.7, 8.5]	30.8; 4.04 [3.4, 7.5]
	-$\underline{CH_2}$-CH=CH-	27.0; 4.66 [6.3, 3.9]	26.6; 4.31 [5.2, 4.0]	26.6; 3.89 [5.4, 3.8]	26.1; 4.44 [5.9, 4.2]	26.1; 3.73 [5.7, 3.9]	27.3; 2.98 [5.6, 3.1]	27.6; 3.52 [6.2, 3.6]
	Ac^{Me}	23.9; 1.27 [3.2, 2.1]	24.0; 1.93 [3.6, 2.6]	23.9; 1.23 [3.3, 2.0]	23.6; 0.83 [2.5, 1.8]	23.6; 1.55 [3.6, 2.5]	24.1; 2.66 [4.5, 3.4]	23.6; 4.00 [6.2, 4.1]

Table S9.1 (cont'd)

	$-CH_3$	20.6; 1.98 [5.3, 2.0]	20.4; 1.11 [3.5, 1.5]	20.5; 2.21 [5.5, 2.1]	20.3; 1.75 [4.7, 2.0]	20.0; 1.98 [5.2, 2.2]	20.4; 1.36 [4.4, 1.8]	19.3; 1.13 [5.1, 1.4]
	$-CH_3^U$	15.3; 2.54 [7.8, 1.7]	15.5; 3.65 [8.4, 2.1]	15.5; 2.21 [7.4, 1.6]	15.6; 2.15 [7.2, 1.7]	15.3; 2.56 [7.8, 1.9]	15.8; 3.01 [8.4, 2.1]	15.1; 2.88 [9.5, 1.9]

Table S9.2 Molecular composition of HF-treated soils based on spectral deconvolution results.

Soil #	Depth (cm)	Carbohyd	Aromatic	Aliphatic	CO	i4/s4 [a]	S/G [b]	[c]-$(CH_2)_n$-	[d]a/b	-OCH_3	Era
1	0-10	29.6%	29.0%	34.6%	6.8%	0.40	1.03	9.3	0.65	9.9%	2018 AD
2	40-50	27.3%	25.9%	39.0%	7.9%	0.11	1.25	15.8	0.72	8.9%	1963 AD
3	80-90	25.0%	28.9%	37.6%	8.6%	0.17	0.93	12	0.65	11.5%	1562 AD
4	100-110	26.4%	29.3%	35.6%	8.7%	0.17	0.94	11.2	0.49	10.0%	1169 AD
5	130-140	21.9%	31.3%	37.8%	9.0%	0.01	0.99	10.2	0.44	10.9%	
6	160-170	22.2%	34.0%	35.8%	8.1%	0.03	1.02	8.3	0.47	11.3%	945 AD
7	170-180	20.1%	35.7%	35.9%	8.4%	0.04	0.80	7.1	0.65	9.6%	

[a] The ratio of interior cellulose to surface cellulose. Using integral of i4 peak over integrals of s4 peak.

[b] Lignin aromatic ring S unit to G unit ratio. Using integrals of S3/5 peaks over integrals of G3/4 peaks.

[c] Summation of integrals of aliphatic polymethylene -$(CH_2)_n$-[a] at 33 ppm and -$(CH_2)_n$-[b] at 31 ppm.

[d] Ratio of integrals of aliphatic polymethylene -$(CH_2)_n$-[a] at 33 ppm and -$(CH_2)_n$-[b] at 31 ppm.

Table S9.3 Parameters of ssNMR experiments measured for soil and plant samples. HF-treated soil samples (Soil #^{HF}); Non-treated soil samples (Soil #^{Non}); NS: number of scans; number of points of time domain for the direct (td2) and indirect (td1) dimensions; the acquisition time of the direct dimension (aq2); the evolution time of indirect dimension (aq1); d1: recycle delay. N/A: not applicable.

	Sample	DNP juice	Experiment	Time	Figure #	MAS (kHz)	NS	td2	td1	aq2 (ms)	aq1 (ms)	d1 (s)
ssNMR	#1-7^{HF}	N/A	^{13}C-MultiCP	35 h	1a, 3a, S6		16k	2494		24.9		1
9.4 T	#1^{Non}	N/A	^{13}C-MultiCP	35 h	1a		16k	2494		41		2
298 K	#1^{HF}	N/A	^{13}C-CP	9.5 h	S3a	14	16k	4096	N/A	41	N/A	2
	#1-7^{HF}	N/A	^{13}C-NQS	9.5 h	S3		16k/8k	4096		41		2
MAS DNP 14.1 T 100 K	#1^{HF}	D_2O/H_2O, 9/1	^{13}C-CP	2 min	1b, 2a, S8a, S8b	8	64	4096		22.9		1.8
	#1^{Non}	d_6-DMSO/H_2O, 9/1	^{13}C-CP	0.6 h	S8a, S8c	8	1k	2048		10.2		2.1
	#4^{Non}	d_6-DMSO/H_2O, 9/1	^{13}C-CP	5 min	S8e	8	128	2048	N/A	10.2	N/A	2.0
	#6^{Non}	d_6-DMSO/H_2O, 9/1	^{13}C-CP	0.25 h	S8f	8	512	2048		10.2		1.7
	plant (edge)	D_2O/H_2O, 9/1	^{13}C-CP	0.5 h	2a, S4a, S4b	10.5	512	4096		22.9		3
	plant (inland)	D_2O/H_2O, 9/1	^{13}C-CP	0.75 h	S4c, S4d	10.5	512	4096		22.9		5.2
	#1^{HF}	D_2O/H_2O, 9/1	^{13}C dipolar-INADEQUATE	16 h	1c	10.5	320	2048	98	17.2	2.7	1.8
	plant (edge)	D_2O/H_2O, 9/1	^{13}C dipolar-INADEQUATE	13 h	2b, 2c	8	160	2048	100	17.2	1.7	2.9
	plant (inland)	D_2O/H_2O, 9/1	^{13}C dipolar-INADEQUATE	23 h	S5a	10.5	160	2048	100	17.2	1.7	5.2
	#1^{HF}	D_2O/H_2O, 9/1	^{1}H-^{13}C HETCOR (PMLG)	1 h	1e	8	40	4096	38	22.9	3.8	1.8
	#1^{HF}	d_6-DMSO/D_2O/H_2O	^{1}H-^{13}C HETCOR (FSLG)	4.5 h	1g, S2	10.5	96	4096	64	22.9	2.4	2.6

Table S9.3 (cont'd)

	#1^{Non}	d_6-DMSO/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	12 h	S7a, S7b	8	512	2048	40	10.2	1.5	2.1
	#1^{Non}	d_6-DMSO/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	24 h	S7a	8	1k	2048	40	10.2	1.5	2.1
	#2^{HF}	D_2O/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	3 h	S7b	10.5	32	4096	128	22.9	4.8	2.6
	#4^{Non}	d_6-DMSO/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	1.4 h	3c, S7b	8	64	2048	40	10.2	1.5	2.0
	#6^{Non}	d_6-DMSO/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	10 h	3c, S7b	8	512	2048	40	10.2	1.5	1.7
	#6^{Non}	d_6-DMSO/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	7.3 h	3c	8	512	2048	30	10.2	1.1	1.7
	plant (edge)	D_2O/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	1.7 h	2d, 2e	10.5	8	4096	128	22.9	4.8	5.8
	plant (inland)	D_2O/H_2O, 9/1	^1H-^{13}C HETCOR (FSLG)	1.7 h	S5b, S5c	10.5	8	4096	128	22.9	4.8	5.8

Table S9.4 ^{13}C and ^{1}H chemical shifts of soil and plant samples. The assignments are identified from ^{13}C INADEQUATE spectra and ^{1}H-^{13}C HETCOR spectra. The ppm values are shown as ^{13}C (^{1}H), e.g., 88.9 (3.3) represents ^{13}C (^{1}H) chemical shifts (ppm) of cellulose i4. Aromatic rings include guaiacyl (G), syringyl (S), *p*-hydroxyphenyl (H) units. Not applicable (/). Unidentified (-). Uncertain (U).

Carbohydrate											**CO**	Sample
Cellulose: Interior (i); Surface (s)						Xylan: 2-fold (Xn^{2f}); 3-fold (Xn^{3f})						
C1	C2, 3, 5	iC4	sC4	iC6	sC6	$C1^{3f}$	$C4^{2f}$	$C4^{3f}$	$C5^{2f}$	$C5^{3f}$		
104.9 (4.3)	74.7 (3.4) 72.8 (3.5)	90.2 (3.6)[e] 89.5 (3.9)[a,b] 87.8 (3.8)[c]	85.5 (3.5)[h] 84.4 (3.5)[f] 83.3 (3.5)[g]	65.9 (3.5)[c] 63.7 (3.5)[a,b]	62.4 (3.6)[f] 60.6 (3.6)[g]	103.1 (4.4)	81.2 (3.6)	79.3 (3.5)	63.6 (-)	63.2 (-)	173.4 175.0 179.6	$\#1^{HF}$
105.1 (4.3)	75.3 (3.9) 74.0 (3.6) 72.5 (3.4)	90.2 (4.0)[e] 88.5 (4.8)[a,b] 87.6 (5.1)[c]	86.0 (4.9)[h] 83.7 (3.9)[g]	65.4 (3.4)[c] 63.8 (3.7)[a,b]	62.3 (3.4)[f] 61.1 (3.2)[g]	102.7 (4.2)	82.0 (4.2)	79.0 (4.3)	-	-	173.3 175.3 180.7	$\#1^{Non}$
104.8 (4.3)	74.6 (3.3) 72.8 (3.4)	89.5 (3.4)[a,b] 88.2 (3.7)[c]	86.7 (3.9)[h] 83.9 (3.5)[f,g]	66.0 (3.8)[c] 64.7 (3.5)[a,b]	63.0 (3.6)[f] 61.6 (3.2)[g]	103.1 (4.4)	82.7 (3.7)	79.4 (3.4)	-	-	172.3 175.6 180.4	$\#2^{HF}$
105.2 (4.4)	75.7 (3.7) 74.6 (3.6) 73.3 (3.7)	89.3 (4.2)[a,b] 88.0 (3.9)[c]	85.7 (3.9)[h] 84.3 (4.0)[f] 83.2 (3.9)[g]	65.7 (3.7)[c] 64.3 (3.7)[a,b]	62.7 (3.4)[f]	103.3 (4.7)	81.5 (4.0)	77.6 (3.7)	-	-	174.0 175.6 180.1	$\#4^{Non}$

Table S9.4 (cont'd)

105.0 (4.9)	75.8 (3.5) 74.6 (3.7) 72.5 (3.5)	89.4 (4.0)[a,b] 87.4 (3.9)[c]	86.0 (3.7)[h] 83.9 (3.8)[f,g]	66.2 (3.1)[c] 64.1 (3.1)[a,b]	62.3 (3.3)[f]		103.4 (4.9)	81.1 (4.4)	79.0 (4.3)	-	-	173.8 176.3 180.9	#6[Non]
105.0 (4.2)	74.5 (3.3) 72.3 (3.3)	88.8 (3.3)[a,b] 87.7 (3.7)[c]	85.8 (3.7)[h] 83.7 (3.3)[f,g]	64.5 (3.6)[a,b]	62.4 (3.6)[f]		101.9 (4.3)	81.4 (3.5)	78.8 (3.5)	65.2 (-)	63.2 (-)	174.5	plant edge
105.0 (4.3)	74.7 (3.3) 72.2 (3.4)	89.0 (3.3)[a,b] 87.4 (3.7)[c]	84.3 (3.3)[f] 83.0 (3.4)[g]	64.6 (3.6)[a,b]	62.7 (3.4)[f]		101.9 (4.3)	81.4 (3.5)	78.9 (3.4)	-	63.6 (-)	173.1	plant inland
Aromatic													
S2/6	G2	G5, H3/5	G6	H2/6	G1/H1	S4	S1	G4	G3	S3/5	H4	Sample	
105.1 (6.0)	111.5 (6.1)	114.9 (6.3)	117.9 (6.6) 121.0 (6.6)	124.7 (6.7)	128.5 (/)	132.3 (/)	136.5 (/)	145.6 (/)	148.1 (/)	153.6 (/)	- (/)	#1[HF]	
104.6 (6.0)	110.1 (6.0)	114.3 (6.2)	118.9 (6.5) 122.2 (6.5)	126.8 (5.8)	129.4 (/)	133.4 (/)	136.4 (/)	145.4 (/)	147.7 (/)	152.6 (/)	162.1 (/)	#1[Non]	
105.1 (6.1)	111.3 (6.3)	115.8 (6.4)	117.7 (6.7) 121.7 (6.7)	125.9 (6.9)	129.7 (/)	132.4 (/)	136.1 (/)	143.4 (/)	147.1 (/)	153.4 (/)	163.3 (/)	#2[HF]	
106.9 (6.2)	111.1 (6.1)	115.5 (6.0)	118.9 (6.1) 122.2 (6.4)	126.7 (6.0)	130.0 (/)	133.1 (/)	136.8 (/)	143.9 (/)	148.4 (/)	152.5 (/)	159.7 (/)	#4[Non]	

Table S9.4 (cont'd)

105.7 (6.1)	113.0 (6.0)	115.6 (6.7)	118.4 (6.2) 122.2 (6.1)	127.3 (6.1)	130.8 (/)	133.7 (/)	136.8 (/)	144.5 (/)	147.7 (/)	152.9 (/)	161.6 (/)	#6^{Non}
-	111.9 (6.7)	116.0 (6.6)	121.6 (6.8)	127.4 (6.7)	130.6 (/)	132.8 (/)	137.0 (/)	144.2 (/)	147.6 (/)	153.0 (/)	-	plant edge
-	111.8 (6.3)	116.0 (6.5)	121.7 (7.3)	126.6 (6.8)	131.8 (/)	133.8 (/)	-	145.1 (/)	147.9 (/)	152.8 (/)	-	plant inland

Aliphatic									
-CH_3	-OCO$\underline{C}H_3$	-$\underline{CH_2}$- CH=CH-	-CH_2-^{b}	-CH_2-^{a}	-CCH	R_3CH & R_4C	-OCH_3^{a}	-OCH_3^{b}	Sample
17.9 (0.8)	23.2 (1.6)	25.7 (1.4)	30.6 (1.4)	33.4 (1.3)	40.8 (1.8)	43.6 (2.9)	50.8 (4.1)	56.6 (3.5)	#1^{HF}
18.7 (0.6)	23.8 (1.3)	25.9 (1.6)	30.1 (1.4)	33.6 (1.3)	38.2 (2.8)	43.3 (3.3)	50.0 (2.5)	56.5 (3.3)	#1^{Non}
18.0 (0.7)	23.1 (1.1)	26.4 (1.1)	30.1 (1.4)	33.4 (1.3)	39.6 (1.8)	44.3 (3.1)	50.0 (4.2)	56.3 (3.6)	#2^{HF}
18.5 (0.5)	22.4 (1.6)	26.7 (1.3)	30.2 (1.1)	34.0 (1.3)	39.4 (2.8)	43.5 (2.8)	50.6 (3.2)	56.3 (3.2)	#4^{Non}
19.5 (0.7)	23.2 (1.6)	27.9 (1.1)	31.0 (1.2)	34.0 (1.3)	40.1 (2.4)	44.6 (2.2)	51.8 (2.5)	56.4 (2.9)	#6^{Non}

Table S9.4 (cont'd)

-	22.0 (1.9)	25.8 (1.4)	30.2 (1.8)	32.6 (1.3)	37.9 (5.2)	-	-	56.2 (3.5)	plant edge
18.2 (1.0)	22.1 (1.0)	26.0 (1.4)	30.5 (1.3)	32.8 (1.3)	38.0 (5.2)	-	-	58.0 (3.4)	plant inland

Table S9.5 Physiochemical properties of soil samples based on the depth. Depth is an average value, for example, 5 cm depth means the sample is collected from 0-10 cm section. Samples #1 to #7 are the samples measured by ssNMR and DNP.

Soil Sample #	Average Depth (cm)	Bulk Density (g/cm^3)	Loss on Ignition (Wt %)	Total Carbon (%)
1	5	0.265	20.5	8.356
	15	0.240	27.1	
	25	0.318	17.9	
	35	0.247	28.5	
2	45	0.294	19.9	10.826
	55	0.368	13.7	
	65	0.141	40.0	
	75	0.166	43.2	
3	85	0.166	40.9	12.326
	95	0.207	31.1	
4	105	0.152	51.6	14.039
	115	0.117	59.4	
	125	0.182	35.4	
5	135	0.114	64.8	10.717
	145	0.119	60.7	
	155	0.146	51.6	
6	165	0.135	39.7	11.165
7	175	0.306	16.1	8.114

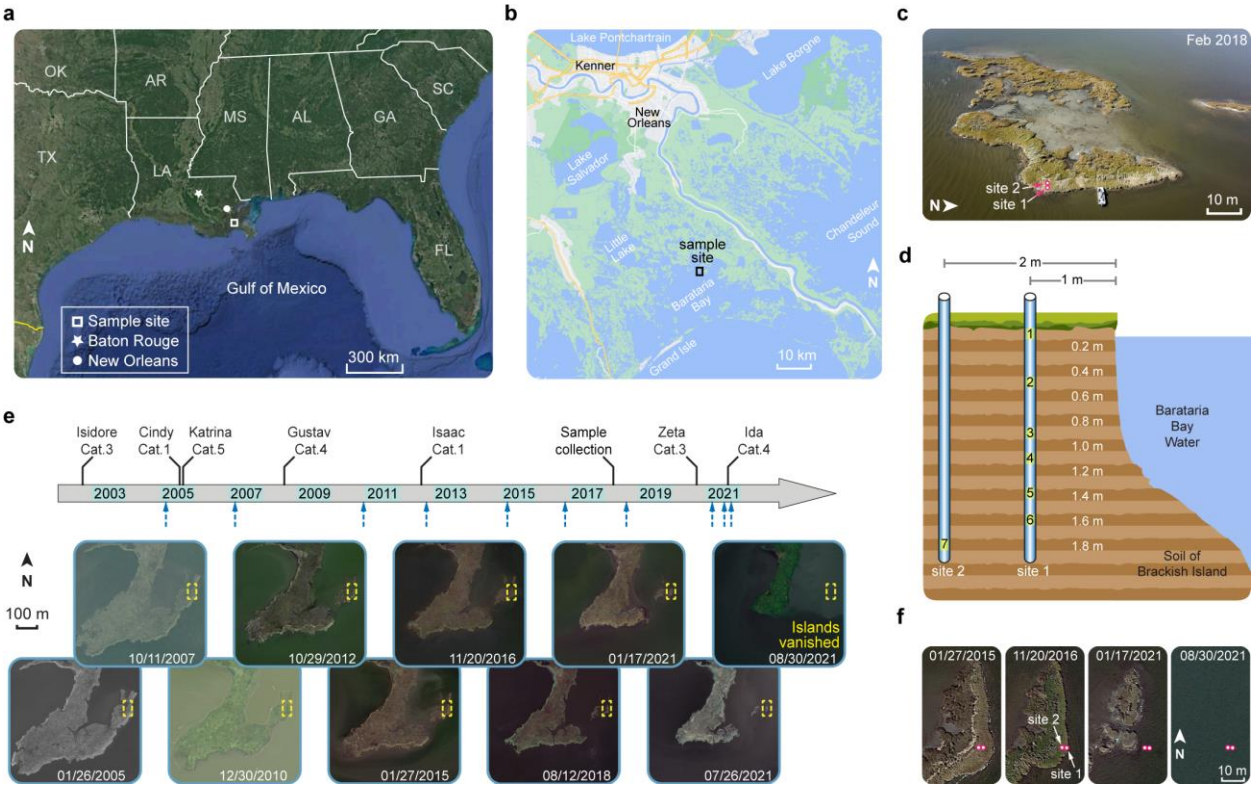

Figure S9.1 Wetland soil from a recently vanished brackish island. **a,** The island is 55 km southeast of New Orleans, and 160 km away from Baton Rouge, the capital city of the state of Louisiana, USA. **b,** Soil materials were collected in February 2018 from a brackish island in Barataria Bay (GPS coordinates: 29°26'36.9"N, 89°53'59.0"W). **c**, Picture of the *Spartina alterniflora*-dominated island with the two sample sites marked. **d,** Location and depth of seven soil samples used for structural characterization. Two poles were used to extract the soil materials, which were divided into 10-cm sections, based on the depth. **e,** Timetable summarizing the landscape change of the island and adjacent lands over two decades. The catastrophic hurricanes that affected this island and the category (Cat.) numbers of these hurricanes are labeled. Blue dashline arrows indicate the ten time points where pictures of the landscape are provided. **f,** Zoom-in view of the dashline boxes in panel (e). Positions of the two sample sites are marked using magenta circles to guide the comparison. The island has been finally disappeared in 2021.

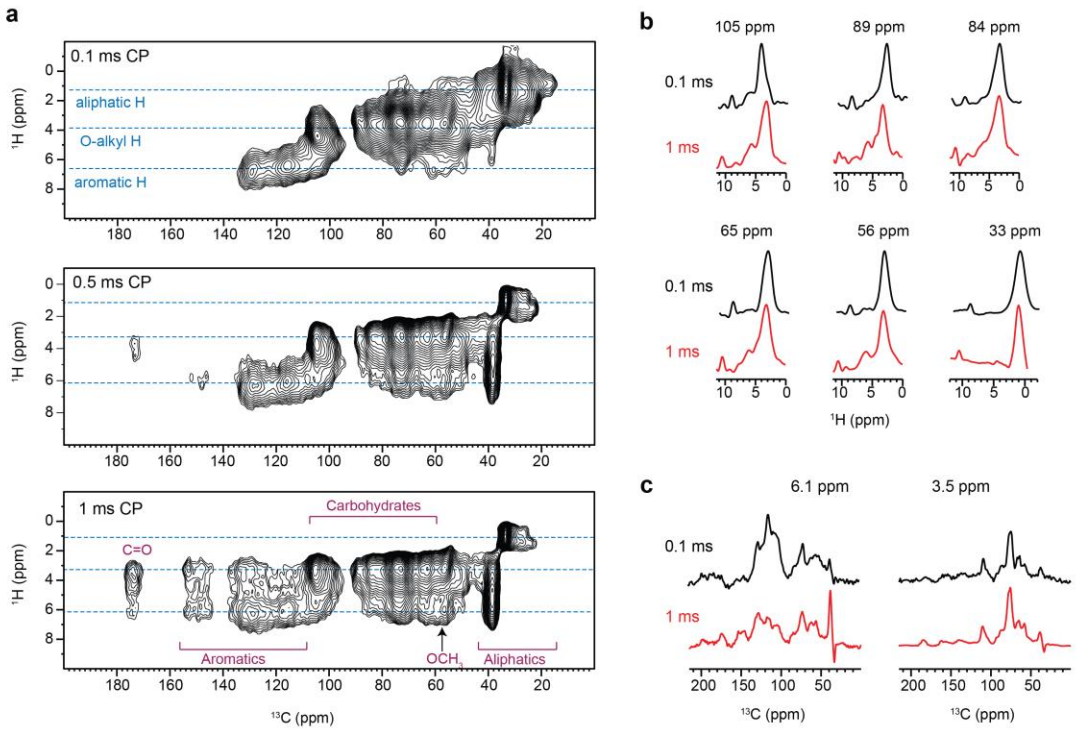

Figure S9.2 DNP 2D ^{1}H-^{13}C correlation spectra of unlabeled HF treated soil. **a,** 2D ^{1}H-^{13}C correlation spectra measured with short (0.1 ms), medium (0.5 ms), and long (1.0 ms) of CP contact times of unlabeled soil sample 1. The blue dash lines show the key proton positions of aliphatics, carbohydrates, and aromatics. The spectrum with 1 ms CP shows intermolecular cross peaks between aromatics and carbohydrates. **b,** Representative ^{1}H cross sections extracted at different carbon sites from the 0.1 ms (black) and 1 ms (red) CP contact times. **c,** Representative ^{13}C cross sections from the 0.1 ms (black) and 1 ms (red) CP contact times. The spectra were measured at 10.5 kHz.

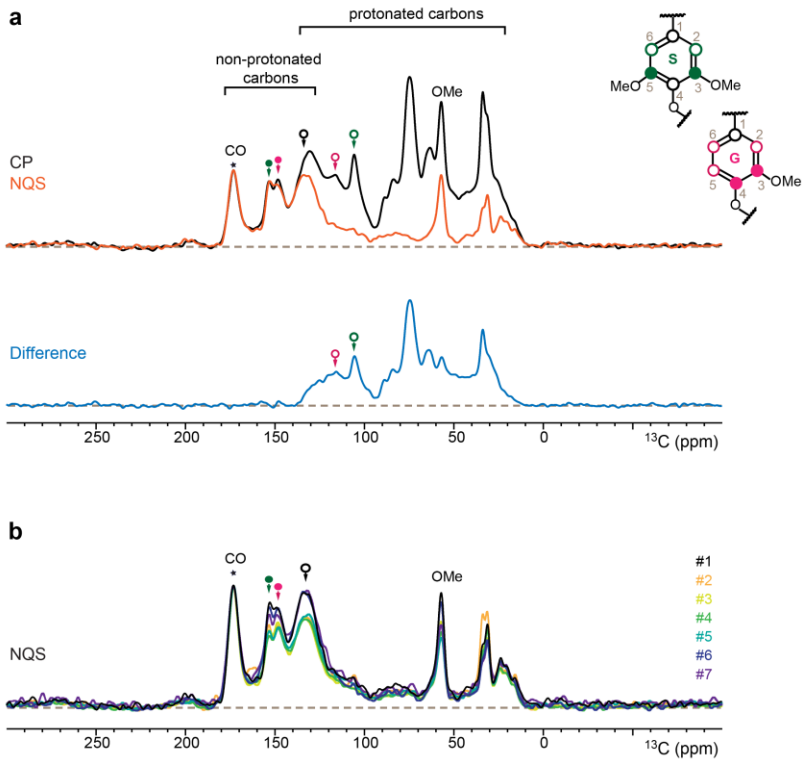

Figure S9.3 1D non-quaternary suppression selectively detecting non-protonated carbons. **a**, Comparison of ^{13}C CP and NQS spectra of the HF treated soil sample #1 at room temperature. The difference spectrum shows only protonated carbons. **b**, Overlay of 1D NQS ^{13}C spectra of seven soil samples. All spectra are normalized by the CO peak (asterisk). QNS spectra mainly shows non-protonated carbons, with methyl carbons as an exception due to their rapid molecular motions.

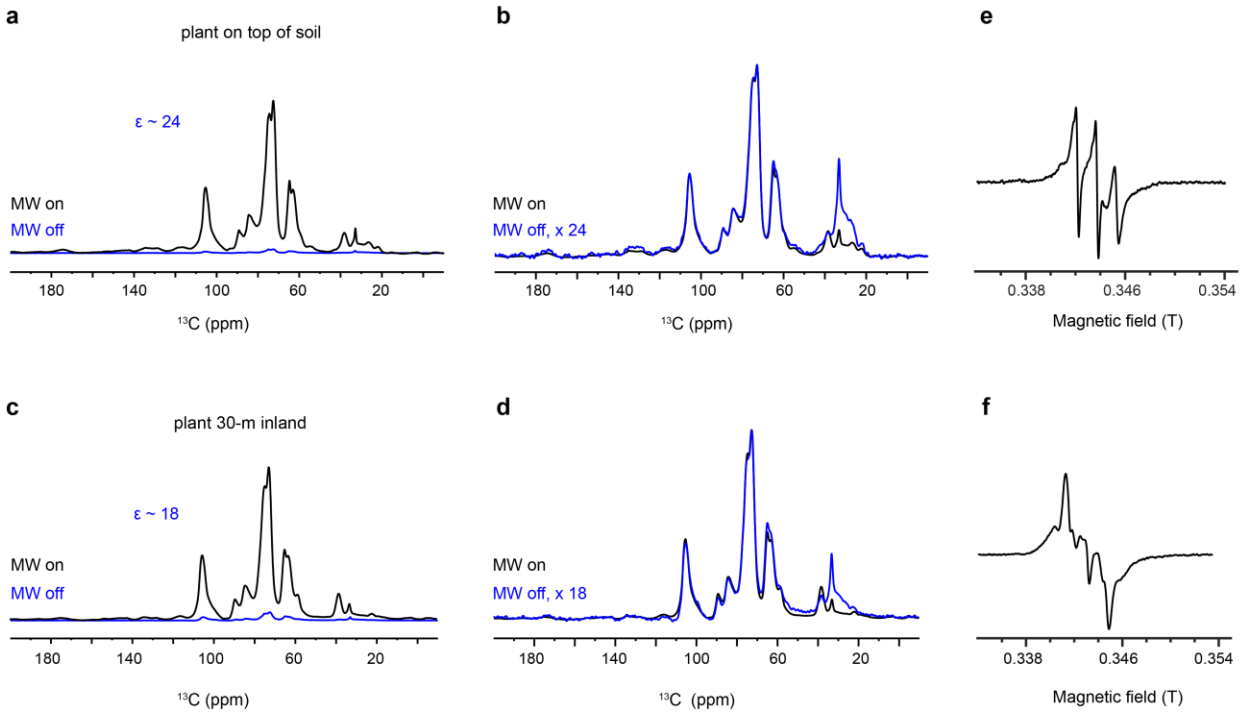

Figure S9.4 DNP and EPR spectra of plant samples. **a,** DNP enhances the sensitivity by 24-fold for the plant on top of soil (on the edge of the island). **b,** Magnification of the microwave (MW) off spectra showed overall consistent pattern with the MW-on spectra, revealing homogeneous polarization by DNP, except for the polymethylene peaks. **c** and **d,** The 30 m inland plant sample also showed 18-fold of DNP enhancement, with homogeneous DNP of carbohydrates and aromatics as shown in panel **e**. **f,** Room temperature EPR spectra of AsymPolPOK (D_2O/H_2O, 90/10 Vol%) at 9.6 GHz for these inland plants.

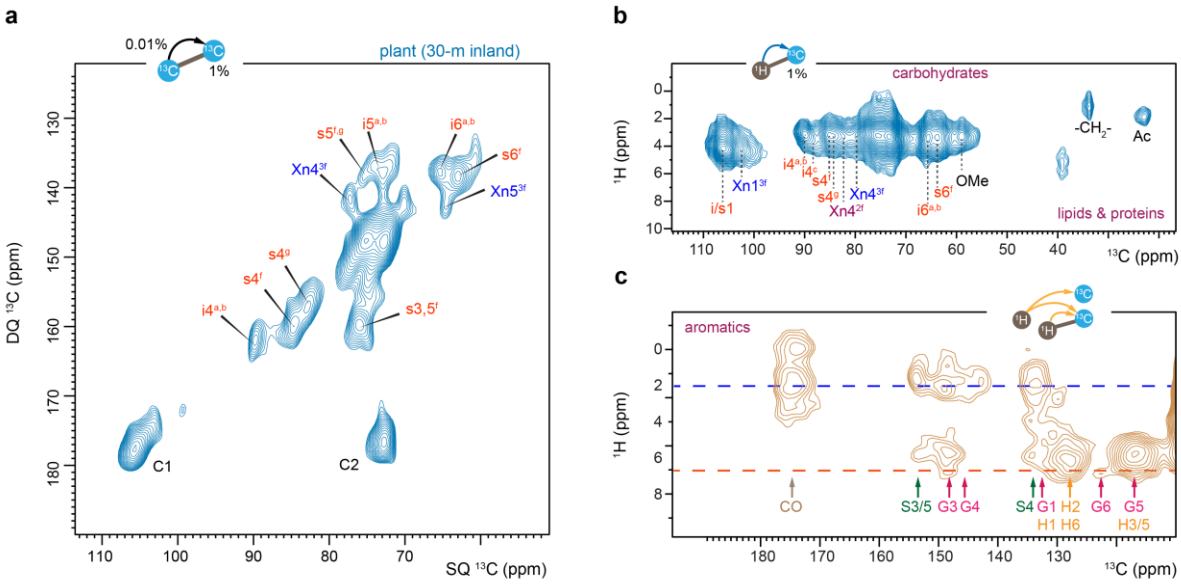

Figure S9.5 2D $^{13}C/^{1}H$-^{13}C spectra of unlabeled plants 30-m inland. **a,** Carbohydrate region of DNP enhanced ^{13}C-^{13}C refocused INADEQUATE spectrum of unlabeled plant samples collected 30-m inland. Signals are resolved for cellulose and xylan. **b**, 2D Carbohydrate and aliphatic region of 2D ^{1}H-^{13}C HETCOR spectrum of the 30-m inland plant. A short 0.1 ms CP was used to emphasize the one-bond correlations. **c**, The aromatic region collected with long (1 ms) CP contact to show aromatic-aliphatic correlations. No cross peaks were observed with carbohydrates. The spectroscopic features are largely consistent in the plants collected at different locations of the island.

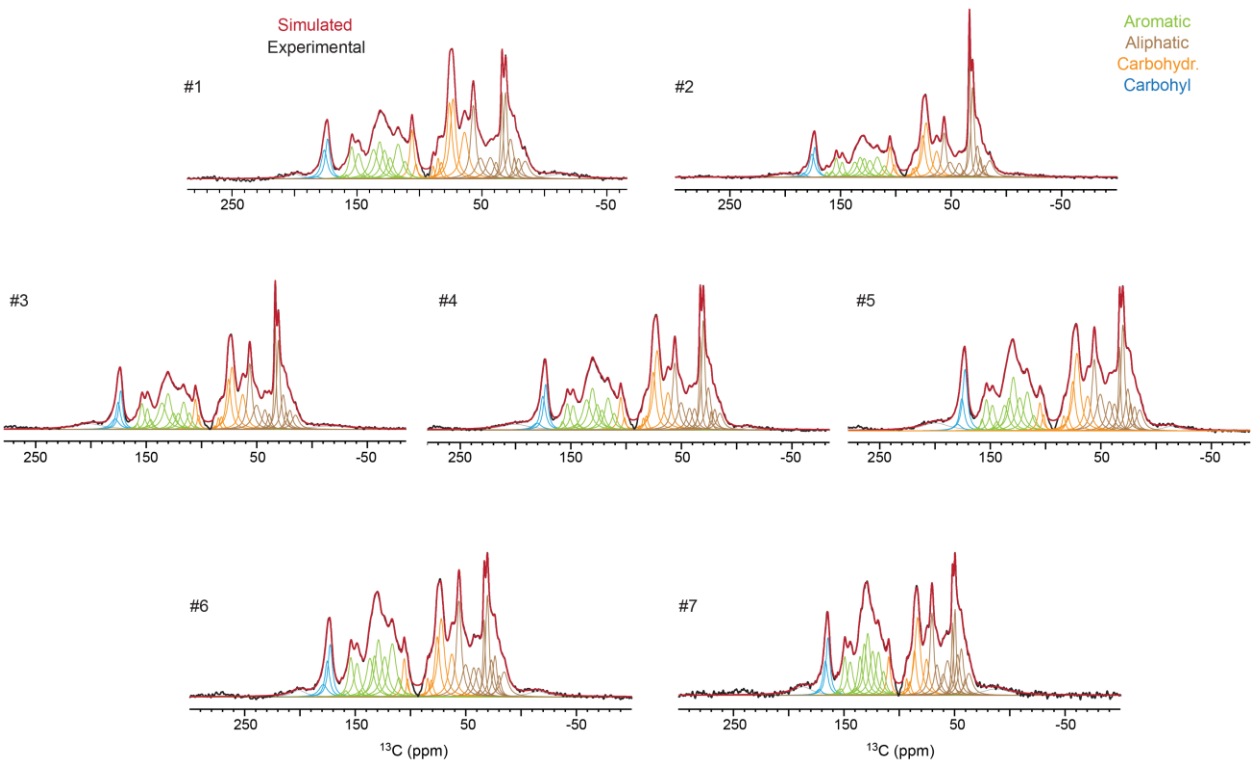

Figure S9.6 Spectral deconvolution of quantitative ^{13}C spectra for molecular composition. For each sample, the simulated spectra (dark red) fit the experimentally measured 1D ^{13}C MultiCP spectra (black). Underneath are the individual peaks that contribute to carbohydrate (orange), aliphatic (brown), aromatic (green) and carbonyl sites (blue). The peak list is guided by the resolvable sites obtained from high-resolution 2D data. Information on the deconvolution was documented in Table S9.1.

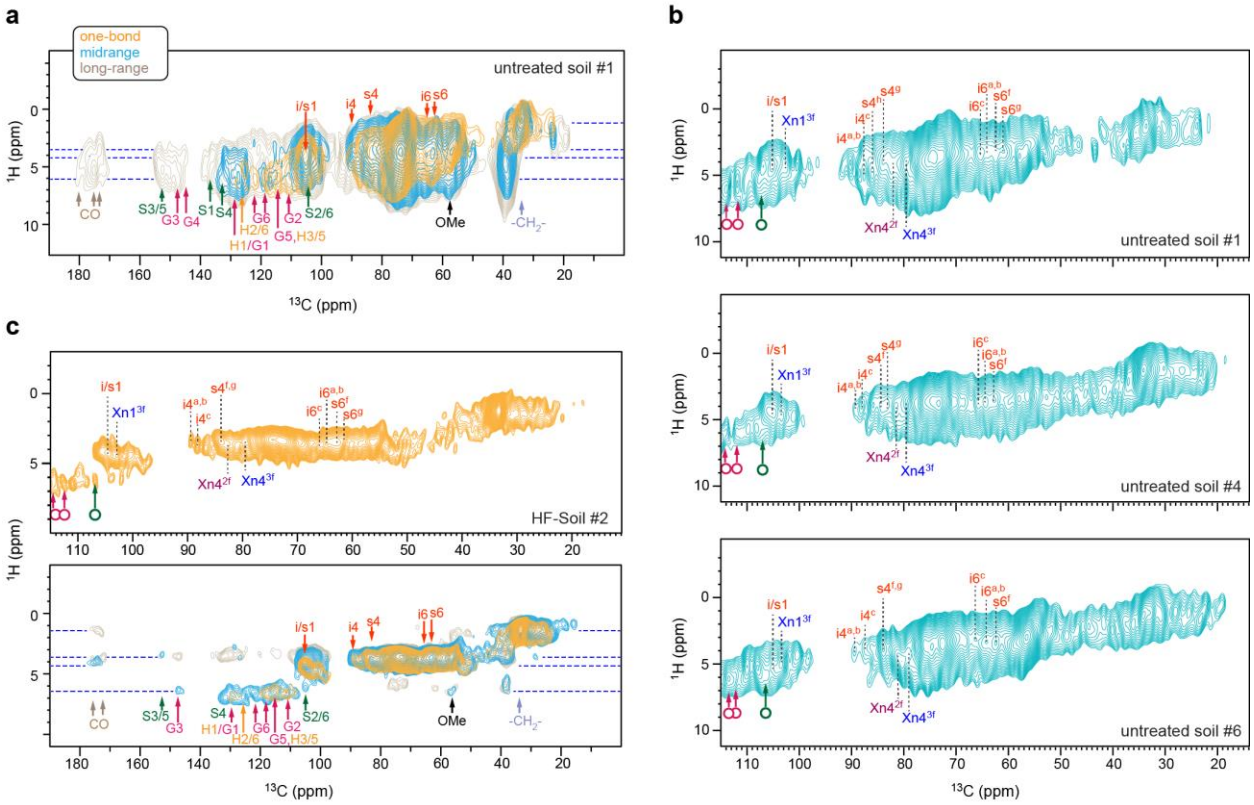

Figure S9.7 2D ^{1}H-^{13}C correlation DNP spectra of untreated soil. **a,** 2D ^{1}H-^{13}C correlation spectra of untreated soil sample 1 measured with 0.1 ms (yellow), 0.5 ms (blue), and 1 ms (grey) CP contact times. **b,** Zoom-in regions of carbohydrate and aliphatic signals in three untreated soil samples (1, 4, and 6). **c,** Additional 2D ^{1}H-^{13}C spectra of HF-treated soil sample 2. The key signals of carbohydrates and aromatics are observable. Top and bottom panels show the aliphatic/carbohydrate and aromatic signals, respectively.

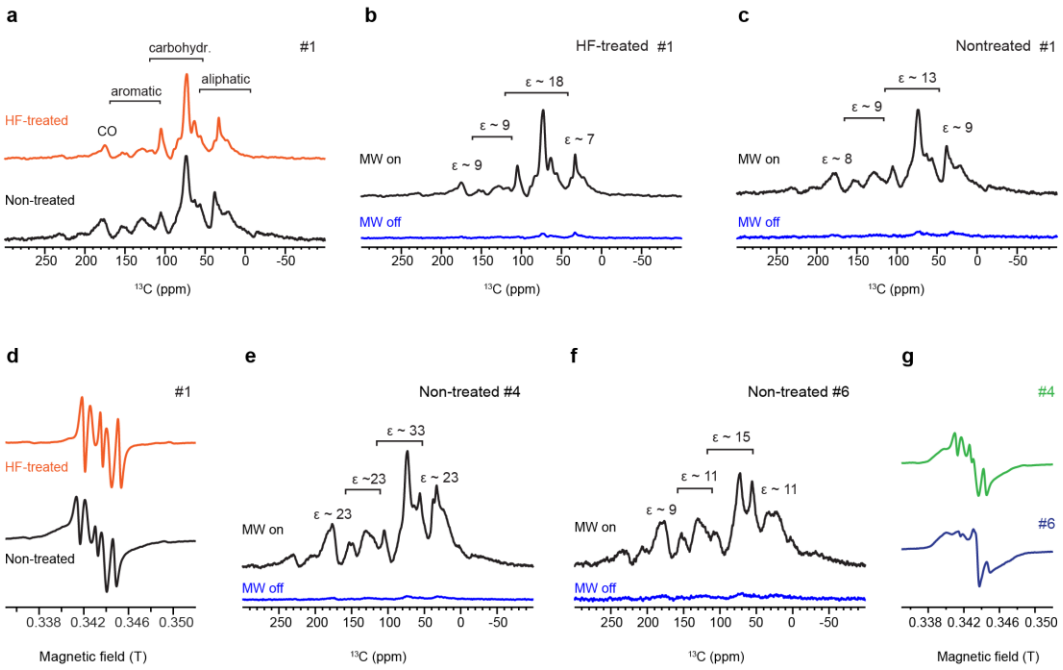

Figure S9.8 1D DNP ^{13}C spectra and EPR of soil samples. **a**, Comparison of ^{13}C spectra of the HF and non-HF treated materials of soil sample #1 under DNP enhancement. **b**, Comparison of ^{13}C spectra with and without microwave (MW) irradiation collected on HF-treated soil sample 1. The enhancement is 18-fold for carbohydrates, 9-fold for aromatics, 7-fold for CO, 9-fold for most aliphatic carbons, and 7-fold for the polymethylene CH_2 peak. **c,** MW-on and MW-off spectra of non-treated soil sample 1. The enhancement is 13-fold for carbohydrates, 9-fold for aromatics, 8-fold for CO, and 8-fold for most aliphatic carbons, and 9-fold for the polymethylene CH_2 peak. **d**, EPR spectra of soil sample #1 with (top) and without (bottom) HF treatment, hydrated using the d_6-DMSO/D_2O/H_2O matrix. **e,** ^{13}C spectra with and without microwave (MW) irradiation collected on native soil sample #4 without HF treatment, showing enhancement factor of 33-fold for carbohydrate and 23-fold for all other carbon sites. **f,** MW-on and MW-off spectra of untreated soil sample #6, showing enhancement of 15 for carbohydrate and 9-11 for other carbon sites. **g**, EPR spectra of AsymPolPOK at 9.6 GHz for untreated sample #4 (green) and #6 (blue), with a solvent of d_6-DMSO/D_2O (90:10 Vol%).

CHAPTER 10: CCMRD: A SOLID-STATE NMR DATABASE FOR COMPLEX CARBOHYDRATES

Research paper reprinted with permission: *J. Biomol. NMR* 74, 239-250 (2020)

Authors: Xue Kan, Wancheng Zhao, Malitha C. Dickwella Widanage, Alex Kirui, Uluc Ozdenvar, and Tuo Wang

10.1 Abstract

Carbohydrates are essential to various life activities in living organisms and serve as the central component in many biomaterials. As an emerging technique with steadily improving resolution, solid-state Nuclear Magnetic Resonance (NMR) spectroscopy has the unique capability in revealing the polymorphic structure and heterogeneous dynamics of insoluble complex carbohydrates. Here, we report the first solid-state NMR database for complex carbohydrates, Complex Carbohydrates Magnetic Resonance Database (CCMRD). This database currently holds the chemical shift information of more than four hundred solid-state NMR compounds and expects rapid expansion. CCMRD provides open portals for data deposition and supports search options based on NMR chemical shifts, carbohydrate names, and compound classes. With the timely implementation, this platform will facilitate spectral analysis and structure determination of carbohydrates and promote software development to benefit the research community. The database is freely accessible at www.ccmrd.org.

10.2 Introduction

Complex carbohydrates play central roles in many biological processes such as energy storage, structural building, and cell recognition[1-3]. These biomacromolecules also form the basis for novel biomaterials such as the scaffolds developed for tissue engineering and the carriers for drug delivery[4, 5]. Polysaccharides are the polymers of the monosaccharide building blocks linked

by glycosidic bonds. Their structural complexity is multifaceted, including substantial variations in the monosaccharide composition, the glycosidic linkages and the anomeric configuration of subunits, the branching pattern of the backbone by sidechains, chemical modifications such as acetylation and methyl esterification, hydrogen bonding patterns, and more subtly, torsional conformations (**Figure 10.1**)[6-11]. The structural complexity determines the physical and chemical properties and leads to the highly heterogeneous dynamics, hydration and intermolecular interactions of carbohydrates in biomaterials.

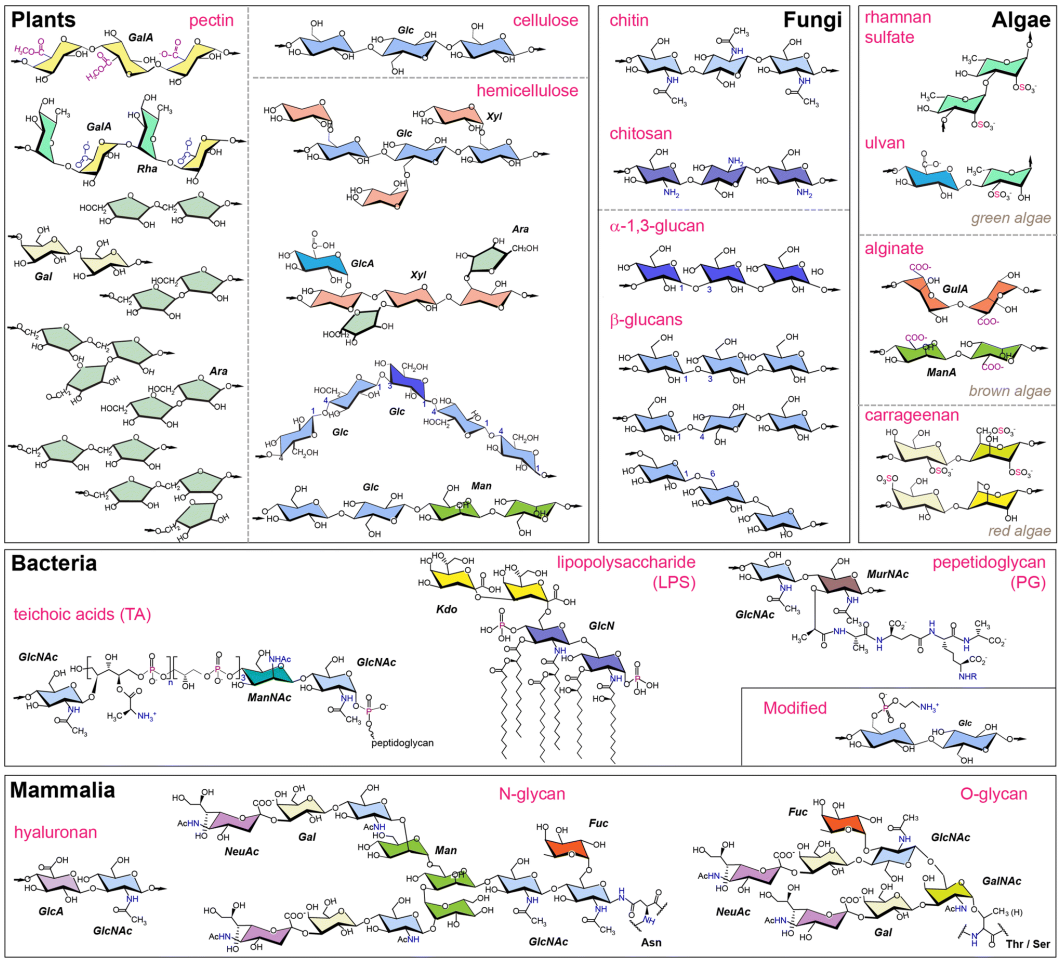

Figure 10.1 Representative structure of complex carbohydrates from different origins. Most of these polysaccharides have been characterized by solid-state NMR. The structures of polysaccharides in their cellular environments are often more complicated and irregular.

Complex carbohydrates are significantly under-investigated compared with other biomacromolecules such as nucleic acids and proteins. Polysaccharides are mostly insoluble in water, partially indigestible, and often non-crystalline, making it practically difficult to analyze their high-resolution structure[12]. The conventional methods typically need to hydrolyze the digestible polysaccharides to smaller, soluble segments[13, 14] and grind or ball-mill the indigestible portion into fine particles to enable structural analysis by solution-NMR and mass spectrometry [15-18]. The biochemical treatments could considerably perturb and even restructure the biomolecules. Due to these technical hurdles, our knowledge of the structure and dynamics of carbohydrates remains inadequate, which has impeded the development of carbohydrate-based biorenewable energy biomaterials[15].

Solid-state NMR (ssNMR) spectroscopy is an emerging technique that is capable of elucidating the molecular structure and dynamics of insoluble polysaccharides in their native cells or environments, without the need for pretreatments. This technique has long been employed to reveal the polymorphic structure of cellulose and matrix polysaccharides, which heavily relies on one-dimensional ^{13}C spectra [19, 20]. With better resolution, multidimensional ssNMR has further revealed biopolymer interactions in plant primary and secondary cell walls, the supramolecular architecture and pigment deposition of pathogenic fungal cell walls and biofilms, the composition of mammalian cells, the structure of bacterial peptidoglycans and lipopolysaccharides and their interactions with antibiotics, the difference of galactolipids, storage polysaccharides and cell wall components in algae, as well as naturally modified or artificially functionalized polysaccharides[21-42]. With the readily improving resolution from the improvement and development of ultrahigh-field magnets[43], the sensitivity enhancement from dynamic nuclear polarization (DNP) and solid-state CryoProbe[44-48], and the assistance from various NMR methods such as ultrafast spinning and

proton detection[49-51] as well as paramagnetic and spectral editing techniques[52-54], many long-standing questions regarding the structure and dynamics of complex carbohydrates now become feasible and await systematic investigations.

At the same time, the rapid expansion of ssNMR applications in carbohydrate research inevitably necessitates the development of a databank to facilitate information storage and sharing as well as statistical analysis and software development. Here we developed the Complex Carbohydrates Magnetic Resonance Database (CCMRD) together with its web interface for public access, information search, and data deposition. The database is freely available at www.ccmrd.org.

10.3 Methods

10.3.1 Database Assembly

The CCMRD system is constructed with a three-tier architecture: server, client and database. The data records are stored in MySql database[55]. The initial data were collected from studies that were published over the past four decades. The database is regularly maintained and updated on a weekly basis. The structure of the database is designed to accommodate the complexity of polysaccharide structure. The data are organized in multiple levels to include as many details as possible. The base unit/monosaccharide serves as a central entity, which is directly related to the NMR chemical shifts and the branching pattern. Such information belongs to a compound, which is an upper-level central entity. The compound is connected to the information of organism, references and experimental details (**Figure 10.2a**). The organizational schema is shown in **Figure S10.1.**

10.3.2 Server/Web Interface

The server/web interface was programmed using PHP/Laravel framework[56]. The website is currently hosted on Amazon Web Services. The web server uses Amazon cloud EC2 service,

and the mysql database uses Amazon cloud RDS service. A physical server is being implemented for long-term maintenance and development. The web interface provides three options for data search based on NMR chemical shifts, polysaccharide names, and compound classes. A data submission portal is also available, which contains a multi-page process for collecting all the required information. All the submitted data are reviewed before the final deposition.

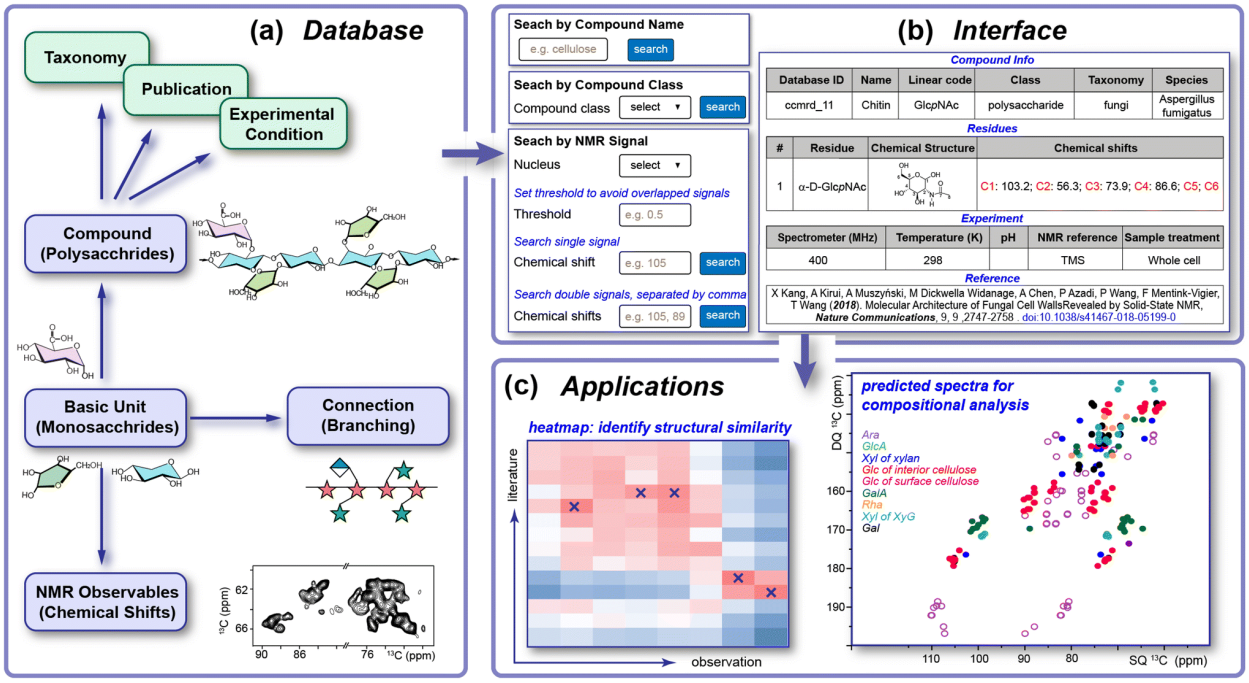

Figure 10.2 Overview of the carbohydrate database. **a,** Structure of the database. **b,** Search interface (left) and a representative entry (right). More detailed instructions for the search and data deposition are in Figure S10.2, S10.3. **c,** For possible applications, a cellulose ^{13}C chemical shift heatmap (left) and a synthesized spectrum of 40 conformers from 9 plant carbohydrates (right) are shown as examples.

10.4 Results and Discussion

To date, CCMRD stores the ssNMR chemical shifts of 435 complex carbohydrates from various organisms (**Figure 10.3a**). Plants account for half of the deposited compounds while fungi,

algae and bacteria share one-third of the publications. The NMR chemical shifts are obtained predominantly from high-resolution 2D/3D correlation solid-state NMR spectra that allows the unambiguous identification of carbon-carbon connectivity in each carbohydrate[26, 57]. Glycosyl linkage and carbohydrate compositional analysis, genetic mutants or chemical extraction of certain components, as well as Density Functional Theory (DFT) calculations are often coupled to implement and effectively verify the NMR assignments. Critical breakthroughs of ssNMR technology have substantially promoted high-resolution studies over the last decade as evidenced by the upsurge of compound numbers (**Figure 10.3b**).

10.4.1 Data Search

The web interface provides three options for data search (**Figure S10.2**):

(1) Search by chemical shifts: This feature is specially designed to help users with unknown signals that are identified in NMR spectra. Users need to specify three parameters include the nucleus of interest, the chemical shift and the threshold. The nucleus is chosen from ^{13}C, ^{1}H, ^{15}N, and ^{31}P. The threshold determines how accurate the search results should match the targeted value. The default threshold is set to 1 ppm for ^{13}C, 3 ppm for $^{15}N/^{31}P$ and 0.5 ppm for ^{1}H based on the typical ssNMR resolution and these values can be changed by the users. The returned result is a list of compounds with matched signals, with the detailed information of carbohydrate structure and experimental conditions underneath each CCMRD compound ID (**Figure 10.2b**).

In addition, CCMRD also support double-signal search using two chemical shifts from the same compound. By inputting two chemical shifts separated by a comma, the server will return the entries that match both the inputs within the specified range of tolerance. For each returned entry, its ID number, compound name, residue name and chemical shifts are listed, with the

matched chemical shifts highlighted. The double-signal search can efficiently eliminate irrelevant compounds.

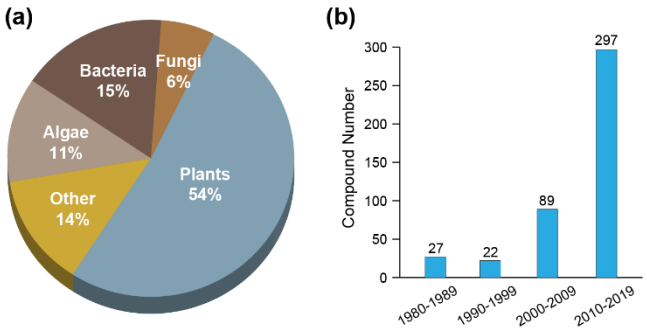

Figure 10.3 Statistics of compounds in CCMRD. The data are categorized by **a,** taxonomical context and **b,** years.

(2) Search by polysaccharide names: The users can conduct a search using a trivial name (**Figure S10.2**). There is no restriction on the input, and both full and partial names are accepted. For example, the users can use either “cel” or “cellulose”. CCMRD will perform a search for the name pattern and return all relevant entries with their ID numbers, with the full details accessible by clicking each ID.

(3) Search by compound classes: Based on the currently available entries, the data deposited in CCMRD are categorized as six classes to facilitate the data search: polysaccharide, oligosaccharide, monosaccharide, lipopolysaccharide, N-acetylglucosamine, and peptidoglycan. This option is useful when the compound type is already known.

To eliminate the potential ambiguity in the atom position labeling of carbohydrates and to provide a more intuitive way to present the data, we have included the molecular structure of each residue with carbon atom numbers so that users can access the data with ease and accuracy. The Symbol Nomenclature for Glycans (SNFG) is also included for full structure presentation but only for those entries with unambiguous information of the linkage patterns[58, 59].

10.4.2 Data Deposition

The web server of CCMRD also provides a portal for data submission (**Figure 10.4 and Figure S10.3)**. The users will need to fill a multi-step form for multiple instances of all the complex carbohydrates in a single publication. The submitted data will be saved in a temporary table and an email will be generated automatically informing the users of the successful submission. An administrator will review, verify, and deposit the submitted record, and a second email containing the CCMRD ID number(s) will be sent automatically to the user to confirm the success of data deposition. The users will need to prepare the following information for data submission:

(1) Provide an email. The email address is only used for communications before the final deposition process.

(2) NMR experimental conditions. The detailed NMR experimental conditions are required to be specified, including magnetic field strengths, sample pH, temperature, NMR reference scale, and sample processing protocols. As some data from the earlier stages of ssNMR has relatively low resolution as measured on low-field magnets, the experimental conditions will assist the research community in evaluating data reliability. Taxonomic domain of the organism and the biological source with full name of genus and species are also required. NCBI TaxID from the NCBI taxonomy database is used for cross-reference. Fields with asterisks are required while the others are optional (**Figure 10.4**).

(3) Compound information. Users are required to provide both the trivial name and linear code of the compound. For the linear code format, we followed the IUPAC recommendations for future compatibility[60]. Compound class also need to be specified, for example, as polysaccharide, monosaccharide, or oligosaccharide.

(4) Residue/Unit. For each residue, the users need to provide the residue/unit name, anomeric configuration, absolute configuration and ring size. Select “unknown” if any field is undetermined. After inputting all the chemical shifts for this residue, additional residues can be added by clicking the “add new residue” button for heteropolysaccharide.

(5) Connection table. The patterns of covalent linkage between different residues need to be clarified by specifying the connecting residues (chosen from previously entered residues in the residues/unit section) and the corresponding connecting position/site. Multiple connection entries can be added for heteropolysaccharide. Until here, user will finish a compound. To add more compounds from the same publication, use “add new compound” and repeat steps (3) – (5).

(6) Reference. The reference type needs to be specified as book chapters or journal articles. Information (title, authors, year etc.) can be entered using the corresponding forms. Both PubMed ID and DOI number are required for cross-referencing to other bibliographic databases.

(7) Review submission. A summary of the information will be provided to the users for review before completing the submission.

10.4.3 Current Status and Limitation

With the recent advances in instrumentation and methodology, ssNMR has become a powerful tool for characterizing the structure and dynamics of complex carbohydrates *in-vivo* and *in-situ*. Multidimensional ssNMR can efficiently identify the constituent monosaccharide units and partially reveal the pattern of covalent linkages of the polysaccharides [6]. The current studies mainly focus on the polysaccharides with highly repetitive structures, including the linear polymers of a single monosaccharide unit or multiple residue types but with an alternating pattern, as well as those polysaccharides with identical patterns of branching. Solving the full structure of more irregular carbohydrates remains challenging. At this stage, only high-resolution data (mostly with

2D/3D correlation spectra) are indexed in CCMRD, and we only accept data from peer-reviewed publications, with every record manually examined for reliability. With the rapid progress in carbohydrate ssNMR, we will also be able to develop statistics-based tools that automatically validate the new entries.

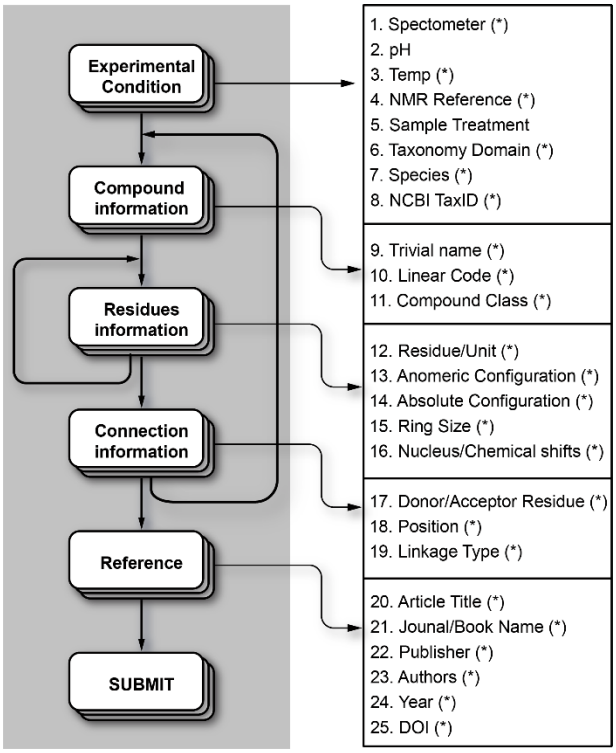

Figure 10.4 Flowchart of data deposition. The entries listed here will be needed for depositing data into CCMRD. The step-by-step user guide is shown in Figure S10.3.

10.5 Conclusion and Outlook

CCMRD presents a timely implementation to the existing databases of carbohydrates, such as Bacterial Carbohydrate Structure DataBase (BCSDB)[61] and the GLYCOSCIENCES.de portal[62], which collected data from traditional methods including x-ray, solution NMR and mass spectrometry. Together, these databases integrate and promote carbohydrate research and its relevant applications in the field of biomedical sciences, biomaterials and bioenergy. We will

continue updating CCMRD with trustworthy data sources and introducing auxiliary functions and software. The ongoing efforts include the implementation of new portals for including additional types of structural data (such as dynamics), the addition of web services in both json and xml formats, and the coding of novel functions for generating chemical shift heatmap[63] and artificial spectra (**Figure 10.2c**). These efforts will facilitate the statistical analysis of carbohydrate composition and structure, which will ultimately lead to high-throughput spectral analysis using various algorithms including deep-learning neural networks.

10.6 Acknowledgement

This work is supported by the National Science Foundation NSF OIA-1833040.

REFERENCES

1. Lowe, J. B.; Marth, J. D., A genetic approach to mammalian glycan function. *Annu. Rev. Biochem.* **2003,** *72* (1), 643-691.

2. Raman, R.; Raguram, S.; Venkataraman, G.; Paulson, J. C.; Sasisekharan, R., Glycomics: an integrated systems approach to structure-function relationships of glycans. *Nat. Methods* **2005,** *2* (11), 817.

3. Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D., Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **2007,** *315* (5813), 804-807.

4. Collins, M. N.; Birkinshaw, C., Hyaluronic acid based scaffolds for tissue engineering-A review. *Carbohyd Polym* **2013,** *92* (2), 1262-1279.

5. Kang, B.; Opatz, T.; Landfester, K.; Wurm, F. R., Carbohydrate nanocarriers in biomedical applications: functionalization and construction. *Chem Soc Rev* **2015,** *44* (22), 8301-8325.

6. Wang, T.; Phyo, P.; Hong, M., Multidimensional solid-state NMR spectroscopy of plant cell walls. *Solid State Nucl Mag* **2016,** *78*, 56-63.

7. Buchanan, B. B.; Gruissem, W.; Jones, R. L., *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, MD: 2000.

8. Latge, J. P., The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.* **2007,** *66* (2), 279-90.

9. Aspinall, G. O., *The Polysaccharides*. 1st ed.; Academic Press, New York, NY: 1983.

10. Lindberg, B., Components of bacterial polysaccharides. *Adv. Carbohydr. Chem. Biochem.* **1990,** *48*, 279-318.

11. Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; A., S., *Plant Cell Walls*. Garland Science, New York 2011.

12. Rytioja, J.; Hildén, K.; Yuzon, J.; Hatakka, A.; de Vries, R. P.; Mäkelä, M. R., Plant-polysaccharide-degrading enzymes from basidiomycetes. *Microbiol. Mol. Biol. Rev.* **2014,** *78* (4), 614-649.

13. Wormald, M. R.; Petrescu, A. J.; Pao, Y.-L.; Glithero, A.; Elliott, T.; Dwek, R. A., Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem. Rev.* **2002,** *102* (2), 371-386.

14. Sluiter, J. B.; Ruiz, R. O.; Scarlata, C. J.; Sluiter, A. D.; Templeton, D. W., Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *J. Agric. Food Chem.* **2010,** *58* (16), 9043-9053.

15. Duus, J. Ø.; Gotfredsen, C. H.; Bock, K., Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev.* **2000,** *100* (12), 4589-4614.

16. Bubb, W. A., NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concept Magn. Reson. A* **2003,** *19* (1), 1-19.

17. Cheng, K.; Sorek, H.; Zimmermann, H.; Wemmer, D. E.; Pauly, M., Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d6/1-ethyl-3-methylimidazolium acetate solvent. *Anal. Chem.* **2013,** *85* (6), 3213-21.

18. Mansfield, S. D.; Kim, H.; Lu, F. C.; Ralph, J., Whole plant cell wall characterization using solution-state 2D NMR. *Nat. Protoc.* **2012,** *7* (9), 1579-1589.

19. Atalla, R. H.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. *Science* **1984,** *223* (4633), 283-285.

20. Newman, R. H.; Davies, L. M.; Harris, P. J., Solid-State 13C Nuclear Magnetic Resonance Characterization of Cellulose in the Cell Walls of Arabidopsis thaliana Leaves. *Plant Physiol.* **1996,** *111* (2), 475-485.

21. Bardet, M.; Emsley, L.; Vincendon, M., Two-dimensional spin-exchange solid-state NMR studies of 13C-enriched wood. *Solid State Nucl Mag* **1997,** *8* (1), 25-32.

22. Renault, M.; Tommassen-van Boxtel, R.; Bos, M. P.; Post, J. A.; Tommassen, J.; Baldus, M., Cellular solid-state nuclear magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (13), 4863-8.

23. Phyo, P.; Wang, T.; Xiao, C.; Anderson, C. T.; Hong, M., Effects of Pectin Molecular Weight Changes on the Structure, Dynamics, and Polysaccharide Interactions of Primary Cell Walls of Arabidopsis thaliana: Insights from Solid-State NMR. *Biomacromolecules* **2017,** *18* (9), 2937-2950.

24. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** *17* (6), 2210-22.

25. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Poppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

26. Kang, X.; Kirui, A.; Muszyński, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9* (1), 2747.

27. Kang, X.; Kirui, A.; Widanage, M. C. D.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 347.

28. Takahashi, H.; Ayala, I.; Bardet, M.; De Paëpe, G.; Simorre, J. P.; Hediger, S., Solid-state NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. *J. Am. Chem. Soc.* **2013,** *135* (13), 5105-10.

29. Kim, S. J.; Chang, J.; Singh, M., Peptidoglycan architecture of Gram-positive bacteria by solid-state NMR. *Bba-Biomembranes* **2015,** *1848* (1), 350-362.

30. Thongsomboon, W.; Serra, D. O.; Possling, A.; Hadjineophytou, C.; Hengge, R.; Cegelski, L., Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. *Science* **2018,** *359* (6373), 334-338.

31. Werby, S. H.; Cegelski, L., Spectral comparisons of mammalian cells and intact organelles by solid-state NMR. *Journal of structural biology* **2018**.

32. Chatterjee, S.; Prados-Rosales, R.; Itin, B.; Casadevall, A.; Stark, R. E., Solid-state NMR Reveals the Carbon-based Molecular Architecture of *Cryptococcus neoformans* Fungal Eumelanins in the Cell Wall. *J. Biol. Chem.* **2015,** *290* (22), 13779-13790.

33. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water-Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis thaliana from Polarization Transfer Solid-State NMR. *J Am Chem Soc* **2014,** *136* (29), 10399-10409.

34. Grantham, N. J.; Wurman-Rodrich, J.; Terrett, O. M.; Lyczakowski, J. J.; Stott, K.; Iuga, D.; Simmons, T. J.; Durand-Tardif, M.; Brown, S. P.; Dupree, R.; Busse-Wicher, M.; Dupree, P., An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **2017,** *3* (11), 859-865.

35. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of Chlamydomonas reinhardtii microalgae. *J. Biol. NMR* **2018,** *70*, 123-131.

36. Cadars, S.; Lesage, A.; Emsley, L., Chemical shift correlations in disordered solids. *J Am Chem Soc* **2005,** *127* (12), 4466-76.

37. Arnold, A. A.; Genard, B.; Zito, F.; Tremblay, R.; Warschawski, D. E.; Marcotte, I., Identification of lipid and saccharide constituents of whole microalgal cells by 13C solid-state NMR. *Biochim. Biophys. Acta* **2015,** *1848* (1 Pt B), 369-77.

38. Laguri, C.; Silipo, A.; Martorana, A. M.; Schanda, P.; Marchetti, R.; Polissi, A.; Molinaro, A.; Simorre, J. P., Solid State NMR Studies of Intact Lipopolysaccharide Endotoxin. *ACS chemical biology* **2018,** *13* (8), 2106-2113.

39. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10* (1), 4978.

40. Kern, T.; Giffard, M.; Hediger, S.; Amoroso, A.; Giustini, C.; Bui, N. K.; Joris, B.; Bougault, C.; Vollmer, W.; Simorre, J. P., Dynamics characterization of fully hydrated bacterial

cell walls by solid-state NMR: evidence for cooperative binding of metal ions. *J Am Chem Soc* **2010,** *132* (31), 10911-9.

41. Kern, T.; Hediger, S.; Müller, P.; Giustini, C.; Joris, B.; Bougault, C.; Vollmer, W.; Simorre, J. P., Toward the characterization of peptidoglycan structure and protein-peptidoglycan interactions by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2008,** *130* (17), 5618-9.

42. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis Primary Cell Walls: Evidence from Solid-State Nuclear Magnetic Resonance. *Plant Physiol.* **2015,** *168* (3), 871-84.

43. Gan, Z. H.; Hung, I.; Wang, X. L.; Paulino, J.; Wu, G.; Litvak, I. M.; Gor'kov, P. L.; Brey, W. W.; Lendi, P.; Schiano, J. L.; Bird, M. D.; Dixon, L. R.; Toth, J.; Boebinger, G. S.; Cross, T. A., NMR spectroscopy up to 35.2 T using a series-connected hybrid magnet. *J Magn Reson* **2017,** *284*, 125-136.

44. Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Coperet, C.; Emsley, L., Dynamic Nuclear Polarization Surface Enhanced NMR Spectroscopy. *Acc. Chem. Res.* **2013,** *46* (9), 1942-1951.

45. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paepe, G., Rapid Natural-Abundance 2D 13C-13C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew Chem Int Edit* **2012,** *51* (47), 11766-11769.

46. Smith, A. N.; Long, J. R., Dynamic Nuclear Polarization as an Enabling Technology for Solid State Nuclear Magnetic Resonance Spectroscopy. *Anal Chem* **2016,** *88* (1), 122-132.

47. Sergeyev, I. V.; Itin, B.; Rogawski, R.; Day, L. A.; McDermott, A. E., Efficient assignment and NMR analysis of an intact virus using sequential side-chain correlations and DNP sensitization. *P Natl Acad Sci USA* **2017,** *114* (20), 5171-5176.

48. Saliba, E. P.; Sesti, E. L.; Scott, F. J.; Albert, B. J.; Choi, E. J.; Alaniva, N.; Gao, C. K.; Barnes, A. B., Electron Decoupling with Dynamic Nuclear Polarization in Rotating Solids. *J Am Chem Soc* **2017,** *139* (18), 6310-6313.

49. Andreas, L. B.; Le Marchand, T.; Jaudzems, K.; Pintacuda, G., High-resolution proton-detected NMR of proteins at very fast MAS. *J Magn Reson* **2015,** *253*, 36-49.

50. Struppe, J.; Quinn, C. M.; Lu, M. M.; Wang, M. Z.; Hou, G. J.; Lu, X. Y.; Kraus, J.; Andreas, L. B.; Stanek, J.; Lalli, D.; Lesage, A.; Pintacuda, G.; Maas, W.; Gronenborn, A. M.; Polenova, T., Expanding the horizons for structural analysis of fully protonated protein assemblies by NMR spectroscopy at MAS frequencies above 100 kHz. *Solid State Nucl Mag* **2017,** *87*, 117-125.

51. Phyo, P.; Hong, M., Fast MAS ^{1}H-^{13}C correlation NMR for structural investigations of plant cell walls. *J Biomol Nmr* **2019**.

52. Jaroniec, C. P., Structural studies of proteins by paramagnetic solid-state NMR spectroscopy.

J Magn Reson **2015,** *253*, 50-59.

53. Schmidt-Rohr, K.; Fritzsching, K. J.; Liao, S. Y.; Hong, M., Spectral editing of two-dimensional magic-angle-spinning solid-state NMR spectra for protein resonance assignment and structure determination. *J Biomol Nmr* **2012,** *54* (4), 343-353.

54. Wang, T.; Chen, Y.; Tabuchi, A.; Cosgrove, D. J.; Hong, M., The Target of b-Expansin EXPB1 in Maize Cell Walls from Binding and Solid-State NMR Studies. *Plant physiology* **2016,** *172* (4), 2107-2119.

55. Greenspan, J.; Bulger, B., *MySQL/PHP database applications*. John Wiley & Sons, Inc.: 2001.

56. Stauffer, M., *Laravel: up and running: a framework for building modern PHP apps*. " O'Reilly Media, Inc.": 2016.

57. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and Dynamics of Brachypodium Primary Cell Wall Polysaccharides from Two-Dimensional C-13 Solid-State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry-Us* **2014,** *53* (17), 2840-2854.

58. Thieker, D. F.; Hadden, J. A.; Schulten, K.; Woods, R. J., 3D implementation of the symbol nomenclature for graphical representation of glycans. *Glycobiology* **2016,** *26* (8), 786-7.

59. Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lutteke, T.; O'Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A.; Woods, R. J.; Group, S. D., Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **2019,** *29* (9), 620-624.

60. McNaught, A. D., Nomenclature of carbohydrates (IUPAC Recommendations 1996). *Pure Appl. Chem.* **1996,** *68* (10), 1919-2008.

61. Toukach, P. V.; Egorova, K. S., Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts. *Nucleic Acids Res.* **2015,** *44* (D1), D1229-D1236.

62. Böhm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M. A.; Lütteke, T., Glycosciences. DB: an annotated data collection linking glycomics and proteomics data (2018 update). *Nucleic Acids Res.* **2018,** *47* (D1), D1195-D1201.

63. Kirui, A.; Ling, Z.; Kang, X.; Dickwella Widanage, M. C.; Mentink-Vigier, F.; French, A. D.; Wang, T., Atomic resolution of cotton cellulose structure enabled by dynamic nuclear polarization solid-state NMR. *Cellulose* **2018**.

APPENDIX

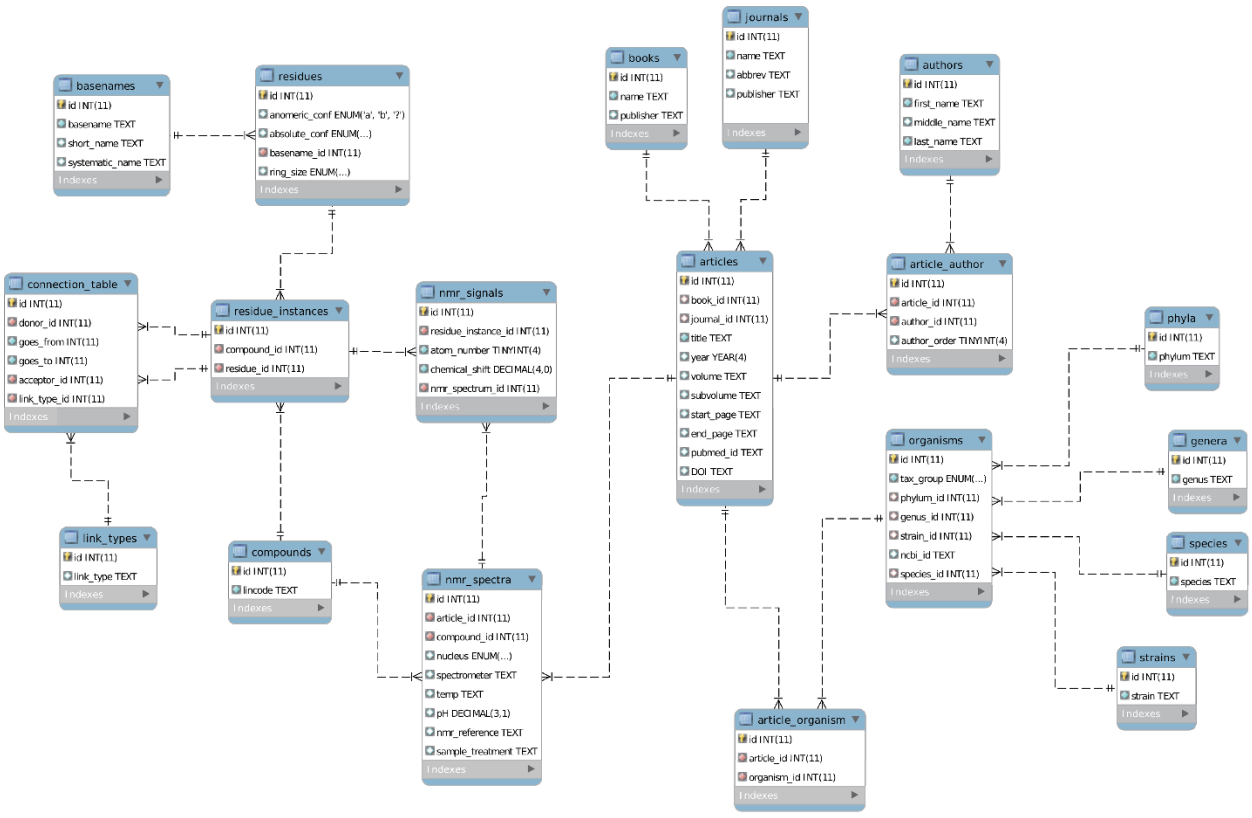

Figure S10.1 Organizational schema of CCMRD. This figure shows the primary organization and how the different entries (NMR signals, compounds, publications, species, etc.) are connected.

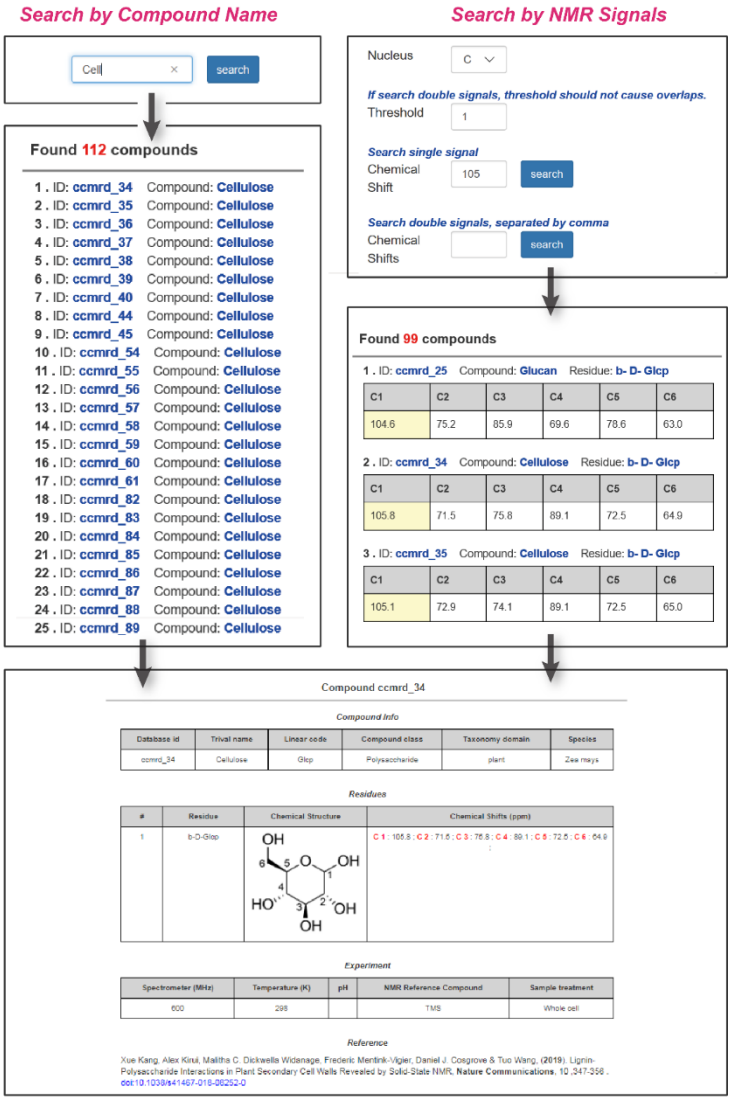

Figure S10.2 Data Search by compound name and NMR signals. Part or the full name of the polysaccharides is needed for searching by the compound name. In the example, 112 entries have been listed for searching cellulose. Two opinions are available for data searching by the NMR signals: a single NMR chemical shift or two chemical shifts in the same compound. The threshold represents the tolerance allowed for the search (e.g. 105 ± 1 ppm for the example shown in the figure). Both the search functions will finally lead to the details of each entry, with information on the compound trivial name, linear code, origin, species, NMR chemical shifts and experimental conditions.

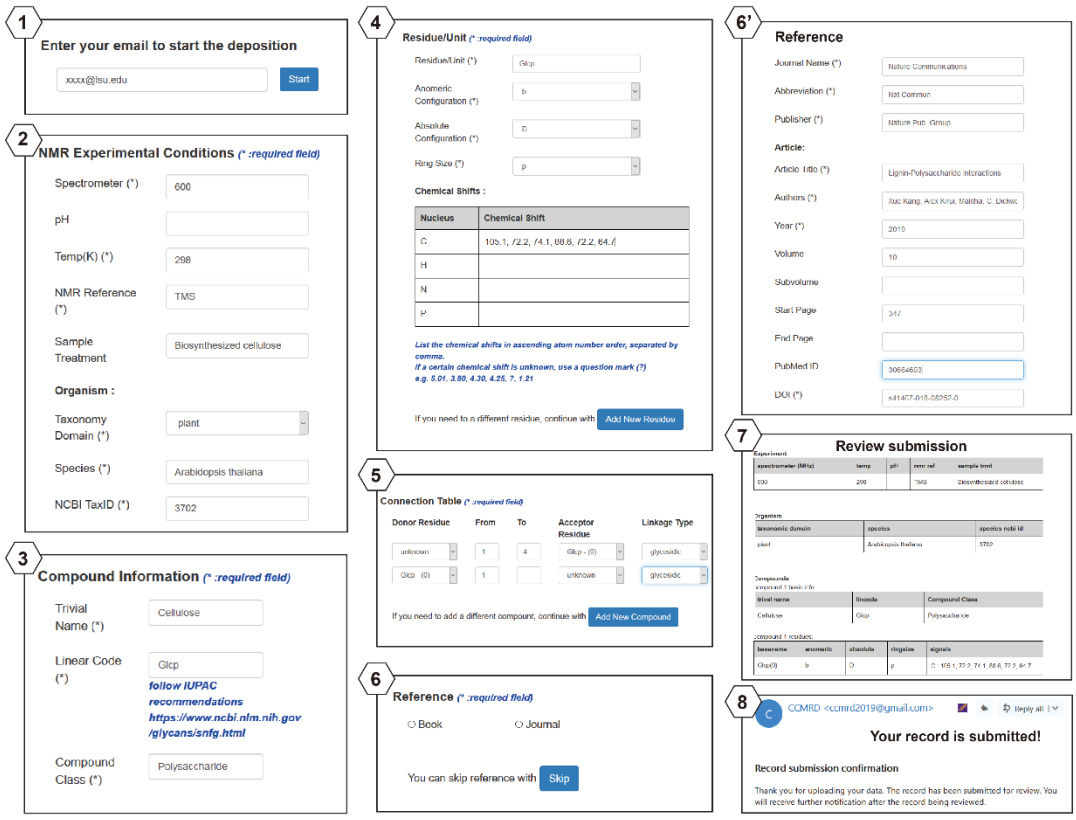

Figure S10.3 User deposition for CCMRD. The flow chart, interface, and an example of step-by-step deposition are given. The user deposition requires approval by the administrator before the record is indexed by the database.

CHAPTER 11: CHARTING THE SOLID-STATE NMR SIGNALS OF POLYSACCHARIDES: A DATABASE-DRIVEN ROADMAP

Research paper reprinted with permission: *Magn. Reson. Chem.*, in press (2023)

Authors: Wancheng Zhao, Debkumar Debnath, Isha Gautam, Liyanage D. Fernando, Tuo Wang

11.1 Abstract

Solid-state NMR (ssNMR) measurements of intact cell walls and cellular samples often generate spectra that are difficult to interpret due to the presence of many coexisting glycans and the structural polymorphism observed in native conditions. To overcome this analytical challenge, we present a statistical approach for analyzing carbohydrate signals using high-resolution ssNMR data indexed in a carbohydrate database. We generate simulated spectra to demonstrate the chemical shift dispersion and compare this with experimental data to facilitate the identification of important fungal and plant polysaccharides, such as chitin and glucans in fungi and cellulose, hemicellulose, and pectic polymers in plants. We also demonstrate that chemically distinct carbohydrates from different organisms may produce almost identical signals, highlighting the need for high-resolution spectra and validation of resonance assignments. Our study provides a means to differentiate the characteristic signals of major carbohydrates and allows us to summarize currently undetected polysaccharides in plants and fungi, which may inspire future investigations.

11.2 Introduction

Polysaccharides are high molecular weight biopolymers that typically consist of hundreds to thousands of sugar units. Elucidating the structure of complex carbohydrates with solid-state NMR (ssNMR) spectroscopy is a challenging task. Early pioneers in this field explored the use of different ssNMR methods, such as one-dimensional (1D) experiments for identifying the structure and quantifying the content of carbohydrate polymers and allomorphs, and Rotational Echo

Double Resonance (REDOR) and multidimensional correlation techniques for revealing polymer cross-linking and spatial proximities[1-6]. There has been a substantial increase in high-resolution ssNMR data that rely on ^{13}C and ^{15}N 2D/3D correlation experiments[7-9], as well as ^{1}H detection[10-12], to provide detailed insights into the structure and function of carbohydrates in a variety of living organisms.

Recent high-resolution ssNMR studies have led to numerous conceptual breakthroughs, including but not limited to the identification of cellulose-pectin spatial contacts in the primary plant cell walls (mostly the seedlings)[13-16], the conformation-dependent function of xylan in associating with lignin and cellulose in the secondary cell walls of grasses and woody stems[17-22], the multi-layered assembly of chitin and glucans in fungal cell walls that provide remarkable structural dynamics to support microbial survival[23-28], as well as the quantification of dynamically heterogeneous carbohydrate components in algal cells and bacterial cell wall and biofilms[29-33]. There is also a growing interest in utilizing ssNMR techniques for the characterization of carbohydrate-based materials, including hydrogels[34, 35].

By taking advantage of the repetitive nature of monosaccharides in a polysaccharide, the structural view of these polymers can be simplified, and their spectral features can be explored. Currently, there are five distinct structural variations of carbohydrates that can be closely tracked in the solid-state, which include the variation in covalent linkages, the compositional differences in cell walls and monosaccharide units that form polysaccharides, the conformational distribution of native polysaccharides, the spatial assembly, and the chemical modifications such as methylation and methyl esterification, as documented in a recent review aricle[36]. The majority of existing ssNMR studies depict polysaccharides in a simplified manner, reducing their complex structures to symbolic representations and only highlighting the most significant constituents

showing strong signals[7]. While this approach enables researchers to concentrate on the primary polysaccharides found in a sample, it overlooks the many sparsely distributed molecules, minor side chains, and irregular substitutions, which may also contribute significantly to the structural stability and functional properties of both the polysaccharide and the cell wall.

In 2020, a complex carbohydrate magnetic resonance database (CCMRD) was established to accommodate the prolific growth of datasets in the upcoming years[37]. Upon its inception, the small-scale database contained 420 entries of high-resolution ssNMR data on carbohydrates. As of early 2023, the number of entries has increased to approximately 850. The database facilitates data search using single or paired chemical shifts, polysaccharide names, and compound classes, and it also supports user deposition. This platform also serves as a vital preparatory step for implementing semi-automated tools to streamline spectral analysis and resonance assignment of carbohydrate systems in the future as the size of the dataset increases.

The present study aims to accomplish three objectives by utilizing the available data. Firstly, we aim to construct artificial NMR spectra of commonly employed experiments to summarize the polysaccharide signals present in plants and fungi, which is akin to the amino acid maps available for protein NMR[38]. Secondly, we will utilize these spectroscopic maps to identify the unique signals that can unequivocally recognize specific polysaccharides. Finally, we will provide a synopsis of the key factors currently constraining our ability and give an overview of the undetected carbohydrates that necessitate further exploration. The pursuit of these objectives intends to encourage the use of high-resolution ssNMR in carbohydrate characterization.

11.3 Materials and Methods

11.3.1 Generation of Simulated 2D Spectra from Chemical Shift Data

We used a dataset containing 569 entries of plant and fungal carbohydrates sourced from CCMRD[37] to construct artificial 2D NMR spectra. Each entry within the dataset comprised ^{13}C chemical shift data for a carbohydrate unit sourced from a peer-reviewed publication. These publications featured high-resolution solid-state NMR data, typically presented in the form of 2D or 3D correlation spectra. The raw ^{13}C chemical shift data were transformed into different scatter plots depending on the targeted type of 2D spectra following the protocol described below.

For the refocused INADEQUATE spectrum that exhibits through-bond correlations[4], the signals of two directly bonded carbons were situated at coordinates (δ_{Ca}, $\delta_{Ca}+\delta_{Cb}$) and (δ_{Cb}, $\delta_{Ca}+\delta_{Cb}$) on the plot. Here, δ_{Ca} and δ_{Cb} denote the single-quantum (SQ) chemical shifts of carbon sites a and b, respectively, and these values were projected to the X-axis of the simulated plot, corresponding to the direct dimension (ω_2) of experimentally measured spectra. The sum of two SQ chemical shifts ($\delta_{Ca}+\delta_{Cb}$) were projected to the Y-axis, which represents the value of the double-quantum (DQ) chemical shift of this pair of carbons, corresponding to the indirect dimension (ω_1) of experimentally measured spectra. Each pyranose unit is represented by 10 scatter dots based on the chemical shifts of six carbons, and an illustrative example is presented in **Figure 11.1a**, showcasing the distribution of these scatter dots for a glucose unit in cellulose. In comparison, each furanose unit will result in 8 scatter dots, due to the lack of C5-C6 spin pair.

More scatter dots are needed for plotting 2D ^{13}C SQ-SQ correlation spectra, such as PDSD/DARR, CORD[39], and PAR[40]. This is because each carbon atom can potentially correlate with all the other carbons within the same monosaccharide unit. Each scatter dot now represents a distinct intramolecular cross peak, with the two coordinates representing the SQ chemical shifts of

two carbons. When examining a pyranose, the anticipated SQ-SQ correlation spectrum consists of 36 derived data points, as illustrated in **Figure 11.1b** for a glucose unit of cellulose. Among these points, 6 are positioned along the spectral diagonal. In contrast, a furanose's spectrum contains 25 data points, with 5 residing on the diagonal. The chemical shift coordinates utilized in generating the scatter plots shown in **Figure 11.1** are provided in **Table S11.1**.

Our analysis extended to encompass 412 entries from plant cell walls and 157 datasets from fungal cell walls obtained from CCMRD. Scatter plots were generated to depict both DQ-SQ and SQ-SQ correlation spectra for various sugar units. Subsequently, density maps were generated from these scatter plots using Kernel Density Estimate in Origin 2019b. These density maps provide an informative overview of the average trends represented by the scatter plots, allowing us to identify the specific regions and patterns related to different sugar units within 2D NMR spectra. **Figures S11.1-S11.26** in the Supplementary Information present individual plots for various plant and fungal carbohydrates. These figures can be used for direct comparison with experimental data, streamlining the analysis process.

11.3.2 Sample Preparation and Packing

Three uniformly ^{13}C-enriched samples were prepared for ssNMR measurement. These samples include the fungal culture of *Aspergillus fumigatus*, and the stems of two plant species: *Arabidopsis thaliana* and *Picea abies* (Norway spruce)[21]. The fungal mycelium was cultivated for two weeks on a liquid minimal medium containing ^{13}C-glucose as the only carbon source[26]. The spruce plant was grown within a closed atmosphere featuring 97 atom% $^{13}CO_2$. This controlled environment was maintained from germination until harvest, spanning approximately four months. Similarly, ^{13}C-enriched *Arabidopsis* was grown in a custom-built transparent chamber for 4 weeks. The harvested fresh plant stems were cut into millimeter-scale pieces by razor blade for better

distribution of mass during MAS.

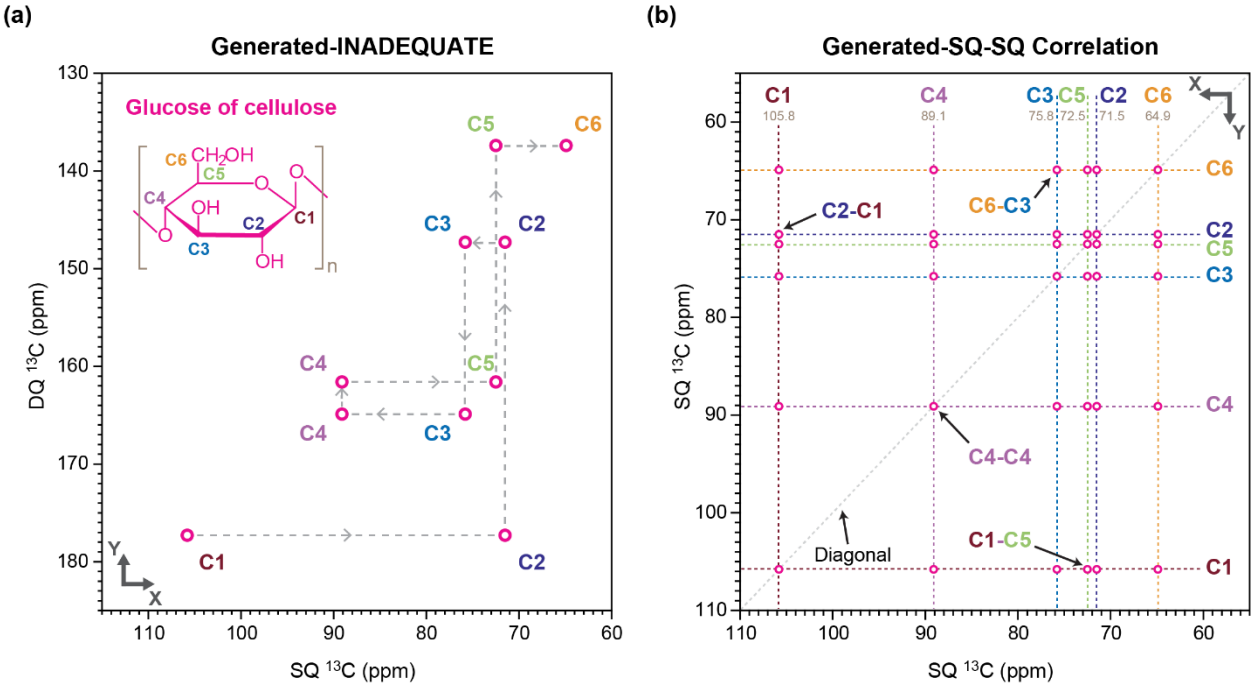

Figure 11.1 Scatter plots of chemical shift values for a glucose unit of cellulose. The plots correspond to **a,** DQ-SQ and **b,** SQ-SQ ^{13}C correlation experiments, with the former showing through-bond correlations and the latter showing all intramolecular through-space correlations. See Table S11.1 for comprehensive details on the dataset used to create these spectral projections.

Isotope-labeling is essential for acquiring high-resolution 2D/3D correlation spectra, particularly in dealing with the significant signal overlap encountered in cellular samples or complex biomaterials. Natural-abundance 2D ^{13}C-^{13}C correlation experiments of unlabeled samples becomes attainable with the utilization of the sensitivity-enhancing technique MAS-DNP, as recently demonstrated on lignocellulosic plant biomass[41-43], fungal mycelia and conidia[44], and functionalized cellulose materials[45-48]. However, it's important to note that at the low temperatures required for MAS-DNP, there is a significant sacrifice in spectral resolution. Therefore, this technique is primarily employed when isotope-labeling presents challenges.

Around 30 mg of *A. fumigatus* was packed into a 3.2 mm MAS rotor and measured on an 800 MHz Bruker Avance Neo Spectrometer at National High Magnetic Field Laboratory using a 3.2 mm HCN MAS probe. Similarly, around 35 mg of *Arabidopsis* sample was packed into a 3.2 mm MAS rotor and measured on a 700 MHz Bruker Avance Neo Spectrometer at Louisiana State University using a 3.2 mm HCN MAS probe. For spruce, approximately 100 mg of material was packed into a 4 mm MAS rotor for measurement on a 400 MHz Bruker Avance spectrometer at Louisiana State University using a 4.0 mm HXY MAS probe.

11.3.3 Solid-State NMR Experiments

All experiments were performed with spinning speeds ranging from 10-13.5 kHz. The radiofrequency field strengths were set to 83.3 kHz for both ^{1}H excitation and decoupling, and 62.5 kHz for ^{13}C hard pulses. The recycle delays were 1.6-2.0s. 2D experiments were recorded over durations of 4 to 30 h, while 1D data were collected in timeframes varying from 5 to 140 min. The 2D refocused J-INADEQUATE scheme was integrated with either direct polarization (DP) and short recycle delays of 1.8 s for the selective detection of mobile components, or with cross-polarization (CP) for the detection of rigid components. A 2D ^{13}C-^{13}C correlation experiment was also measured using the CORD scheme with a 53-ms recoupling period. Within each spectrum, both the sidechains and primary constituents of a polysaccharide, if they exhibit similar dynamics, were detected simultaneously. The acquisition time was 12-29 ms, and the evolution time for the indirect dimension of 2D experiments was 4-10 ms. The key parameters used in these ssNMR experiments are listed in **Table S11.2**.

Accurate experimental chemical shift data are essential for structural analysis[49]. All ^{13}C NMR spectra are consistently referenced to the tetramethylsilane (TMS) scale by externally calibrating the CH_2 signal of adamantane to 38.48 ppm at room temperature, which is a standard

calibration procedure widely employed in solid-state NMR studies[50], including carbohydrate research. In the chemical shift data deposited in CCMRD, NMR chemical shifts are only reported on the TMS scale, the information of which is included in each entry[37]. DSS scale is another reference scale that yields a ^{13}C chemical shift difference of 2.0 ppm compared to the TMS scale. Chemical shifts reported on the DSS scale can be adjusted by subtracting 2.0 ppm for the purpose of comparison with data on the TMS scale, which has been reported in recent studies of amyloid fibrils[51]. However, it finds primary usage in protein solid-state NMR and solution NMR studies. It should be noted that variations in chemical shifts may arise due to different calibration procedures employed by various research groups, as reported in recent cellulose studies[49].

11.4 Results and Discussion

11.4.1 Unique Signals of Different Carbohydrates

Initial assignments of carbohydrate signals heavily rely on 2D ^{13}C refocused INADEQUATE correlation experiments[4] that correlate DQ and SQ chemical shifts to obtain diagonal-free spectra. We extracted 412 plant carbohydrate entries and constructed artificial INADEQUATE spectra (**Figure 11.2a**). The scatter plots were derived from data deposited in CCMRD, with all through-bond correlations simulated for each carbohydrate unit. For example, for each hexose, the DQ and SQ chemical shifts were calculated for six spin pairs of C1-C2, C2-C3, C3-C4, C4-C5, C5-C6, while the last spin pair is absent in each pentose. Furthermore, we created corresponding color-coded contour lines based on computed Kernel Density Estimate to represent the probability of different sugar residues in certain spectral regions. These maps included sugar residues found in context of a plant polysaccharide, such as the glucose units (Glc) found in cellulose and seven sugar residues commonly found in the matrix polysaccharides (hemicellulose and pectin), namely xylose (Xyl), galacturonic acid (GalA), glucuronic acid

(GlcA), arabinose (Ara), galactose (Gal), rhamnose (Rha), and mannose (Man) (**Figure 11.2b**).

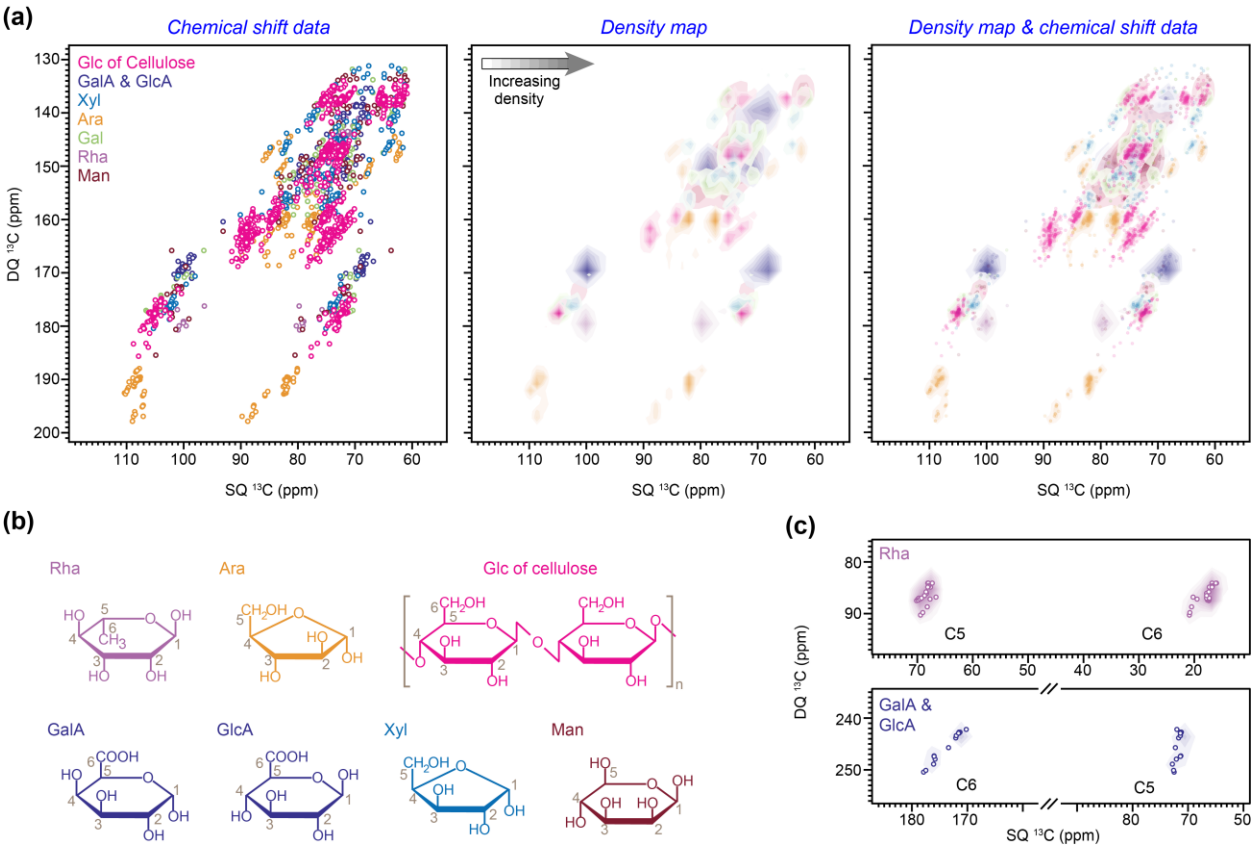

Figure 11.2 Density map of plant polysaccharides projected into INADEQUATE spectra. **a,** ^{13}C chemical shift data of 412 datasets plotted as either scatter plot (left) or density map (middle), or both (right). The data of eight types of monosaccharide units were color coded. **b,** Representative monosaccharide units with carbon numbers labeled. Two glucose units were shown for cellulose. **c,** Unique spectral regions of the C5-C6 signals of Rha (top) and acidic residues (bottom), where the methyl and carbonyl groups can help to track the signals of these carbohydrates.

In particular, the Ara units are typically resolved by their C1-C2 spin pair, with unique DQ chemical shifts ranging from 188 ppm to 198 ppm. Although the cellulose signals are tightly clustered with peaks from other carbohydrates, their C3-C4 and C4-C5 spin pairs can still be used as the starting point to track their signals. Xyl signals have a broad distribution but can be tracked

using the C4-C5 spin pair: α-linked Xyl has a unique DQ chemical shift at 130-135 ppm while β-linked Xyl typically shows up at 140-150 ppm, although the latter depends on the helical screw conformation. Rha can be unambiguously tracked starting from its C6, which is a methyl group, and the corresponding C5-C6 spin pair (**Figure 11.2c**). Acidic sugars, such as GalA and GlcA, rely on the spin pair between C5 and C6, which is the carbonyl group, for initiating the resonance assignment and tracking the other carbons within the same sugar unit.

It should be noted that the chemical shift of the carbonyl groups in acidic sugars can vary significantly depending on the state of protonation, the deprotonation and ion coordination (e.g., with Calcium in HG and with Boron in RG-II)[15, 52], as well as the degree of methyl esterification. Peak multiplicity in this spectral region has been consistently observed in many plant species, such as *Arabidopsis*, *Zea mays*, and *Brachypodium*[16, 53-55], with 176-178 ppm for carboxylate ($-COO^-$ or -COOH), 173-175 ppm for acetyl ($-OCOCH_3$), and 170-172 ppm for methyl ester ($-COOCH_3$). Due to the low bulk pKa of GalA (~3.5)[56], the carboxylate is in the deprotonated form under the near-neutral pH of most plant cell wall samples. However, the carboxylate signal vanishes GalA residues are neutralized during cell wall acidification, which has been observed at pH 4.0 in an *Arabidopsis* sample[57]. This further impedes GalA's capability to coordinate Ca^{2+} for crosslinking adjacent HG chains. Therefore, the pKa of these sugars and the pH of the samples being investigated should be considered and the carbonyl signals should be carefully tracked.

Signals of Gal residues are difficult to separate from the peaks of other sugar residues when using 2D ^{13}C-^{13}C correlation spectra; therefore, it requires special attention in analysis. The SQ chemical shift of carbon 4 site (~78 ppm) and the DQ chemical shifts of C3-C4 and C4-C5 (152-155 ppm) offer partial resolution in the refocused INADEQUATE spectrum, serving as an initial reference for Gal resonance assignment, especially for those in the mobile phase[55]. Proton-

detection techniques such as 2D TOCSY and 3D HCC INEPT-TOCSY have been employed to distinguish Gal signals from other sugar residues in mobile matrix polysaccharides[10].

We also presented these datasets as a 2D ^{13}C SQ-SQ correlation spectrum, as shown in **Figure 11.3**. This spectrum encompasses all intramolecular cross-peaks, resembling the spectral patterns commonly observed with ~100 ms mixing periods under experimental schemes like CORD[39], DARR/PDSD, or other analogous approaches conducted at moderately slow magic-angle spinning (MAS) frequencies, e.g., in the 10-20 kHz range. The simulated spectrum is more congested than the INADEQUATE spectra shown earlier due to the presence of numerous multibond correlations. Nevertheless, discernable spectral regions are still available for almost every type of carbohydrate, except for Gal, where complementary techniques like the fast MAS INEPT spectra with ^{1}H detection, as demonstrated in a recent study, may become necessary[10].

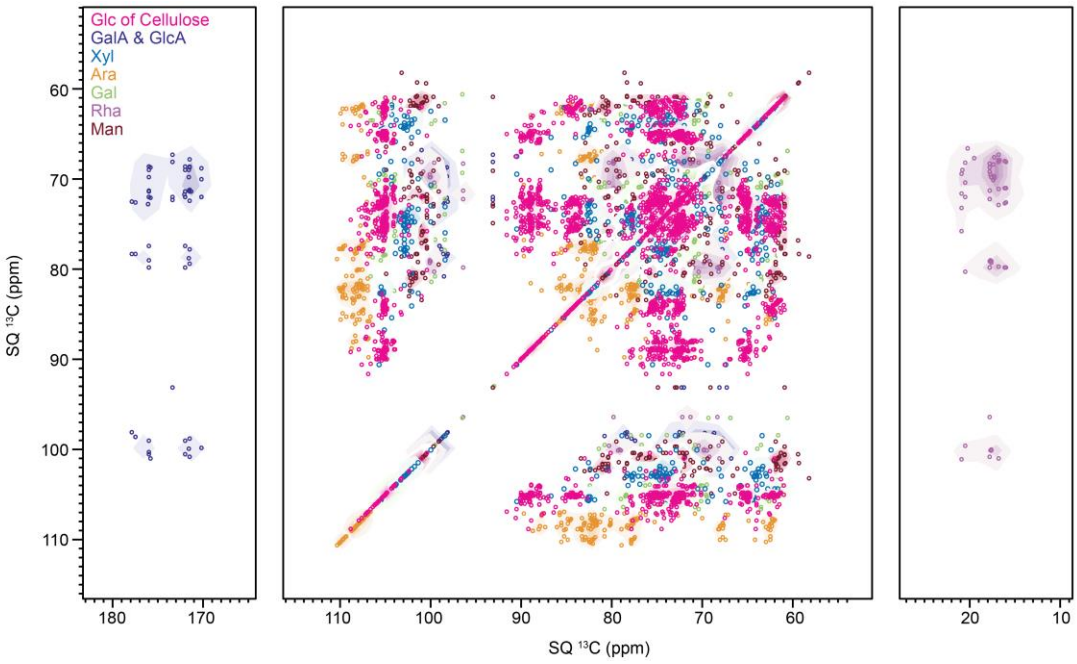

Figure 11.3 Projected NMR signals of plant polysaccharides as the 2D SQ-SQ correlation spectra using the same datasets for plotting Figure 11.2.

11.4.2 Reconstituted Spectra for Polysaccharide Identification

Monosaccharide units can associate with others to form various polysaccharides. We synthesized artificial spectra to depict the potential range of chemical shift dispersion for polysaccharides (**Figure 11.4a**). The first spectrum cataloged all the signals identified in plant cellulose, which possesses a highly polymorphic structure characterized by up to seven types of Glc subforms within each plant sample[58]. The second spectrum encompassing Xyl and Ara signals, contained the spectra from various matrix polysaccharides, including hemicellulose arabinoxylan, pectic sidechain arabinan, as well as the sidechains of xyloglucan. Notably, the α-Xyl units of xyloglucan and β-Xyl residues of xylan were not differentiated in this plot, but the distinction can be easily achieved, if necessary. The third spectrum displayed the signals of acidic residues and Rha, which covered two primary pectic polymers: homogalacturonan (HG) and rhamnogalacturonan-I (RG-I), as well as the GlcA sidechains of the hemicellulose glucuronoxylan. The last spectrum displayed Gal and Man residues, which could be part of mannan polymers like galactoglucomannan (GGM), or the galactan sidechains of RG-I. The corresponding carbohydrate structures are summarized in **Figure 11.4b**.

The comparison of simulated carbohydrate maps with experimentally measured 2D spectra can aid in the initial evaluation of the sample and facilitate the identification of the most probable carbohydrates. As an example, we analyzed the refocused J-INADEQUATE spectrum of ^{13}C-labeled *Arabidopsis* primary cell walls, which were measured using DP and short recycle delays of 1.8 s to select the mobile molecules with rapid ^{13}C-T_1 relaxation (**Figure 11.5a**). Consequently, only dynamic matrix polysaccharides were detected, including the well-resolved signals of Rha, acidic sugars, and Ara from pectin, as well as the signature C4-C5 signals of α-linked Xyl units from xyloglucan.

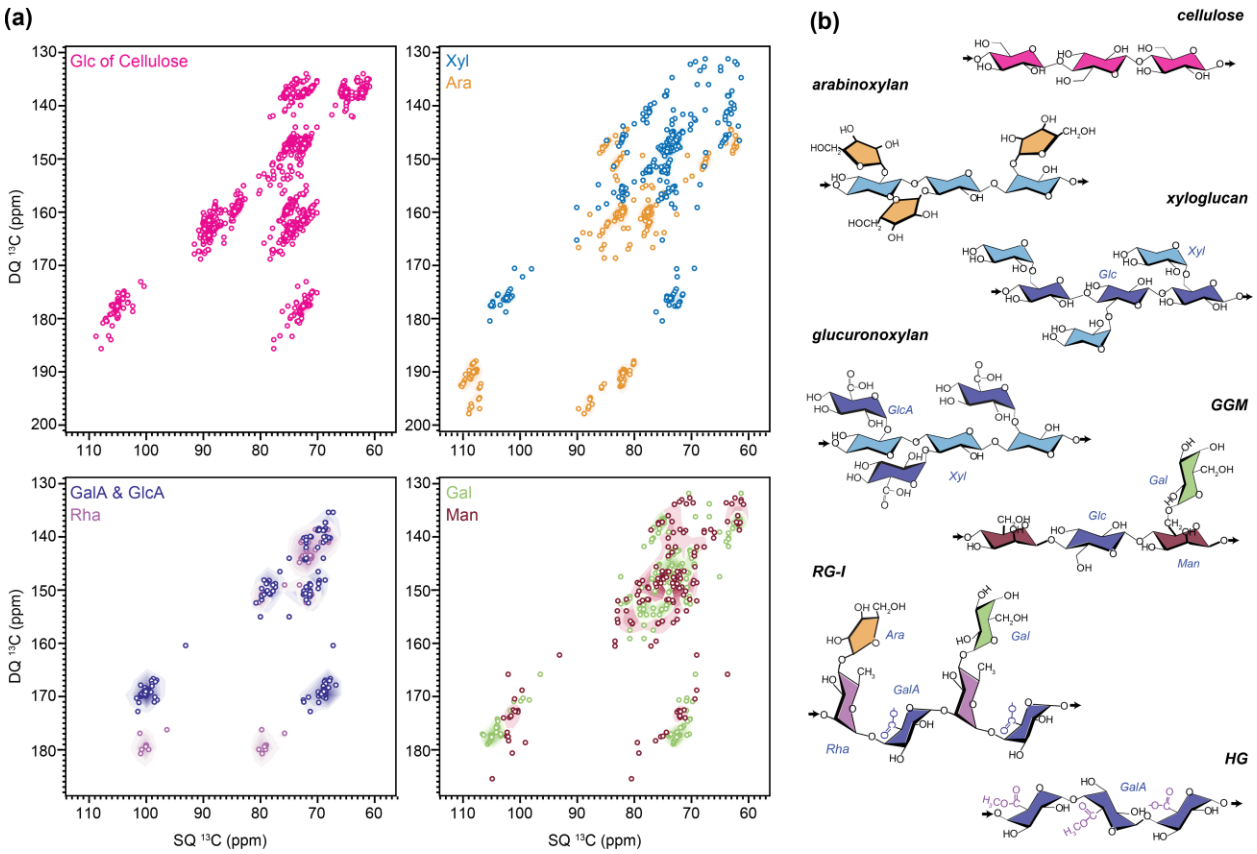

Figure 11.4 Simulated NMR spectra of individual carbohydrate units. **a,** Chemical shift dispersion and density maps of cellulose glucose units, Xyl, Ara, GalA and GlcA, Rha, Gal and Man. **b,** Simplified structure of polysaccharides containing these residues widely found in plant primary and secondary cell walls.

Ara units are distinguishable by their distinct signals falling within the DQ ^{13}C chemical shift range of 188-200 ppm. Ara peaks are also evident in all other discernible spectral regions indicated by the yellow regions in the simulated density maps, which include the DQ chemical shift range of 160-170 ppm for C2-C3 and C3-C4, as well as the DQ chemical shift range of 145-155 ppm for C4-C5. These observations confirmed the presence of Ara residues in *Arabidopsis*. Similarly, the presence of Rha is verified by its characteristic C5-C6 signals, where C6 represents a methyl carbon, situated within the DQ chemical shift range of 80-90 ppm. The existence of acidic

residues (GalA/GlcA) is substantiated by their unique signals with DQ chemical shifts extending beyond 240 ppm, corresponding to the covalently linked C5 and carbonyl group.

Cellulose is not detected in this spectrum, as evidenced by the absence of peaks in the magenta regions, where DQ chemical shifts typically fall within the range of 160-170 ppm, a characteristic range for C3-C4 and C4-C5 carbon pairs in cellulose. This outcome aligns with expectations since cellulose is a rigid component and is not found in this spectrum, which is designed to detect mobile components. Intensities are observed in other magenta regions, but these sharp signals originate from mobile polysaccharides exhibiting similar chemical shifts at these indistinguishable carbon positions.

The rigid portion of the spruce secondary cell wall was examined using ^{13}C CP-based INADEQUATE and CORD experiments, where the spectra were dominated by cellulose signals (**Figure 11.5a, b**). Weak peaks were also spotted for the β-linked Xyl from xylan backbones and the Ara residues from xylan sidechains, as well as some Man signals from GGM. The α-Xyl residues, which are indicative of xyloglucans, are not observed in this spectrum as shown by the absence of C4-C5 signals within the DQ chemical shift range of 130-136 ppm. This outcome is anticipated, considering that the spruce sample under examination is primarily composed of secondary cell wall component and the current spectrum selectively targets rigid polysaccharides, whereas xyloglucan, being partially mobile, exists in primary cell walls.

The use of carbohydrate fingerprints and the combination of selected datasets from certain glycans or organisms turned out to be an efficient approach for identifying polysaccharides, which should be applicable to both purified biopolymers and cellular samples. We have included a collection of individual scatter and density plots as **Figures S11.1-S11.26** for the purpose of enabling direct comparisons with experimentally acquired spectra. This qualitative analysis will

aid in identifying potential carbohydrates in the plant sample of interest as the initial step, which will require further validation or complete resonance assignment as subsequent procedures. As the *Arabidopsis* and spruce spectra were collected on 400-700 MHz NMR instruments, it is evident that a moderate-to-high magnetic field is adequate for the implementation of this method. A current limitation is the lack of quantitative methods to assess the probability of specific resonance assignments, a capability akin to what is available for proteins and small molecules[59, 60]. Further development of new software and algorithms[61, 62] will be essential.

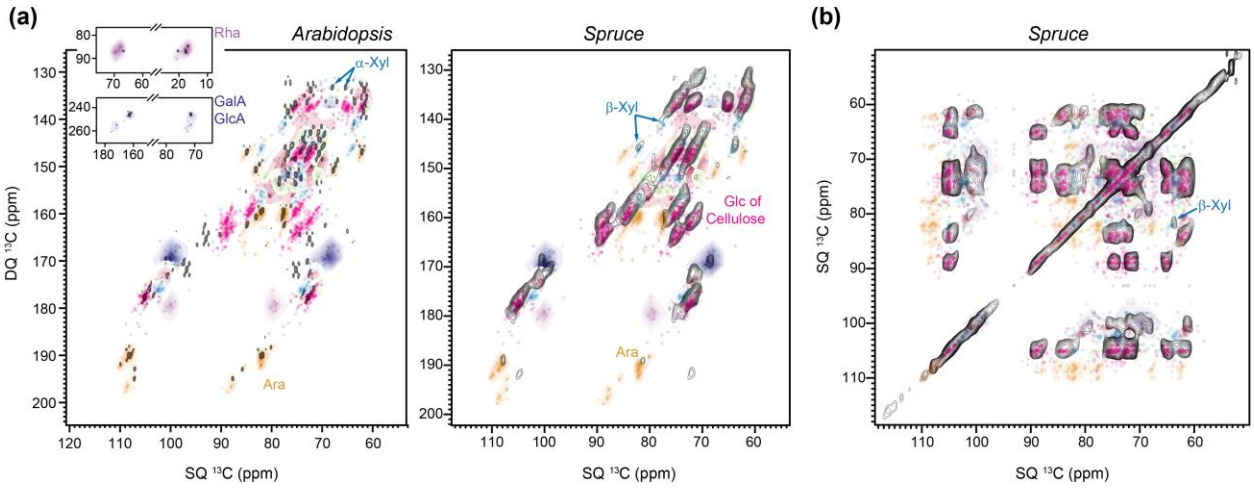

Figure 11.5 Overlay of carbohydrate density map and experimental spectra for **a,** refocused J-INADEQUATE and **b,** 53 ms CORD spectra. The DP-based J-INADEQUATE spectrum measured on *Arabidopsis* primary cell walls, and the CP-based J-INADEQUATE and CORD experiments were conducted on spruce stems that are rich in secondary cell walls. *Arabidopsis* and spruce samples were measured on 700 MHz and 400 MHz NMR spectrometers, respectively.

11.4.3 Plant Carbohydrates Evading Detection

Many polysaccharides remain elusive to solid-state NMR detection, even for the polysaccharides in plants, which have been relatively well studied. Although xyloglucan (XyG) is the primary hemicellulose in the primary cell wall, most studies have to rely on its α-Xyl sidechains

to track XyG signals and examine interactions with other cell wall components[13, 63]. Its β-1,4-glucan backbone, as well as the occasional substitutions of β-galactose (Gal) and α-fucose (Fuc)[64], have not been unambiguously identified. The signals corresponding to the β-1,4-glucan backbones are predominantly shielded by the surface glucan chains of cellulose, which are chemically identical, while the Gal and Fuc residues are relatively scarce. In addition, the third abundant pectic polymer, RG-II, has not been detected in intact cell wall samples either. RG-II contains at least 12 sugar types with complex linkages and has a relatively lower content compared with HG and RG-I, making it very difficult to characterize[65, 66]. Moreover, chemical data suggest that pectic polymers exist in integrated networks containing polysaccharides and proteoglycans and might be covalently interconnected and further crosslinked to other cell wall macromolecules such as hemicellulose and protein components[66-68]. At this moment, such structural features have not been recognized using intact cell walls. The integration of solid-state NMR analysis with functional genomics, incorporating engineered mutants that selectively remove individual carbohydrate components, holds promise for the unambiguous identification of previously unresolved polysaccharides. This approach has recently proven successful in fungal species, particularly in resolving the ambiguity associated with the identification of galactosaminogalactan (GAG) and galactomannan (GM) within the cell walls of *Aspergillus* species[25]. Augmenting this approach with MAS-DNP can further mitigate detection issues associated with the low levels of certain carbohydrate components, such as RG-II and the Gal and Fuc sidechains of xyloglucan. Improvement in 2D/3D correlation experiments to achieve better signal dispersion[69, 70], coupled with the assessment of ^{1}H chemical shifts using ultrafast MAS techniques[10, 71], are poised to surmount the existing resolution limitations.

11.4.4 Limited Dataset of Fungal Carbohydrates

Fungal glycans have received inadequate attention, and most of the available data has been collected in the past five years, as summarized in **Figure 11.6**. We can unambiguously identify the signals from several amino sugars, such as the N-acetylglucosamine (GlcNAc) from chitin and the glucosamine (GlcN) from chitosan. Both chitin and chitosan have exhibited broad distribution owing to the high level of structural polymorphism in their crystallites[72, 73]. However, the relationship between the polymorphism of chitin and chitosan and the diversity of chitin synthase gene families remains unclear[74]. β-glucans are prominent molecules widely distributed in different fungi, and have a high level of linkage complexity and can exist as different structures, such as linear β-1,3-glucan, branched β-1,3/1,6-glucan, linear β-1,3/1,4-glucan, and linear β-1,6-glucan, depending on the location within the cell wall and the fungal species being studied[75, 76]. Different linkages can exist in the same fungi with variable percentages. For example, the β-glucans of yeast consist (~85%) of branched β-1,3-glucan, which contains ~3% β-1,3,6 interchain branching sites, with a smaller amount of linear β-1,6-glucans (~15%)[76]. Meanwhile, the filamentous mold *Aspergillus fumigatus* also have predominantly branched β-1,3-glucan accounts with ~4% β-1,6 branching sites, but the minor component has changed to β-1,3/1,4-glucan, which constitutes approximately 10% of the cell wall β-glucans in this species[75]. Solid-state NMR studies mainly focus on the function of β-1,3-glucan and ongoing research aims to identify the signals of each linkage type and relate them to their structural functions in the cell wall. Recent solid-state NMR studies have mainly focused on the function of β-1,3-glucan and ongoing research aims to identify the signals of each linkage type and relate them to their structural functions in the cell wall.

Recent studies have revealed the roles of α-1,3-glucan in stabilizing the cell wall assembly, but information regarding the α-1,4-linked glucose residue is lacking[77]. Other spotted

monosaccharide units, such as galactofuranose (Gal*f*), galactosamine (GalN), N-acetylgalactosamine (GalNAc), Man, and Fuc, belong to galactosaminogalactan (GAG), galactomannan (GM), phosphomanan, mannoproteins, and other minor cell wall polysaccharides, which can further complicated spectral analysis. The contribution of glycoproteins and lipid components to the cell wall architecture requires further investigation[78].

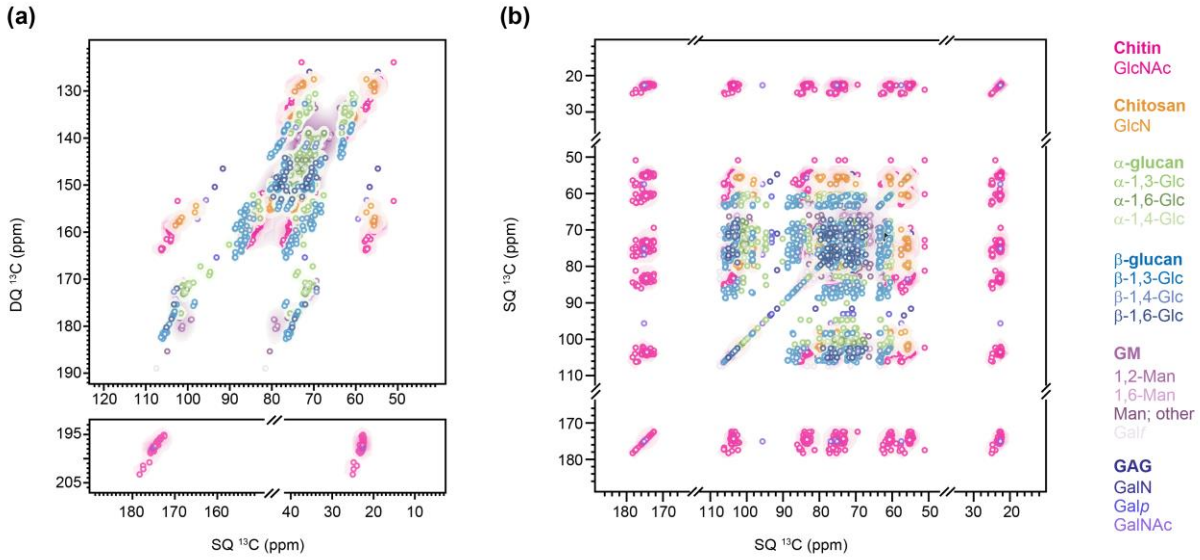

Figure 11.6. Summary of ^{13}C chemical shifts from fungal carbohydrates. The simulated spectra were presented as **a,** DQ-DQ and **b,** SQ-SQ correlation spectra. The current plot contains data from 157 entries, which is relatively limited and cannot efficiently represent the complex nature of fungal polysaccharides. The name of each polysaccharide is shown in bold followed by the monosaccharide units including GlcNAc, GlcN, Glc, Man, Gal*f*, GalN, Gal*p*, and GalNAc.

Compared to their plant counterparts, fungal polysaccharides exhibit much higher structural diversity, and extensive solid-state NMR studies are necessary to document the carbohydrate signals across a wide range of fungal species to facilitate antifungal development and the biotechnology applications of these microorganisms. These endeavors may encompass the

utilization of multidimensional correlation experiments involving ^{13}C, ^{15}N, and ^{1}H nuclear spins to elucidate the carbohydrate signals in significant human pathogenic species, such as *Aspergillus*, *Candida*, and *Cryptococcus* species, as well as non-pathogenic species that could potentially be used for nutritional resources and biomaterials[28].

11.4.5 Distinct Carbohydrates Can Have Identical Signals

Caution should be taken when utilizing NMR maps to investigate novel species with unknown composition. In certain instances, two distinct carbohydrates may demonstrate strictly overlapping signals, for example, between arabinofuranose (Ara*f*) and Gal*f*[25]. While Ara*f* (typically called Ara in our studies) is a common sugar forming the side chains of pectin and hemicellulose in plants, Gal*f* is a rare form of galactose and is sporadically identified in some fungal species, for example, as a constituent of the GM present in the cell wall[79-81]. Although Gal*f* is a hexose and Ara is a pentose, both sugar units possess five-membered rings, resulting in identical ^{13}C chemical shifts in their first four carbons, with representative chemical shifts of approximately 108 ppm, 82 ppm, 77 ppm, and 84 ppm, respectively. The only unique signals of Gal*f* in a 2D INADEQUATE spectrum come from its C5-C6 carbon pair, however, it showed up in a heavily crowded spectral region and cannot be utilized as a reliable tracker of molecules (**Figure 11.7a**). In a standard 1D spectrum, the sole distinctive signal for Gal*f*, when compared to Araf, is the C5 signal at 69 ppm (**Figure 11.7b**). However, this peak falls within a densely populated region, making it impractical for distinguishing signals from other carbohydrates. The occurrence of such indistinguishable spectral patterns between two chemically and structurally distinct monosaccharide units was unanticipated as we extend our reach of carbohydrate ssNMR and expand our knowledge and understanding of these complex biopolymers. Hence, it is essential to carefully consider the source of the samples being analyzed, be it from plants, fungi, or other

organisms. In the case of a novel species, it becomes imperative to complement NMR analysis with the chemical assays that provide information on the linkage patterns and chemical composition of carbohydrate constituents. Moreover, the reliance on 1D spectra for analyzing previously uncharted samples should be minimized.

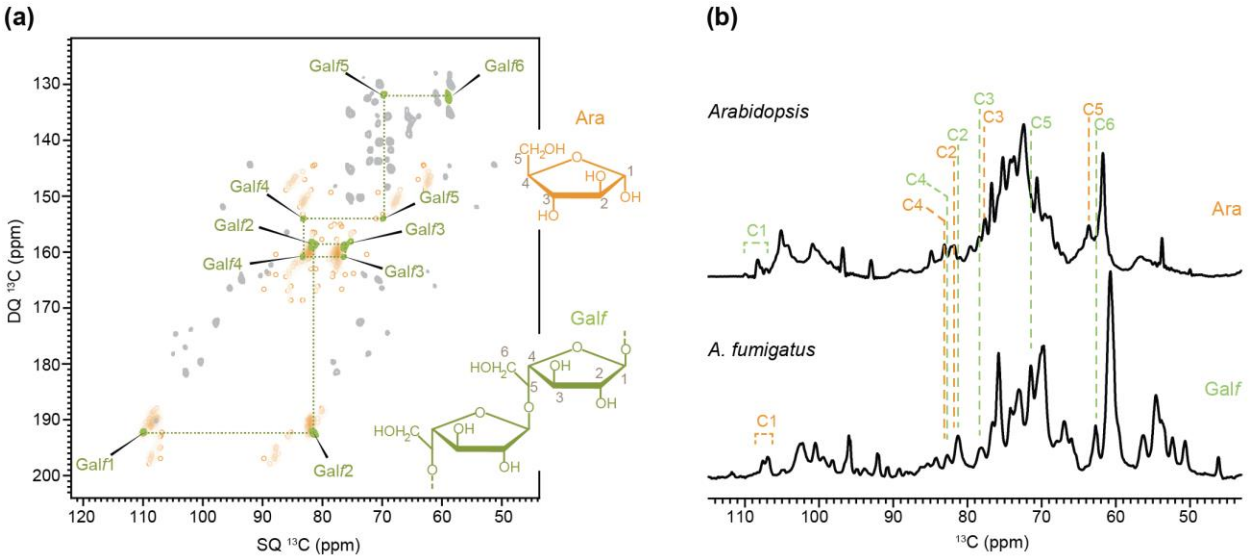

Figure 11.7 Comparison of the NMR signals of Gal*f* and Ara. **a,** The C1-C2-C3-C4-C5 signals of Gal*f* in the 2D DP refocused J-INADEQUATE spectrum of *A. fumigatus* falls within the expected signals for arabinose residues from plant cell walls. Only the C5-C6 carbon pair of Gal*f* is unique. **b,** Comparison of 1D ^{13}C quantitative DP spectrum measured on *Arabidopsis* (top) and *A. fumigatus* (bottom). Orange and green lines indicate the signals from Ara (only in *Arabidopsis*) and Gal*f* residues (only in *A. fumigatus*), respectively.

11.5 Conclusions and Perspectives

Significant advances have been made in solid-state NMR analysis of carbohydrates over the past decade, allowing us to identify key spectroscopic features of major plant and fungal glycans. It is critical to ensure that there is sufficient spectral resolution to track the minor

differences in molecules, carefully track the full carbon connectivity in 2D/3D correlation spectra without missing any carbon in the linkage pattern, seek additional validation from other techniques, such as mass spectrometry, and cross-compare with mutants or other strains lacking the carbohydrate of interest. Further datasets are required before statistical analysis and automated tools can be developed to facilitate the analysis of these complex carbohydrate polymers. It is anticipated that tools similar to those already available for other molecules, such as the DEEP picker for protein 2D spectral deconvolution and peak picking, and the probabilistic assignment tool for organic crystals[59, 82], will lead to significant progress in solid-state glycoNMR.

11.6 Acknowledgments

This work was supported by the National Science Foundation (NSF) MCB-2308660. Part of the NMR data was collected on a 700 MHz NMR spectrometer funded by NSF Major Research Instrumentation Program under the award number DBI-2019046. A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by National Science Foundation Cooperative Agreement No. DMR-2128556 and the State of Florida.

REFERENCES

1. Fukamizo, T.; Kramer, K. J.; Mueller, D. D.; Schaefer, J.; Carbow, J.; Jacob, G. S., Analysis of Chitin Structure by Nuclear Magnetic Resonance Spectroscopy and Chitinolytic Enzyme Digestion. *Arch. Biochem. Biophys.* **1986,** *249*, 15-26.

2. Schaefer, J.; Stejskal, E. O.; Beard, C. F., Carbon-13 Nuclear Magnetic Resonance Analysis of Metabolism in Soybeans Labeled by $^{13}CO_2$. *Plant Physiol.* **1975,** *55*, 1048-1053.

3. Bardet, M.; Emsley, L.; Vincendon, M., Two-dimensional spin-exchange solid-state NMR studies of 13 C-enriched wood *Solid State Nucl. Magn. Reson.* **1997,** *8*, 25-32.

4. Lesage, A.; Bardet, M.; Emsley, L., Through-bond carbon-carbon connectivities in disordered solids by NMR. *J. Am. Chem. Soc.* **1999,** *188*, 10987-10993.

5. Tong, G.; Pan, Y.; Dong, H.; Pryor, R.; Wilson, G. E.; Schaefer, J., Structure and Dynamics of Pentaglycyl Bridges in the Cell Walls of Staphylococcus aureus by ^{13}C-^{15}N REDOR NMR. *Biochemistry* **1997,** *36*, 9859-9866.

6. Atalla, R. H.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. *Science* **1984,** *223*, 283-285.

7. Ghassemi, N.; Poulhazan, A.; Deligey, F.; Mentink-Vigier, F.; Marcotte, I.; Wang, T., Solid-State NMR Investigations of Extracellular Matrices and Cell Walls of Algae, Bacteria, Fungi, and Plants. *Chem. Rev.* **2022,** *122*, 10036–10086.

8. Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M., Solid-state NMR spectroscopy. *Nat. Rev. Methods Primers* **2021,** *1* , 2.

9. Zhao, W.; Fernando, L. D.; Kirui, A.; Deligey, F.; Wang, T., Solid-state NMR of plant and fungal cell walls: a critical review. *Solid State Nucl. Magn. Reson.* **2020,** *107*, 101660.

10. Phyo, P.; Hong, M., Fast MAS ^{1}H-^{13}C correlation NMR for structural investigations of plant cell walls. *J. Biomol. NMR* **2019,** *73*, 661-674.

11. Safeer, A.; Kleiburg, F.; Bahri, S.; Beriashvili, D.; Veldhuizen, E. J. A.; van Neer, J.; Tegelaar, M.; de Cock, H.; Wosten, H. A. B.; Baldus, M., Probing Cell-Surface Interactions in Fungal Cell Walls by High-Resolution 1H-Detected Solid-State NMR Spectroscopy. *Chem. Euro. J.* **2022,** *29*, e202202616.

12. Bougault, C.; Ayala, I.; Vollmer, W.; Simorre, J. P.; Schanda, P., Studying intact bacterial peptidoglycan by proton-detected NMR spectroscopy at 100 kHz MAS frequency. *J. Struct. Biol.* **2019,** *206*, 66-72.

13. Dick-Perez, M.; Zhang, Y.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and Interactions of Plant Cell-Wall Polysaccharides by Two- and Three-Dimensional Magic-Angle-Spinning Solid-State NMR. *Biochemistry* **2011,** *50*, 989-1000.

14. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water–Polysaccharide Interactions in the Primary Cell Wall of Arabidopsis thaliana from Polarization Transfer Solid-State NMR. *J. Am. Chem. Soc.* **2014,** *136*, 10399-10409.

15. Temple, H.; Phyo, P.; Yang, W.; Lyczakowski, J. J.; Echevarria-Poza, A.; Yakunin, I.; Parra-Rojas, J. P.; Terrett, O. M.; Saez-Aguayo, S.; Dupree, R.; Orellana, A.; Hong, M.; Dupree, P., Golgi-localized putative S-adenosyl methionine transporters required for plant cell wall polysaccharide methylation *Nat. Plants* **2022,** *8*, 656-669.

16. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis* primary cell walls: evidence from solid-state nuclear magnetic resonance. *Plant Physiol.* **2015,** *168*, 871-884.

17. Terrett, O. M.; Lyczakowski, J. J.; Yu, L.; Iuga, D.; Franks, W. T.; Brown, S. P.; Dupree, R.; Dupree, P., Molecular architecture of softwood revealed by solid-state NMR. *Nat. Commun.* **2019,** *10*, 4978.

18. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Pöppler, A.-C.; Brown, S. P.; Deazevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** *7*, 13902.

19. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015,** *54*, 2335-2345.

20. Kang, X.; Kirui, A.; Widanage, M. C. D.; Mentink-Vigier, F.; Cosgrove, D. J.; Wang, T., Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nat. Commun.* **2019,** *10*, 347.

21. Kirui, A.; Zhao, W.; Deligey, F.; Yang, H.; Kang, X.; Mentink-Vigier, F.; Wang, T., Carbohydrate-aromatic interface and molecular architecture of lignocellulose. *Nat. Commun.* **2022,** *13*, 538.

22. Gao, Y.; Lipton, A. S.; Wittmer, Y.; Murray, D. T.; Mortimer, J. C., A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nat. Commun.* **2020,** *11*, 6081.

23. Kelly, J. E.; Chrissian, C.; Stark, R. E., Tailoring NMR experiments for structural characterization of amorphous biological solids: A practical guide. *Solid State Nucl. Magn. Reson.* **2020,** *109*, 101686.

24. Ehren, H. L.; Appels, F. V. W.; Houben, K.; Renault, M. A. M.; Wösten, H. A. B.; Baldus, M., Characterization of the cell wall of a mushroom forming fungus at atomic resolution using solid-state NMR spectroscopy. *Cell Surf.* **2020,** *6*, 100046.

25. Chakraborty, A.; Fernando, L. D.; Fang, W.; Dickwella Widanage, M. C.; Wei, P.; Jin, C.; Fontaine, T.; Latgé, J. P.; Wang, T., A molecular vision of fungal cell wall organization by

functional genomics and solid-state NMR. *Nat. Commun.* **2021,** *12*, 6346.

26. Kang, X.; Kirui, A.; Muszynski, A.; Widanage, M. C. D.; Chen, A.; Azadi, P.; Wang, P.; Mentink-Vigier, F.; Wang, T., Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat. Commun.* **2018,** *9*, 2747.

27. Chrissian, C.; Lin, C. P.-C.; Camacho, E.; Casadevall, A.; Neiman, A. M.; Stark, R. E., Unconventional Constituents and Shared Molecular Architecture of the Melanized Cell Wall of *C. neoformans* and Spore Wall of *S. cerevisiae*. *J. Fungi* **2020,** *6*, 329.

28. Pylkkanen, R.; Werner, D.; Bishoyi, A.; Weil, D.; Scoppola, E.; Wagermaier, W.; Safeer, A.; Bahri, S.; Baldus, M.; Paananen, A.; Penttila, M.; Szilvay, G.; Mohammadi, P., The complex structure of Fomes fomentarius represents an architectural design for high-performance ultralightweight materials. *Sci. Adv.* **2023,** *8*, ade5417.

29. Nygaard, R.; Romaniuk, J. A. H.; Rice, D. M.; Cegelski, L., Spectral snapshots of bacterial cell-wall composition and the influence of antibiotics by whole-cell NMR. *Biophys. J.* **2015,** *108*, 1380-1389.

30. Arnold, A. A.; Bourgouin, J. P.; Genard, B.; Warschawski, D. E.; Tremblay, R.; Marcotte, I., Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. *J. Biomol. NMR* **2018,** *70*, 123-131.

31. Romaniuk, J. A. H.; Cegelski, L., Bacterial cell wall composition and the influence of antibiotics by cell-wall and whole-cell NMR. *Philos. Trans. R. Soc. B* **2015,** *370*, 20150024.

32. Poulhazan, A.; Dickwella Widanage, M. C.; Muszynski, A.; Arnold, A. A.; Warschawski, D. E.; Azadi, P.; Marcotte, I.; Wang, T., Identification and Quantification of Glycans in Whole Cells: Architecture of Microalgal Polysaccharides Described by Solid-State Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* **2021,** *143*, 19374-19388.

33. Reichhardt, C.; Cegelski, L., Solid-state NMR for bacterial biofilms. *Mol. Phys.* **2013,** *7*, 887-894.

34. El Hariri El Nokab, M.; Lasorsa, A.; Sebakhy, K. O.; Pincchioni, F.; van der Wel, P. C. A., Solid-state NMR spectroscopy insights for resolving different water pools in alginate hydrogels. *Food Hydrocol.* **2022,** *127*, 107500.

35. El Hariri El Nokab, M.; van der Wel, P. C. A., Use of solid-state NMR spectroscopy for investigating polysaccharide-based hydrogels: A review. *Carbohydr. Polym.* **2020,** *240*, 116276.

36. Zhao, W.; Deligey, F.; Shekar, S. C.; Mentink-Vigier, F.; Wang, T., Current limitations of solid-state NMR in carbohydrate and cell wall research. *J. Magn. Reson.* **2022,** *341*, 107263.

37. Kang, X.; Zhao, W.; Widanage, M. C. D.; Kirui, A.; Ozdenvar, U.; Wang, T., CCMRD: a solid-state NMR database for complex carbohydrates. *J. Biomol. NMR* **2020,** *74*, 239-245.

38. Fritzsching, K. J.; Hong, M.; Schmidt-Rohr, K., Conformationally selective multidimensional chemical shift ranges in proteins from a PACSY database purged using intrinsic quality criteria. *J. Biomol. NMR* **2016,** *64*, 115-130.

39. Hou, G.; Yan, S.; Trebosc, J.; Amoureux, J. P.; Polenova, T., Broadband Homonuclear Correlation Spectroscopy Driven by Combined R2nv Sequences under Fast Magic Angle Spinning for NMR Structural Analysis of Organic and Biological Solids. *J. Magn. Reson.* **2013,** *232*, 18-30.

40. De Paepe, G.; Lewandowski, J. R.; Loquet, A.; Bockmann, A.; Griffin, R. G., Proton assisted recoupling and protein structure determination. *J. Chem. Phys.* **2008,** *129*, 245101.

41. Zhao, W.; Kirui, A.; Deligey, F.; Mentink-Vigier, F.; Zhou, Y.; Zhang, B.; Wang, T., Solid-state NMR of unlabeled plant cell walls: high-resolution structural analysis without isotopic enrichment. *Biotechnol. Biofuels* **2021,** *14*, 14.

42. Viger-Gravel, J.; Lan, W.; Pinon, A. C.; Berruyer, P.; Emsley, L.; Bardet, M.; Luterbacher, J., Topology of Pretreated Wood Fibers Using Dynamic Nuclear Polarization. *J. Phys. Chem. C* **2019,** *123*, 30407-30415.

43. Perras, F. A.; Luo, H.; Zhang, X.; Mosier, N. S.; Pruski, M.; Abu-Omar, M. M., Atomic-Level Structure Characterization of Biomass Pre- and Post-Lignin Treatment by Dynamic Nuclear Polarization-Enhanced Solid-State NMR. *J. Phys. Chem. A* **2017,** *121*, 623-630.

44. Fernando, L. D.; Dickwella Widanage, M. C.; Shekar, S. C.; Mentink-Vigier, F.; Wang, P.; Wi, S.; Wang, T., Solid-state NMR analysis of unlabeled fungal cell walls from Aspergillus and Candida species. *J. Struct. Biol. X* **2022,** *6*, 100070.

45. Deligey, F.; Frank, M. A.; Cho, S. H.; Kirui, A.; Mentink-Vigier, F.; Swulius, M. T.; Nixon, B. T.; Wang, T., Structure of In Vitro-Synthesized Cellulose Fibrils Viewed by Cryo-Electron Tomography and ^{13}C Natural-Abundance Dynamic Nuclear Polarization Solid-State NMR. *Biomacromolecules* **2022,** *23*, 2290-2301.

46. Takahashi, H.; Lee, D.; Dubois, L.; Bardet, M.; Hediger, S.; De Paepe, G., Rapid Natural-Abundance 2D ^{13}C-^{13}C Correlation Spectroscopy Using Dynamic Nuclear Polarization Enhanced Solid-State NMR and Matrix-Free Sample Preparation. *Angew. Chem. Int. Ed.* **2012,** *51*, 11766-11769.

47. Kumar, A.; Watbled, B.; Baussanne, I.; Hediger, S.; Demeunynck, M.; De Paepe, G., Optimizing chemistry at the surface of prodrug-loaded cellulose nanofibrils with MAS-DNP. *Commun. Chem.* **2023,** *6*, 58.

48. Berruyer, P.; Gericke, M.; Moutzouri, P.; Jakobi, D.; Bardet, M.; Karlson, L.; Schantz, S.; heinze, T.; Emsley, L., Advanced characterization of regioselectively substituted methylcellulose model compounds by DNP enhanced solid-state NMR spectroscopy. *Carbohydr. Polym.* **2021,** *262*, 117944.

49. Brouwer, D. H.; Mikolajewski, J. G., Resolving the discrepancies in reported ^{13}C solid state

NMR chemical shifts for native celluloses. *Cellulose* **2023,** *30*, 4827-4839.

50. Hoffman, R., Solid-state chemical-shift referencing with adamantane. *J. Magn. Reson.* **2022,** *340*, 107231.

51. Elkins, M. R.; Wang, T.; Nick, M.; Jo, H.; Lemmin, T.; Prusiner, S. B.; DeGrado, W. F.; Stohr, J.; Hong, M., Structural polymorphism of Alzheimer's β-amyloid fibrils as controlled by an E22 switch: A solid-state NMR study. *J. Am. Chem. Soc.* **2016,** *138*, 9840-9852.

52. Kirui, A.; Du, J.; Zhao, W.; Barnes, W.; Kang, X.; Anderson, C. T.; Xiao, C.; Wang, T., A pectin methyltransferase modulates polysaccharide dynamics and interactions in Arabidopsis primary cell walls: Evidence from solid-state NMR. *Carbohydr. Polym.* **2021,** *270*, 118370.

53. Wang, T.; Chen, Y.; Tabuchi, A.; Cosgrove, D. J.; Hong, M., The target of β-expansin EXPB1 in maize cell walls from binding and solid-state NMR studies. *Plant Physiol.* **2016,** *172*, 2107-2119.

54. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and Dynamics of Brachypodium Primary Cell Wall Polysaccharides from Two-Dimensional ^{13}C Solid-State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2014,** *53*, 2840-2854.

55. Phyo, P.; Wang, T.; Kiemle, S. N.; O'Neill, H.; Pingali, S. V.; Hong, M.; Cosgrove, D. J., Gradients in wall mechanics and polysaccharides along growing inflorescence stems. *Plant Physiol.* **2017,** *175*, 1593-1607.

56. Kohn, R.; Kovac, P., Dissociation-constants of D-galacturonic and D-glucuronic acid and their O-methyl derivatives. *Chem. Pap.* **1978,** *32*, 478-485.

57. Phyo, P.; Gu, Y.; Hong, M., Impact of acidic pH on plant cell wall polysaccharide structure and dynamics: insights into the mechanism of acid growth in plants from solid-state NMR. *Cellulose* **2018,** *26*, 291-304.

58. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations. *Biomacromolecules* **2016,** *17*, 2210-2222.

59. Cordova, M.; Balodis, M.; de Almeida, B. S.; Ceriotti, M.; Emsley, L., Bayesian probabilistic assignment of chemical shifts in organic solids. *Sci. Adv.* **2021,** *7*, abk2341.

60. Fritzsching, K. J.; Yang, Y.; Schmidt-Rohr, K.; Hong, M., Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information. *J. Biomol. NMR* **2013,** *56*, 155-167.

61. Yang, Y.; Fritzsching, K. J.; Hong, M., Resonance assignment of the NMR spectra of disordered proteins using a multi-objective non-dominated sorting genetic algorithm. *J. Biomol. NMR* **2013,** *57*, 281-296.

62. Hu, K. N.; Qiang, W.; Tycko, R., A general Monte Carlo/simulated annealing algorithm for

resonance assignment in NMR of uniformly labeled biopolymers. *J. Biomol. NMR* **2011,** *50*, 267-276.

63. Dick-Perez, M.; Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Multidimensional solid-state NMR studies of the structure and dynamics of pectic polysaccharides in uniformly ^{13}C-labeled Arabidopsis primary cell walls. *Magn. Reson. Chem.* **2012,** *50*, 539-550.

64. Scheller, H. V.; Ulvskov, P., Hemicelluloses. *Annu. Rev. Plant Biol.* **2010,** *61*, 263-289.

65. Ridley, B. L.; O'Neill, M. A.; Mohnen, D., Pectins: Structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **2001,** *57*, 929-967.

66. Atmodjo, M. A.; Hao, Z.; Mohnen, D., Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* **2013,** *64*, 747-779.

67. Biswal, A. K.; Atmodjo, M. A.; Li, M.; Baxter, H. L.; Geun Yoo, C.; Pu, Y.; Lee, Y. C.; Mazarei, M.; Black, I. M.; Zhang, J. Y.; Ramanna, H.; Bray, A. L.; King, Z. R.; LaFayette, P. R.; Pattathil, S.; Donohoe, B. S.; Mohanty, S. S.; Ryno, D.; Yee, K.; Thompson, O. A.; Rodriguez, M.; Dumitrache, A.; Natzke, J.; Winkeler, K.; .....; Mohnen, D., Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. *Nat. Biotechnol.* **2018,** *36*, 249-257.

68. Tan, L.; Zhang, L.; Black, I. M.; Glushka, J.; Urbanowicz, B.; Heiss, C.; Azadi, P., Most of the rhamnogalacturonan-I from cultured Arabidopsis cell walls is covalently linked to arabinogalactan-protein. *Carbohydr. Polym.* **2023,** *301*, 120340.

69. Shekar, S. C.; Zhao, W.; Fernando, L. D.; Hung, I.; Wang, T., A ^{13}C three-dimensional DQ-SQ-SQ correlation experiment for high-resolution analysis of complex carbohydrates using solid-state NMR. *J. Magn. Reson.* **2022,** *336*, 107148.

70. Li, S.; Zhang, Y.; Hong, M., 3D ^{13}C–^{13}C–^{13}C correlation NMR for de novo distance determination of solid proteins and application to a human α-defensin. *J. Magn. Reson.* **2010,** *202*, 203-210.

71. Bahri, S.; Safeer, A.; Adler, A.; Smedes, H.; Ingen, H. v.; Baldus, M., ^{1}H-detected characterization of carbon-carbon networks in highly flexible protonated biomolecules using MAS NMR. *J. Biomol. NMR* **2023,** *77*, 111-119.

72. Fernando, L. D.; Dickwella Widanage, M. C.; Penfield, J.; Lipton, A. S.; Washton, N.; Latge, J. P.; Wang, P.; Zhang, L.; Wang, T., Structural Polymorphism of Chitin and Chitosan in Fungal Cell Walls From Solid-State NMR and Principal Component Analysis. *Front. Mol. Biosci.* **2021,** *8*, 727053.

73. Pillai, C. K. S.; Paul, W.; Sharma, C. P., Chitin and Chitosan Polymers: Chemistry, Solubility and Fiber Formation. *Prog. Polym. Sci.* **2009,** *34*, 641-678.

74. Lenardon, M. D.; Whitton, R. K.; Munro, C. A.; Marshall, D.; Gow, N. A. R., Individual Chitin Synthase Enzymes Synthesize Microfibrils of Differing Structure at Specific Locations

in the Candida Albicans Cell Wall. *Mol. Microbiol.* **2007,** *66*, 1164-1173.

75. Latgé, J.-P., The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.* **2007,** *66*, 279-290.

76. Rui-Herrera, J.; Ortiz-Castellanos, L., Cell Wall Glucans of Fungi. A Review. *Cell Surf.* **2019,** *5*, 100022.

77. Grun, C. H.; Hochstenbach, F.; Humbel, B.; Verkleij, A. J.; Sietsma, J. H.; Klis, F. M.; Kamerling, J. P.; Vliegenthart, J. F. G., The structrue of cell wall a-glucan from fission yeast. *Glycobiology* **2005,** *15*, 245-257.

78. Gow, N. A. R.; Lenardon, M. D., Architecture of the dynamic fungal cell wall. *Nat. Rev. Microbiol.* **2022,** *21*, 248-259.

79. Fontaine, T.; Latgé, J. P., Galactomannan Produced by Aspergillus fumigatus: An Update on the Structure, Biosynthesis and Biological Functions of an Emblematic Fungal Biomarker. *J. Fungi* **2020,** *6*, 283.

80. Latge, J. P., Galactofuranose containing molecules in Aspergillus fumigatus. *Med. Mycol.* **2009,** *47*, S104-S109.

81. Heiss, C.; Skowyra, M. L.; Liu, H.; Klutts, J. S.; Wang, Z.; Williams, M.; Srikanta, D.; Beverley, S. M.; Azadi, P.; Doering, T. L., Unusual galactofuranose modification of a capsule polysaccharide in the pathogenic yeast Cryptococcus neoformans. *J. Biol. Chem.* **2013,** *288*, 10994-11003.

82. Li, D. W.; Hansen, A. L.; Yuan, C.; Bruschweiler-Li, L.; Bruschweiler, R., DEEP picker is a deep neural network for accurate deconvolution of complex two-dimensional NMR spectra. *Nat. Commun.* **2021,** *12*, 5229.

APPENDIX

Table S11.1 An example of simulated data for a glucose unit in cellulose. The representative dataset of actual chemical shift values of cellulose includes the six 13C chemical shifts sequentially from C1 to C6: 105.8 ppm, 71.5 ppm, 75.8 ppm, 89.1 ppm, 72.5 ppm, 64.9 ppm. The resultant INADEQUATE and SQ-SQ correlation data could be employed to construct graphical depictions portraying the projected 2D NMR spectra. The X and Y coordinates are indicative of the chemical shifts (ppm) of NMR signals/cross-peaks within a 2D spectrum. Not applicable (/).

[a] Cellulose (ppm)		DQ-SQ Correlation				SQ-SQ Correlation											
		X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
C1	**105.8**	C1 105.8	C1+C2 177.3	C2 71.5	C1+C2 177.3	C1 105.8	C1 105.8	C1 105.8	C2 71.5	C1 105.8	C3 75.8	C1 105.8	C4 89.1	C1 105.8	C5 72.5	C1 105.8	C6 64.9
C2	**71.5**	C2 71.5	C2+C3 147.3	C3 75.8	C2+C3 147.3	C2 71.5	C1 105.8	C2 71.5	C2 71.5	C2 71.5	C3 75.8	C2 71.5	C4 89.1	C2 71.5	C5 72.5	C2 71.5	C6 64.9
C3	**75.8**	C3 75.8	C3+C4 164.9	C4 89.1	C3+C4 164.9	C3 75.8	C1 105.8	C3 75.8	C2 71.5	C3 75.8	C3 75.8	C3 75.8	C4 89.1	C3 75.8	C5 72.5	C3 75.8	C6 64.9
C4	**89.1**	C4 89.1	C4+C5 161.6	C5 72.5	C4+C5 161.6	C4 89.1	C1 105.8	C4 89.1	C2 71.5	C4 89.1	C3 75.8	C4 89.1	C4 89.1	C4 89.1	C5 72.5	C4 89.1	C6 64.9
C5	**72.5**	C5 72.5	C5+C6 177.3	/		C5 72.5	C1 105.8	C5 72.5	C2 71.5	C5 72.5	C3 75.8	C5 72.5	C4 89.1	C5 72.5	C5 72.5	C5 72.5	C6 64.9
C6	**64.9**	C6 64.9	C5+C6 177.3			C6 64.9	C1 105.8	C6 64.9	C2 71.5	C6 64.9	C3 75.8	C6 64.9	C4 89.1	C6 64.9	C5 72.5	C6 64.9	C6 64.9

[a] The chemical shifts data for the six carbons of glucose in cellulose is from CCMRD.

Table S11.2 Parameters of ssNMR experiments applied for plant and fungi samples. Corresponding figure numbers are indicated. Not applicable (/).

	Figure 11.5a	**Figure 11.5a**	**Figure 11.5b**	**Figure 11.7a**	**Figure 11.7b**	**Figure 11.7b**
Sample	*Arabidopsis thaliana*	Spruce	Spruce	*Aspergillus fumigatus*	*Aspergillus fumigatus*	*Arabidopsis thaliana*
Experiment	DP J-INADEQUATE	CP J-INADEQUATE	CORD	DP J-INADEQUATE	1D 2s-DP	1D 2s-DP
Experiment Time	30 h	9 h	12 h	4 h	5 min	140 min
Number Scans	128	192	128	8	128	2048
Field (T)	16.5	9.4	9.4	18.8	18.8	16.5
Spinning Speed (kHz)	20	10	10	13.5	13.5	20
Temperature (K)	293	298	298	298	298	293
Recycle Dalay (s)	1.8	1.8	2.0	1.55	2.0	2.0
^{1}H Excitation (kHz)	/	83.3	83.3	83.3	/	/
^{13}C Excitation (kHz)	62.5	62.5	62.5	62.5	62.5	62.5
CP Contact Time (ms)	/	1.0	1.0	/	/	/
^{1}H Decoupling (kHz)	83.3	83.3	83.3	83.3	83.3	83.3
Acquisition Time (aq2; ms)	12.8	14.0	16.0	19.4	28.7	12.8
Evolution Time (aq1; ms)	5.0	5.3	4.2	10.2	/	/
Number of Points For Direct Dimensions (td2)	1600	1400	1600	2600	4096	1600
Number of Points For Indirect Dimensions (td1)	464	90	166	1024	/	/

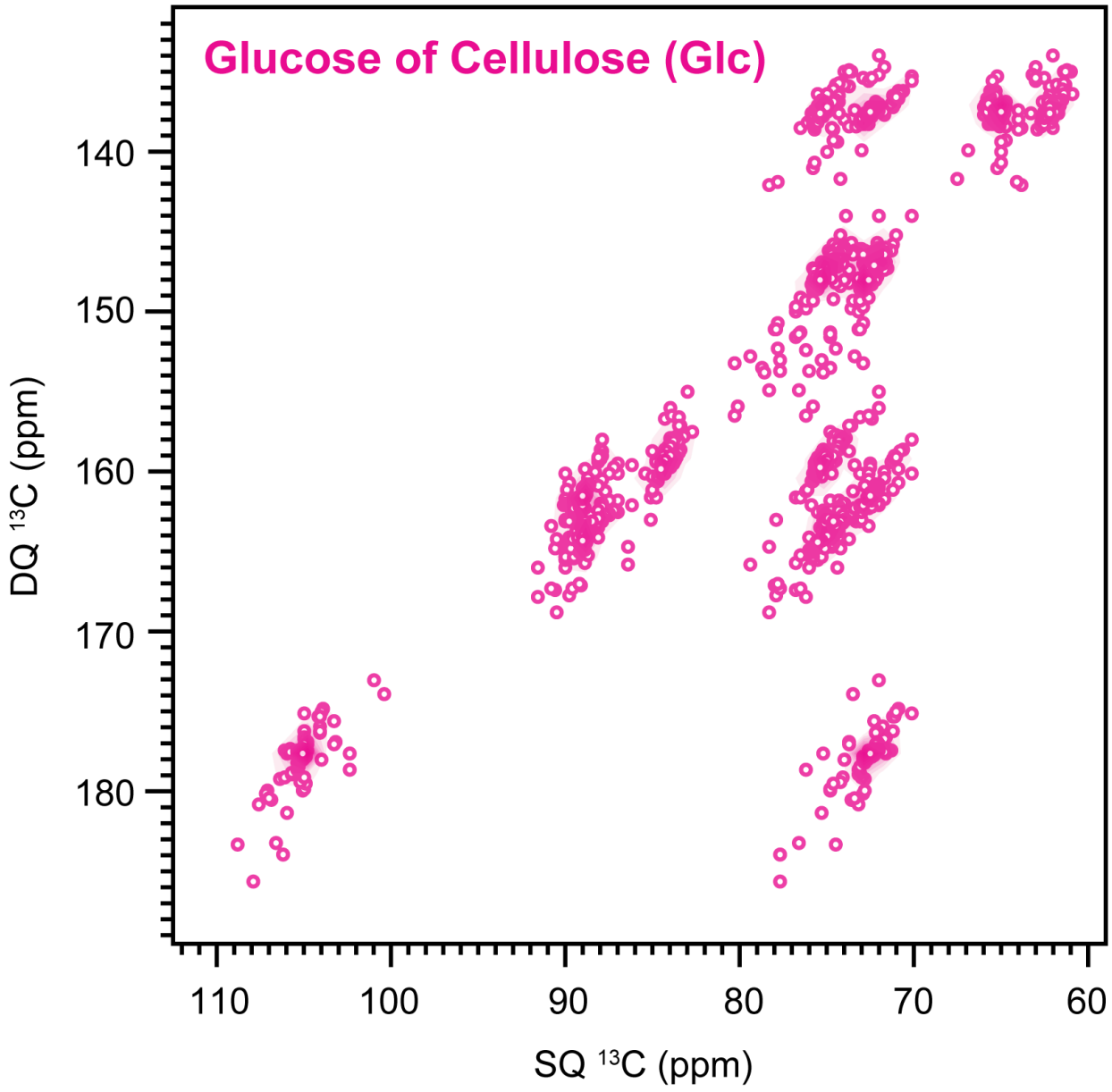

Figure S11.1 Projected 2D INADEQUATE correlation spectra of plant glucose residue in cellulose.

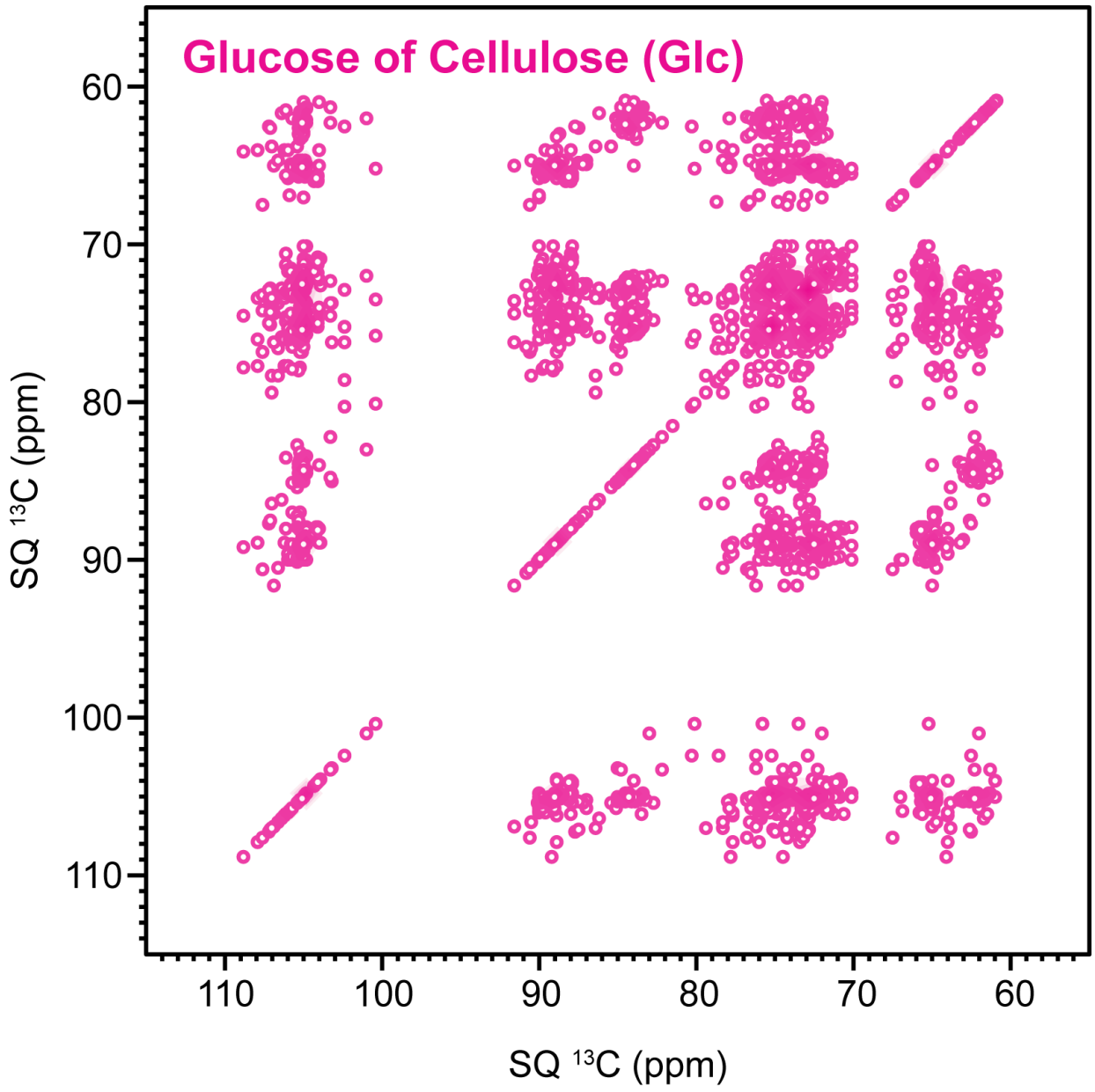

Figure S11.2 Projected 2D SQ-SQ ^{13}C correlation spectra of glucose residues in cellulose.

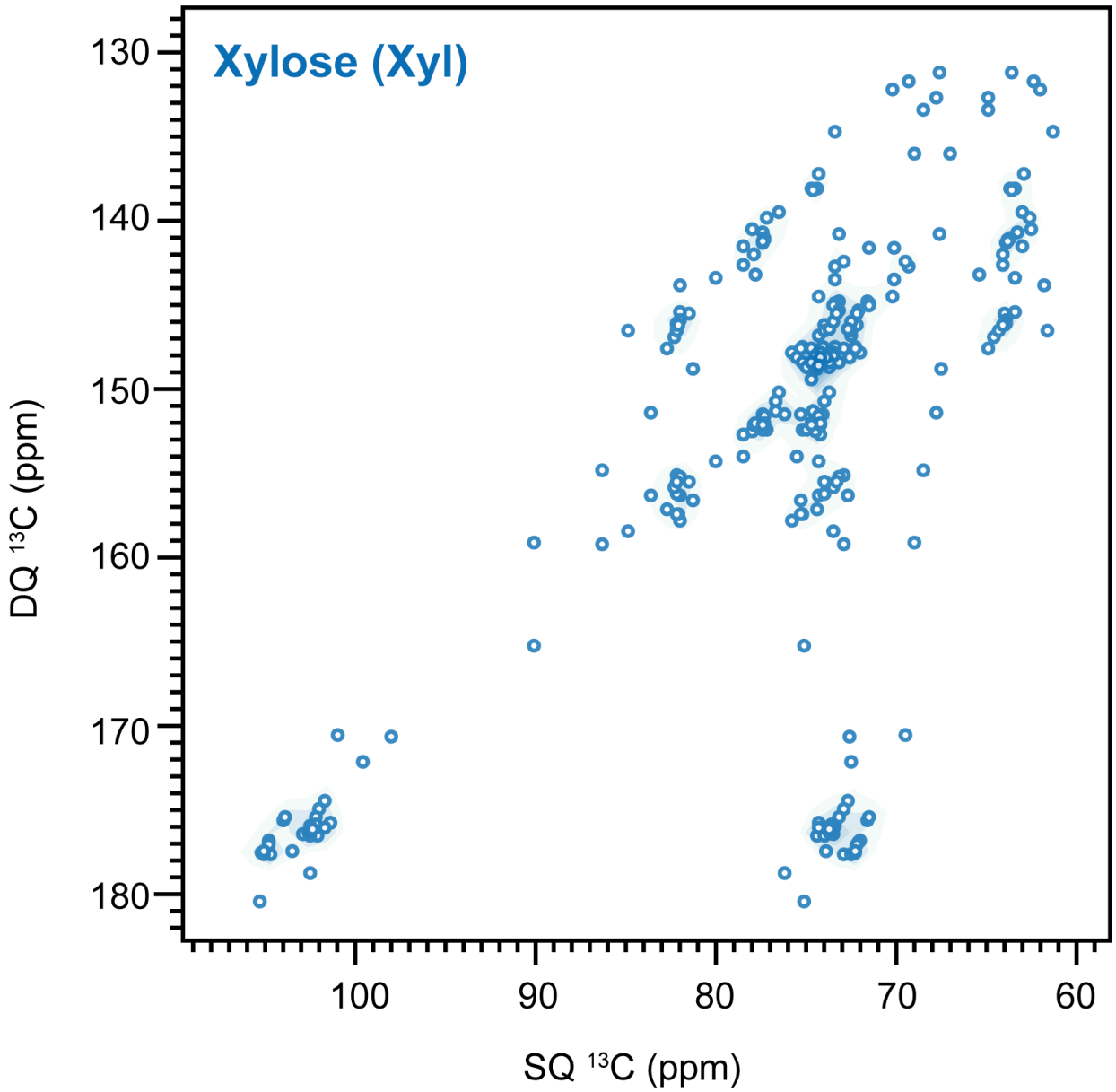

Figure S11.3 Projected 2D INADEQUATE DQ-SQ ^{13}C correlation spectra of plant xylose residue.

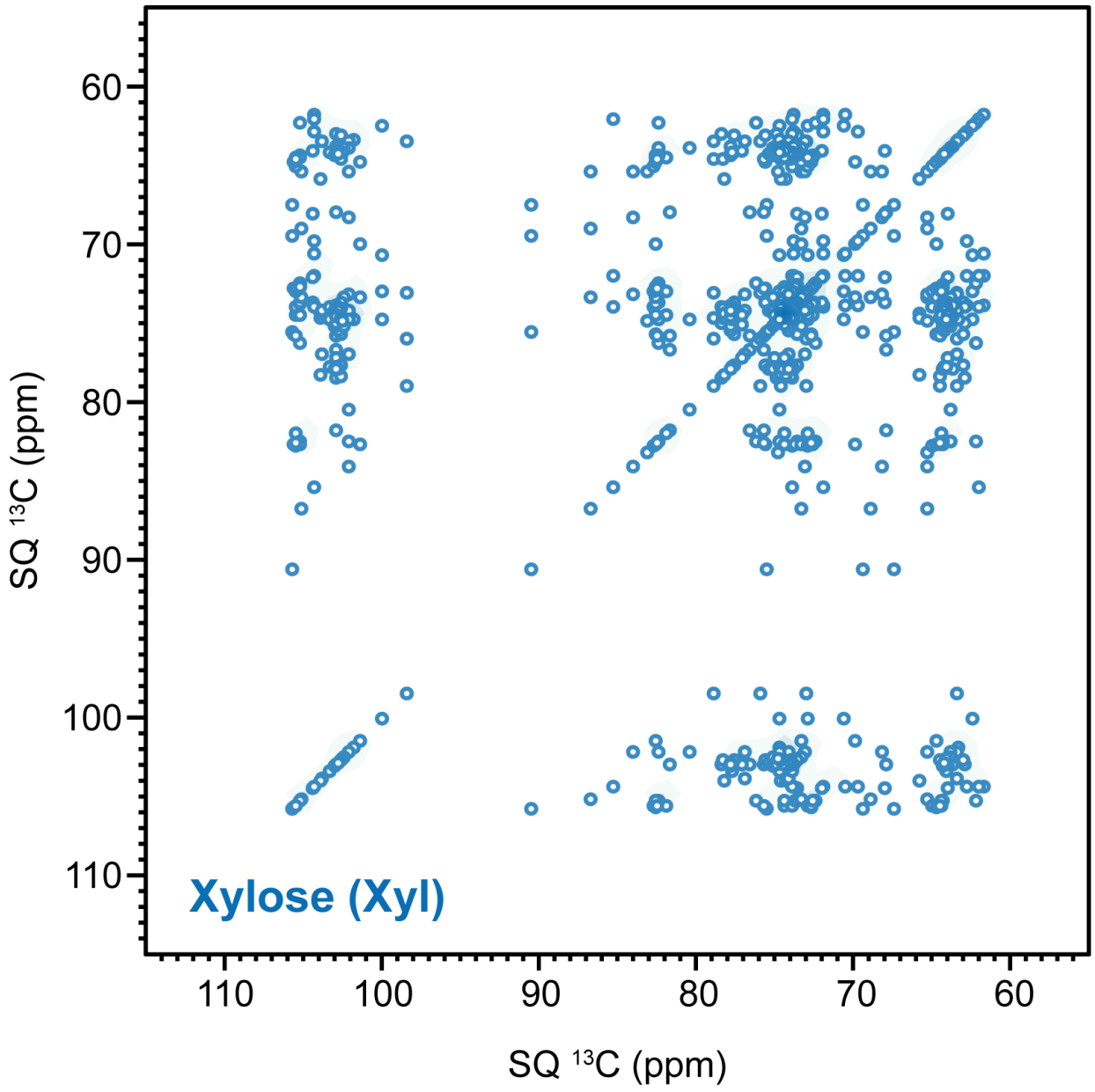

Figure S11.4 Projected 2D SQ-SQ ^{13}C correlation spectra of plant xylose residue.

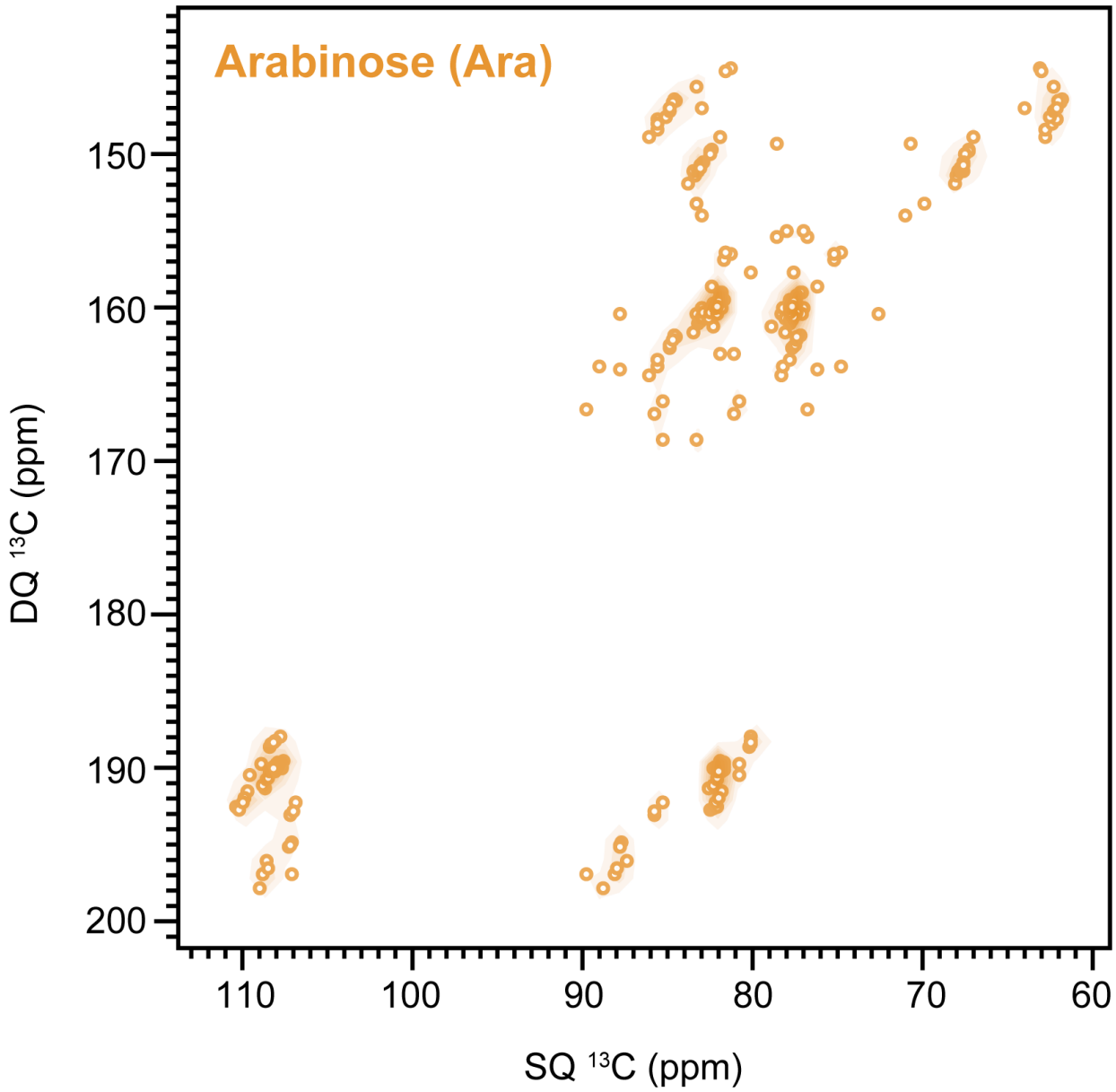

Figure S11.5 Projected 2D INADEQUATE DQ-SQ ^{13}C correlation spectra of plant arabinose residue.

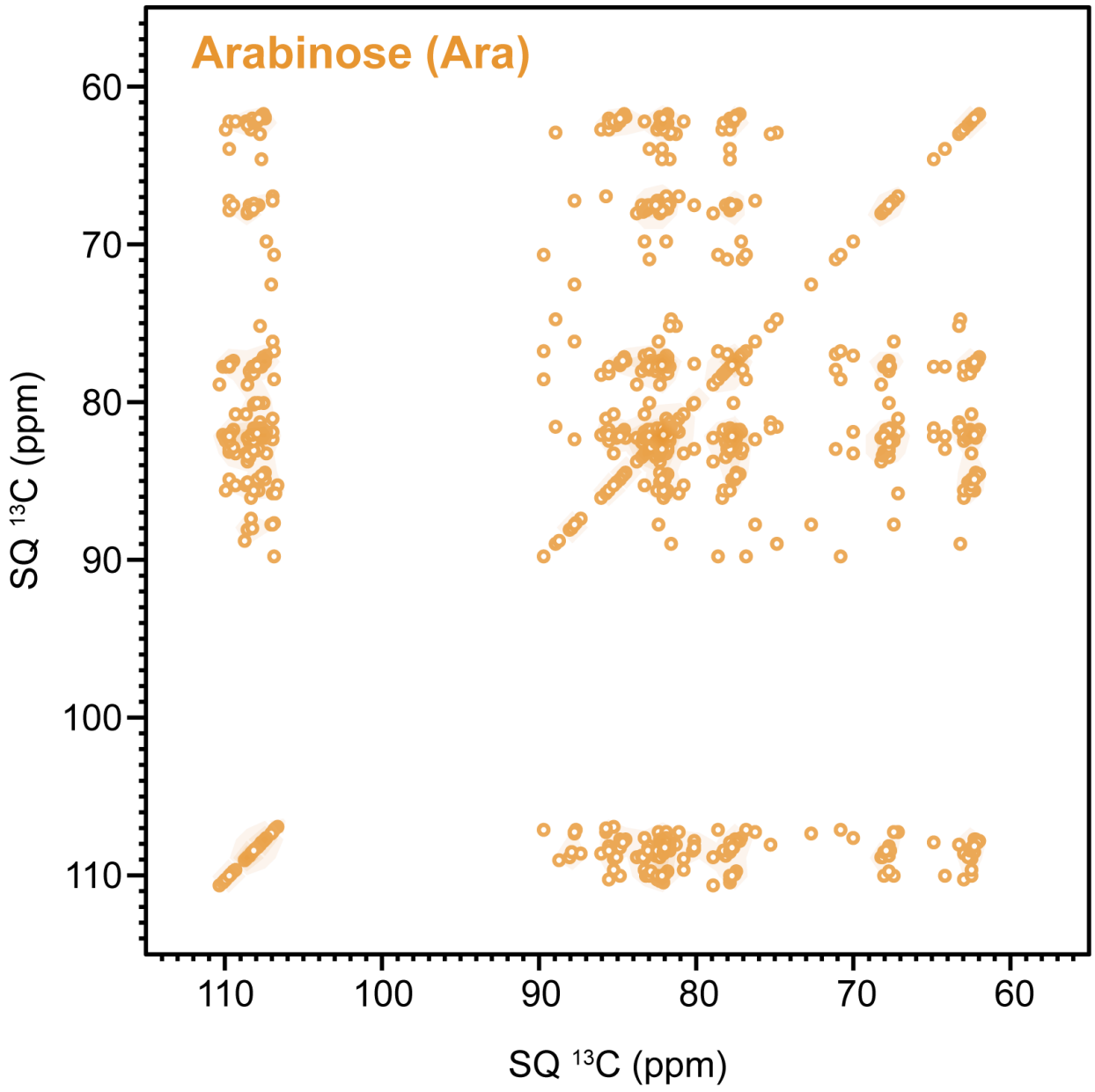

Figure S11.6 Projected 2D SQ-SQ ^{13}C correlation spectra of plant arabinose residue.

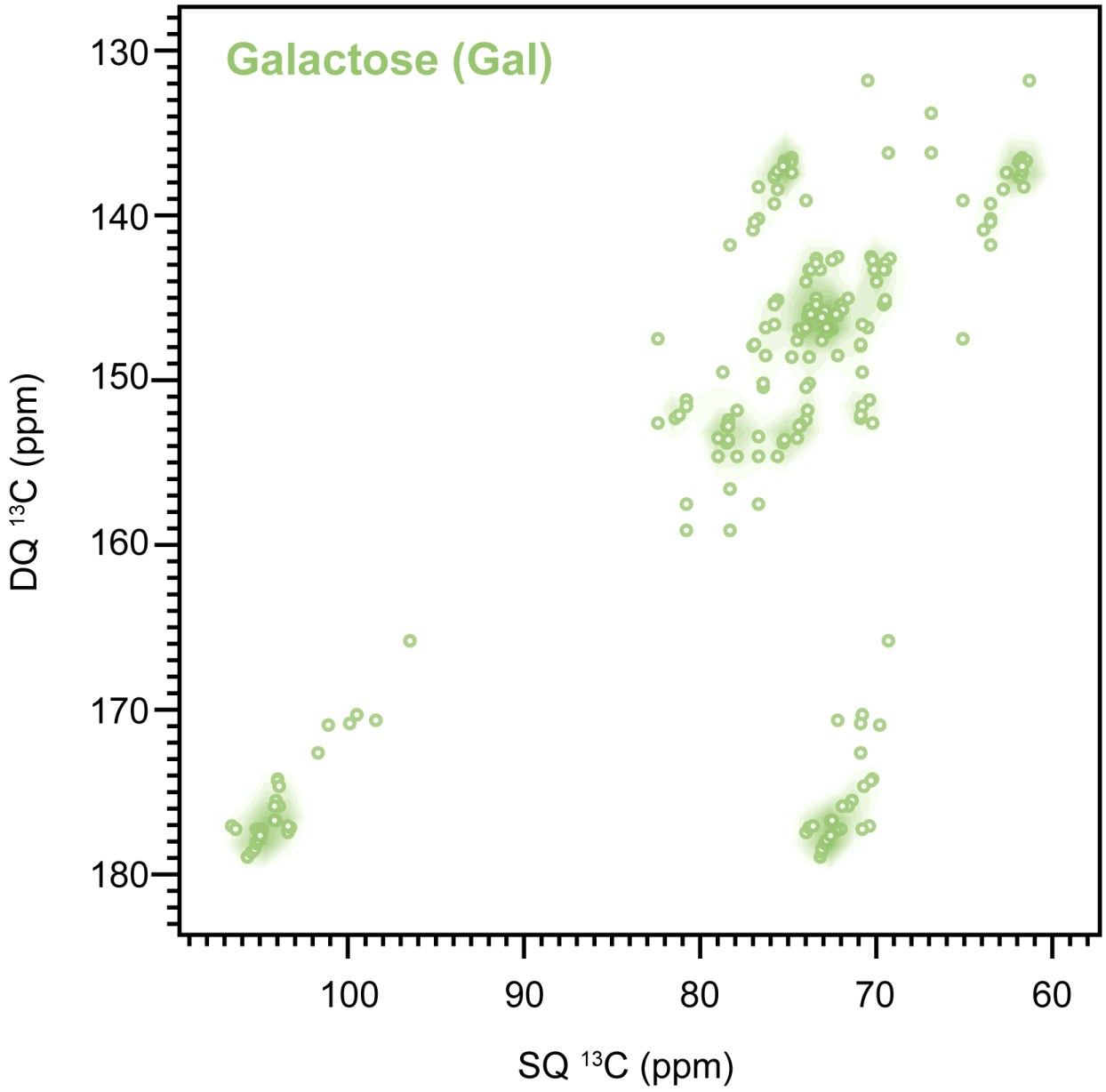

Figure S11.7 Projected 2D ^{13}C INADEQUATE correlation spectra of plant galactose residue.

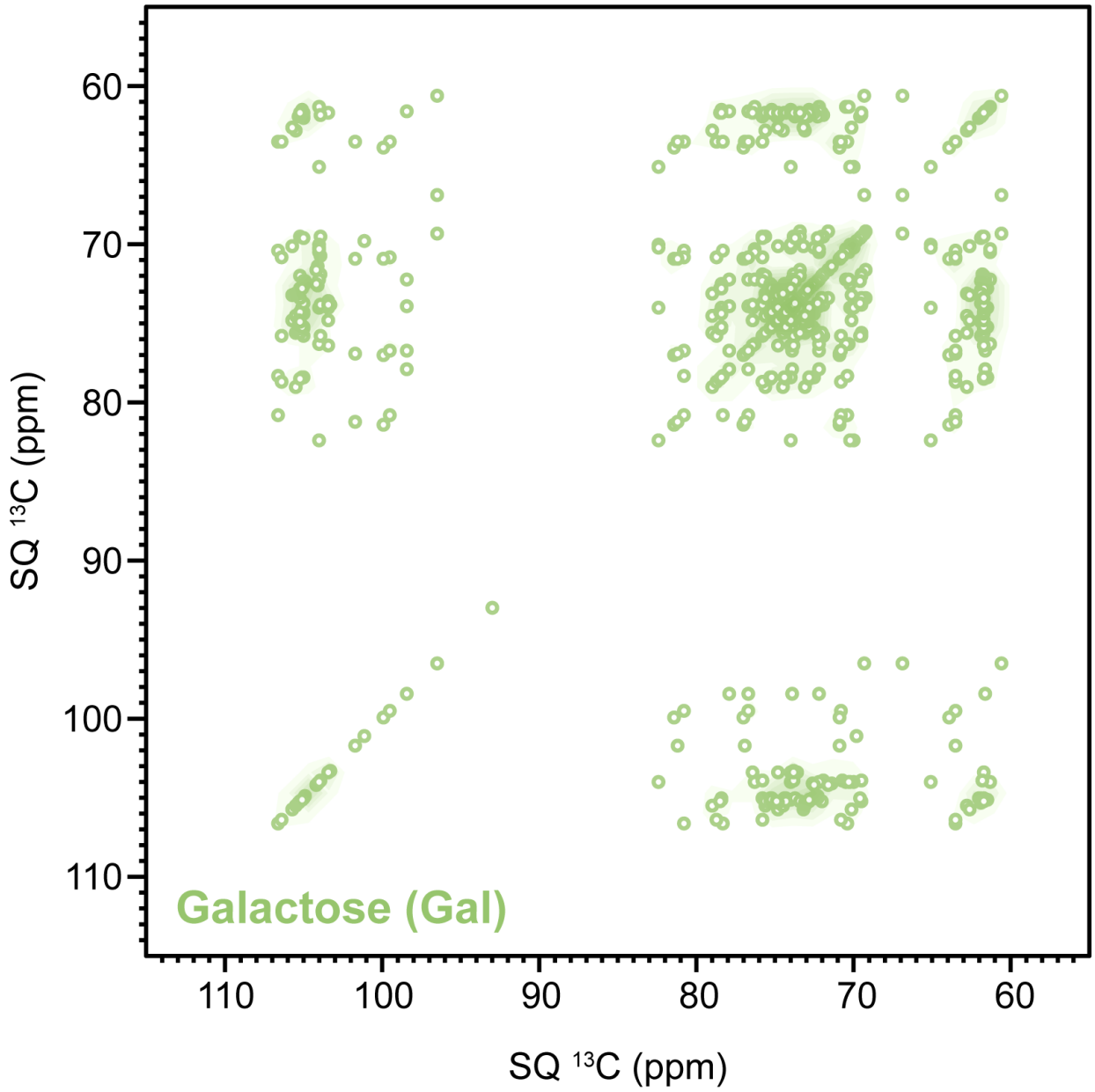

Figure S11.8 Projected 2D SQ-SQ correlation spectra of plant galactose residue.

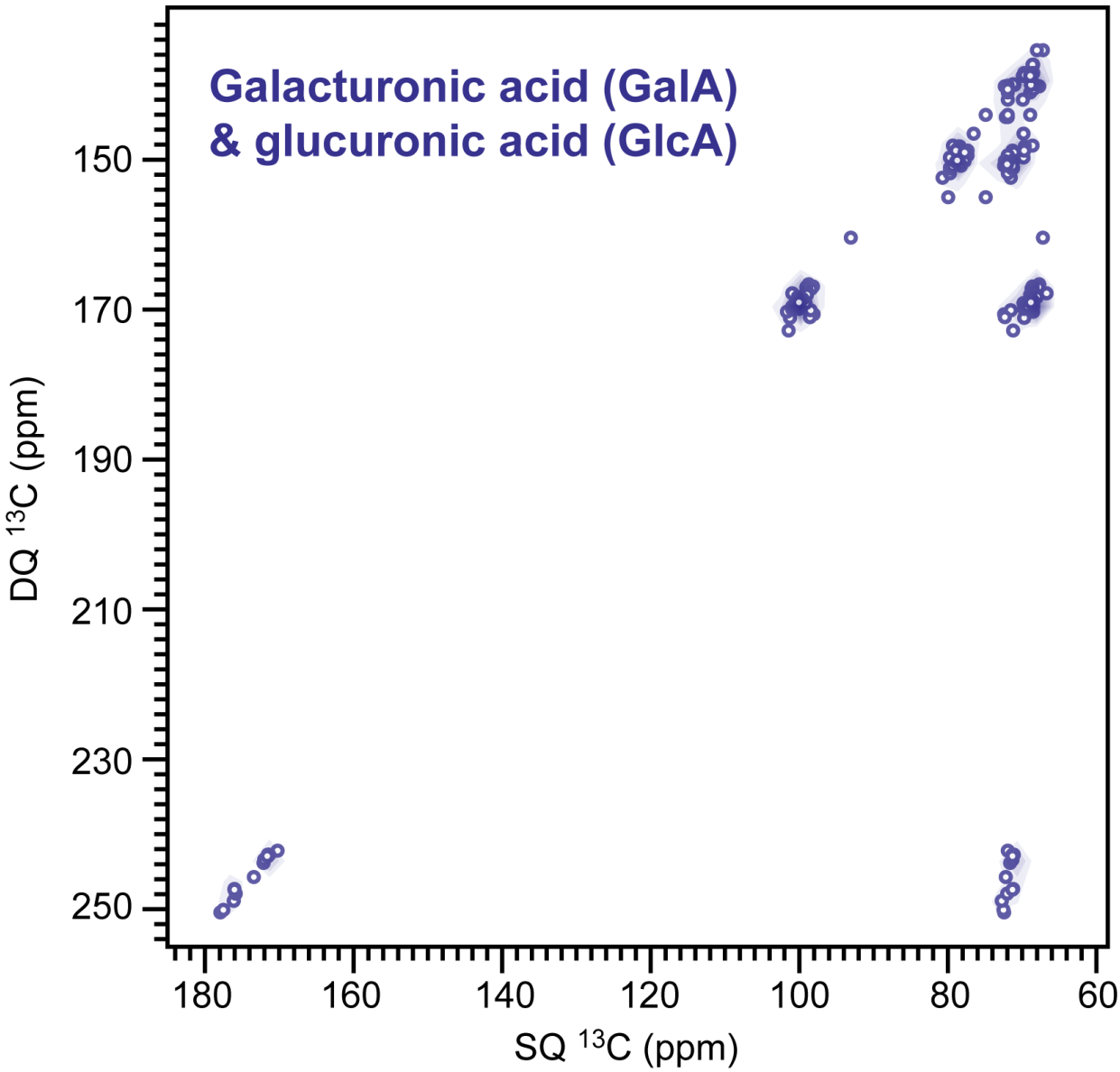

Figure S11.9 Projected 2D ^{13}C INADEQUATE correlation spectra of plant galacturonic acid and glucuronic acid residues.

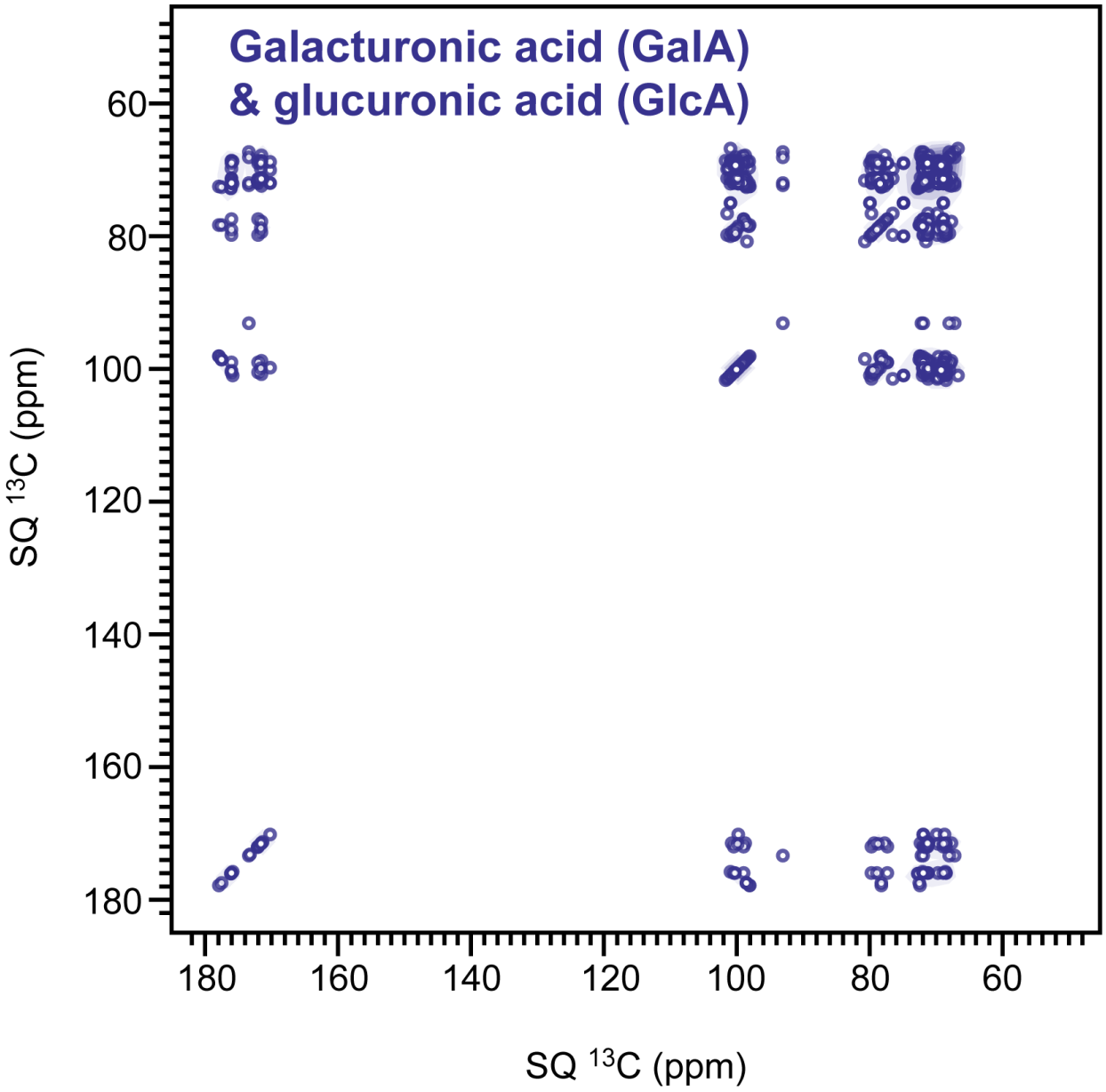

Figure S11.10 Projected 2D ^{13}C SQ-SQ correlation spectra of plant galacturonic acid and glucuronic acid residues.

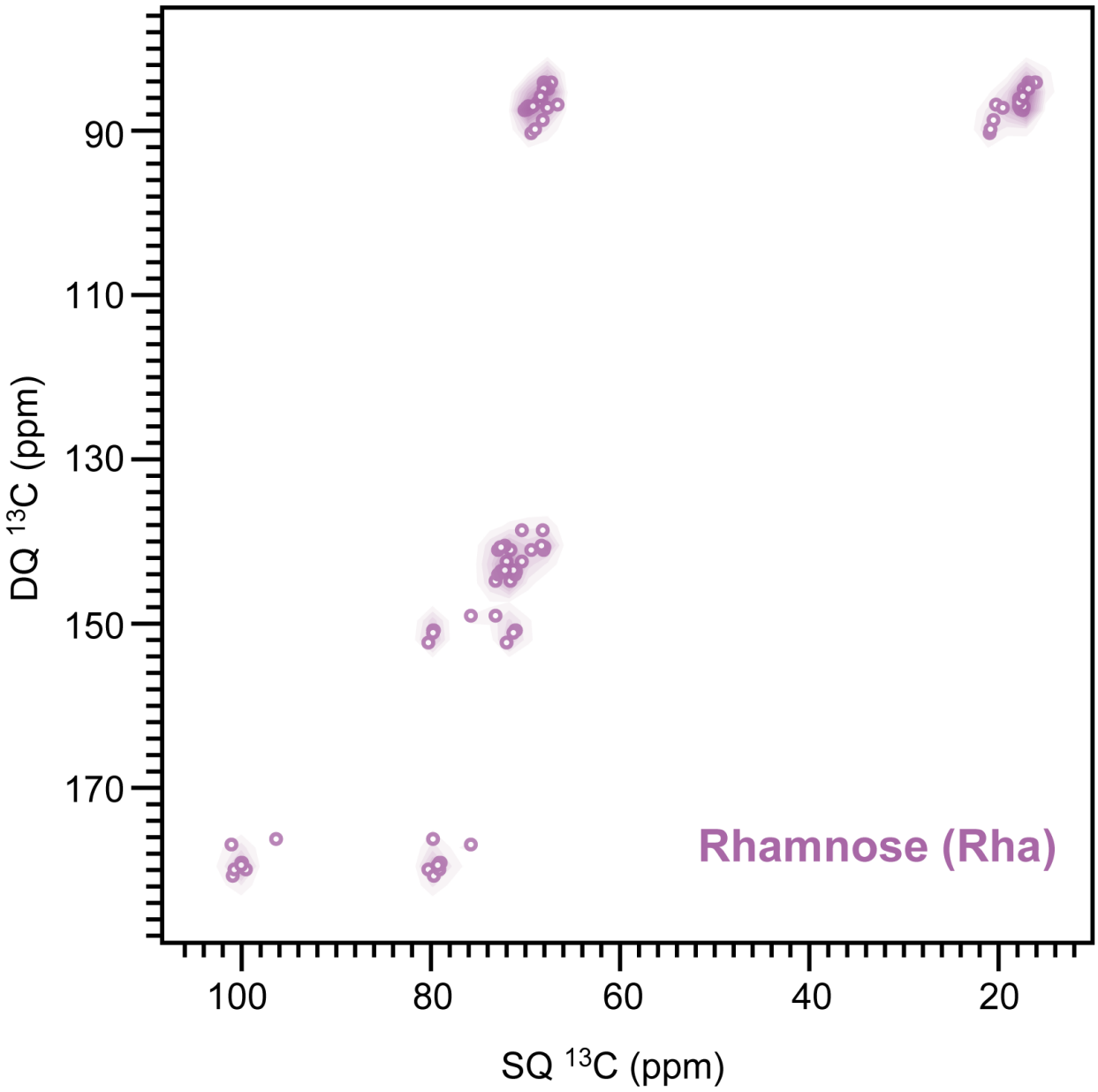

Figure S11.11 Projected 2D ^{13}C INADEQUATE correlation spectra of plant rhamnose residue.

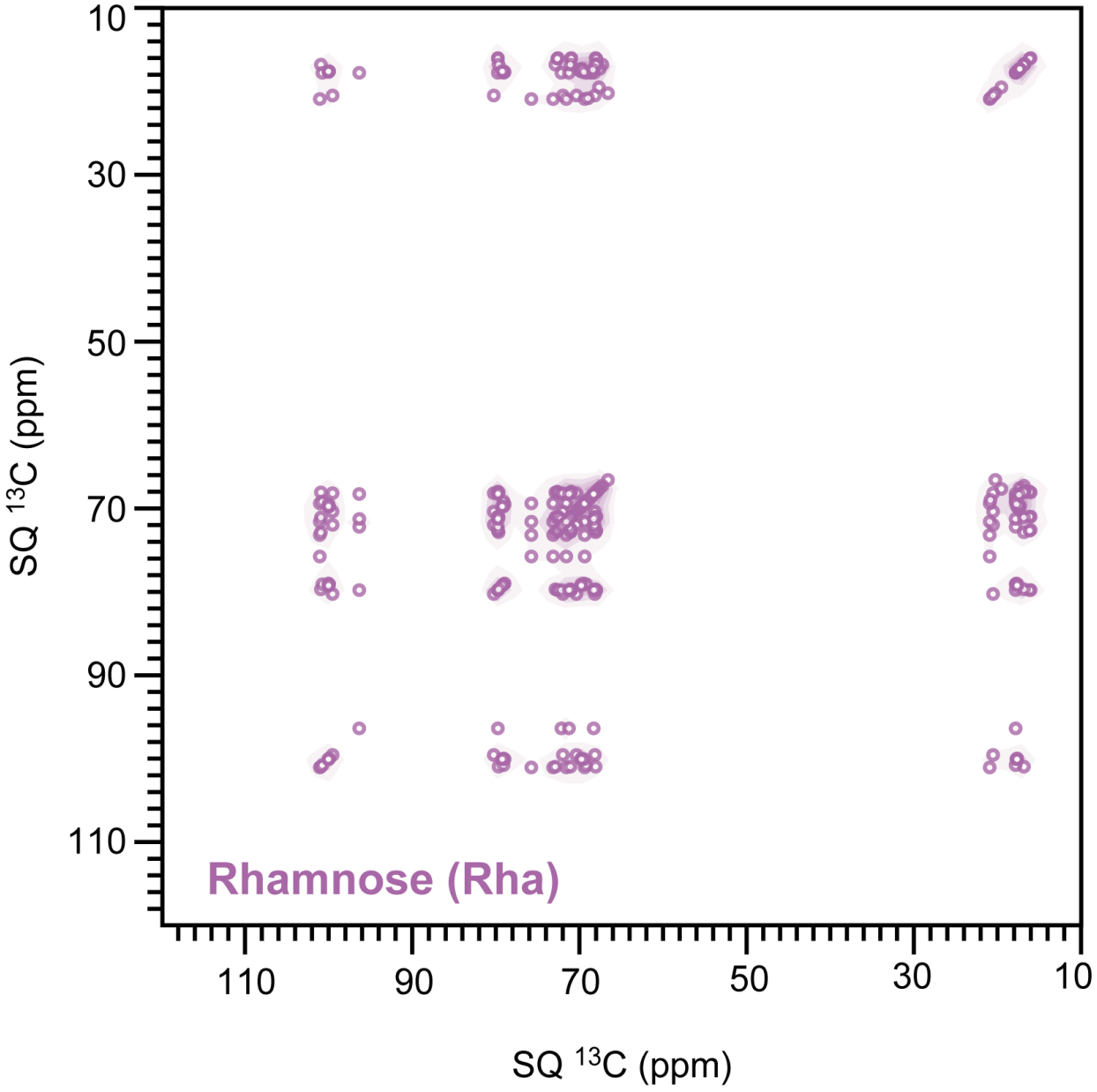

Figure S11.12 Projected 2D ^{13}C SQ-SQ correlation spectra of plant rhamnose residue.

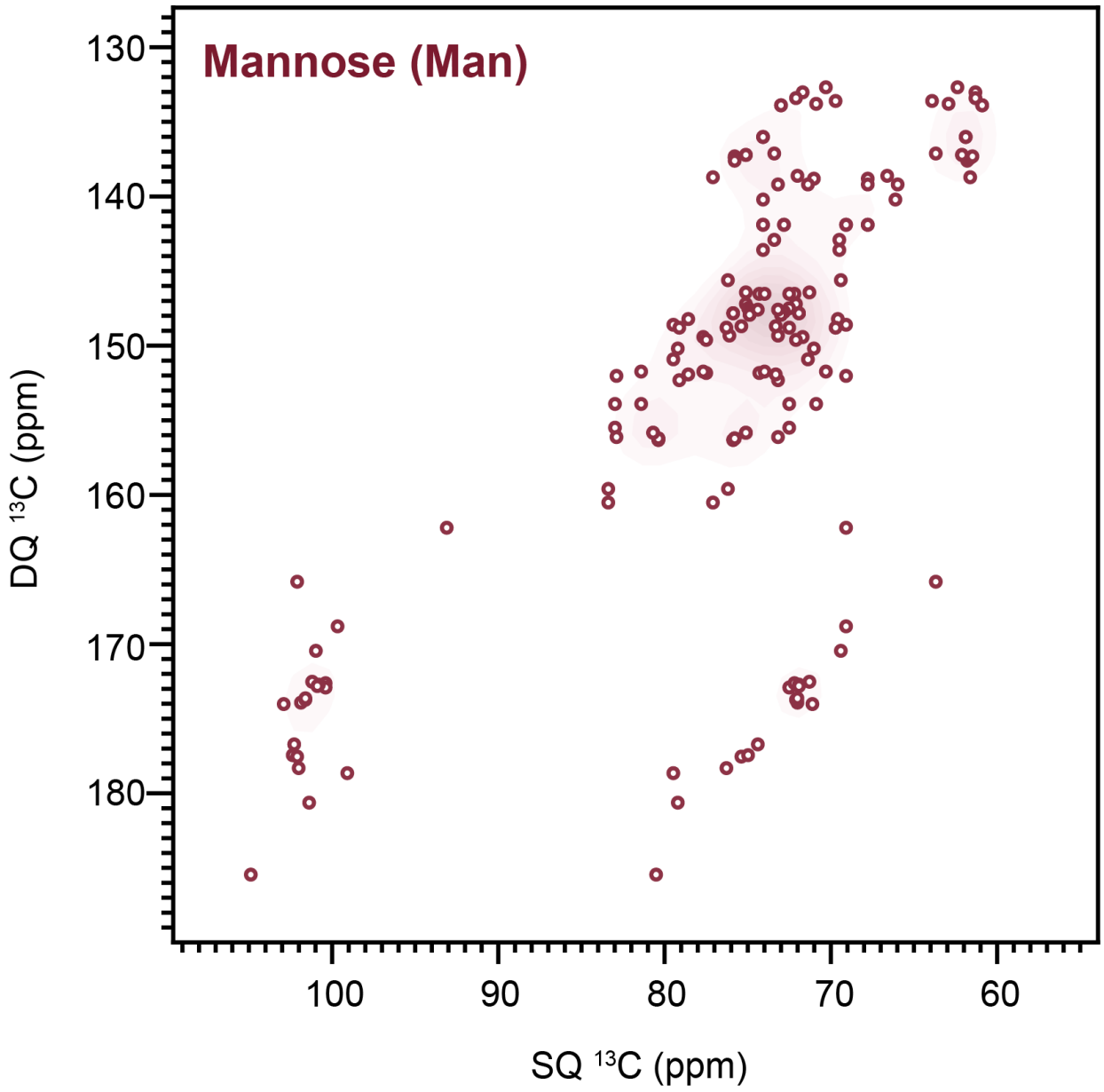

Figure S11.13 Projected 2D ^{13}C INADEQUATE correlation spectra of plant mannose residue.

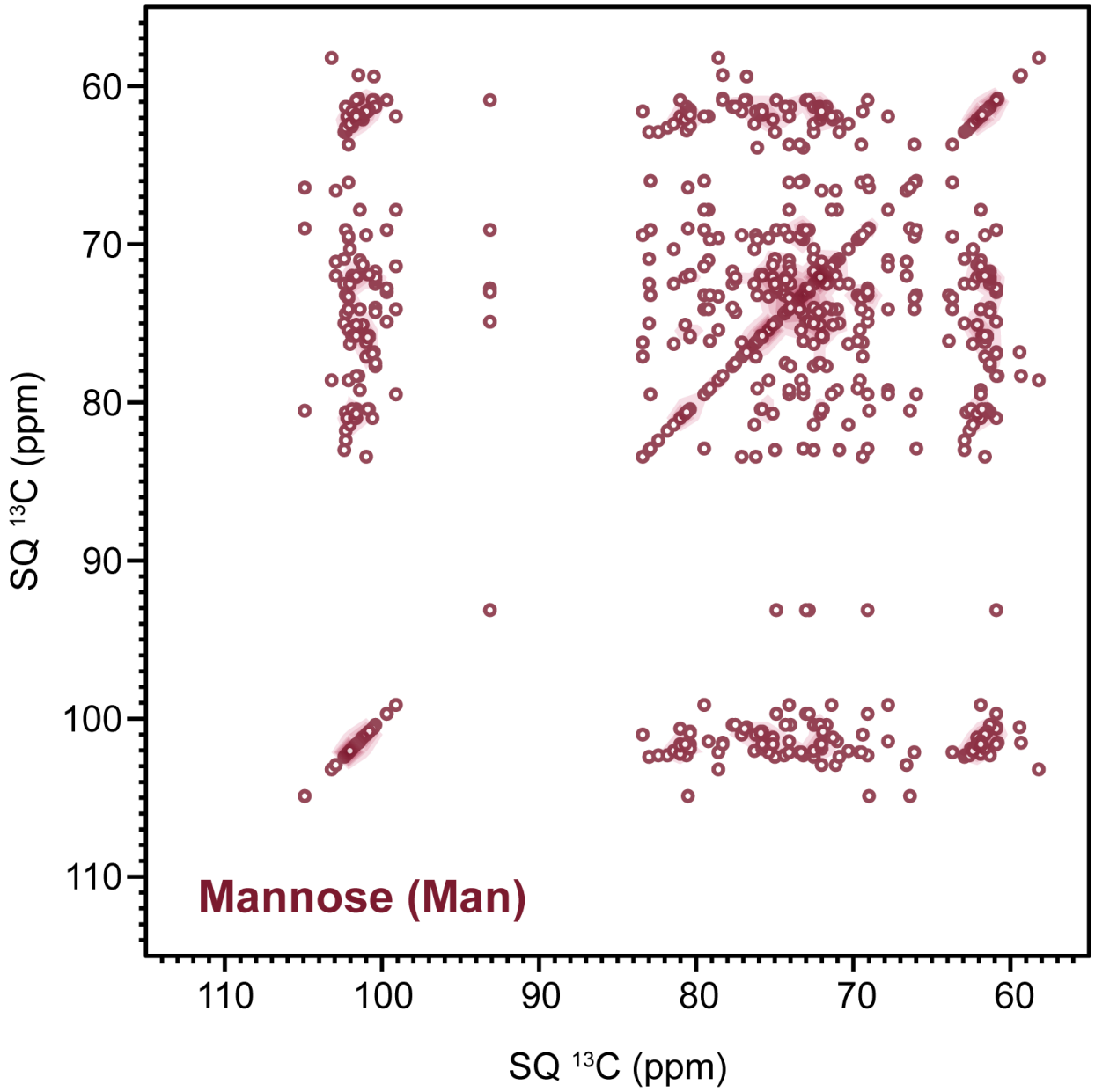

Figure S11.14 Projected 2D ^{13}C SQ-SQ correlation spectra of plant mannose residue.

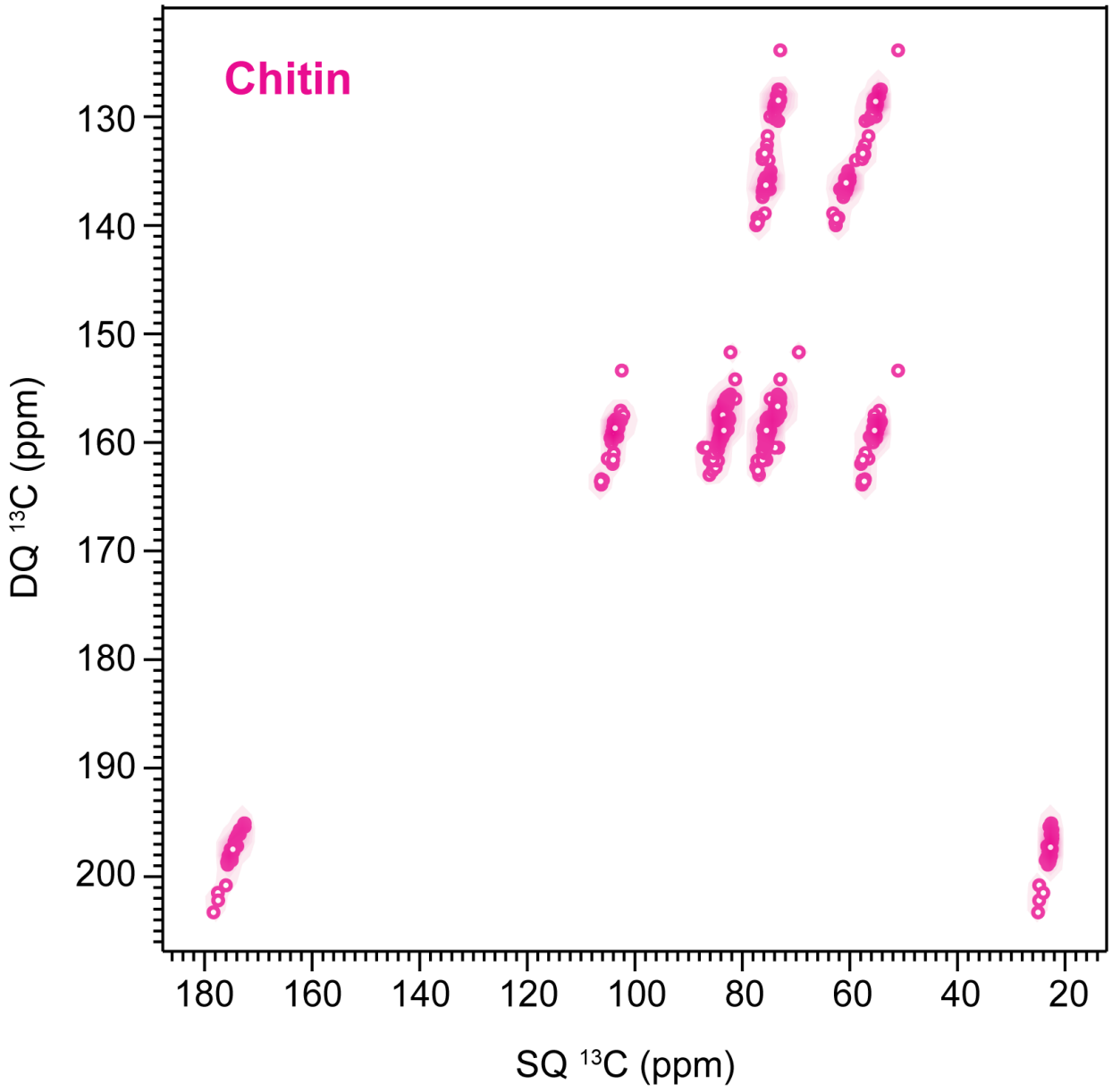

Figure S11.15 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi chitin.

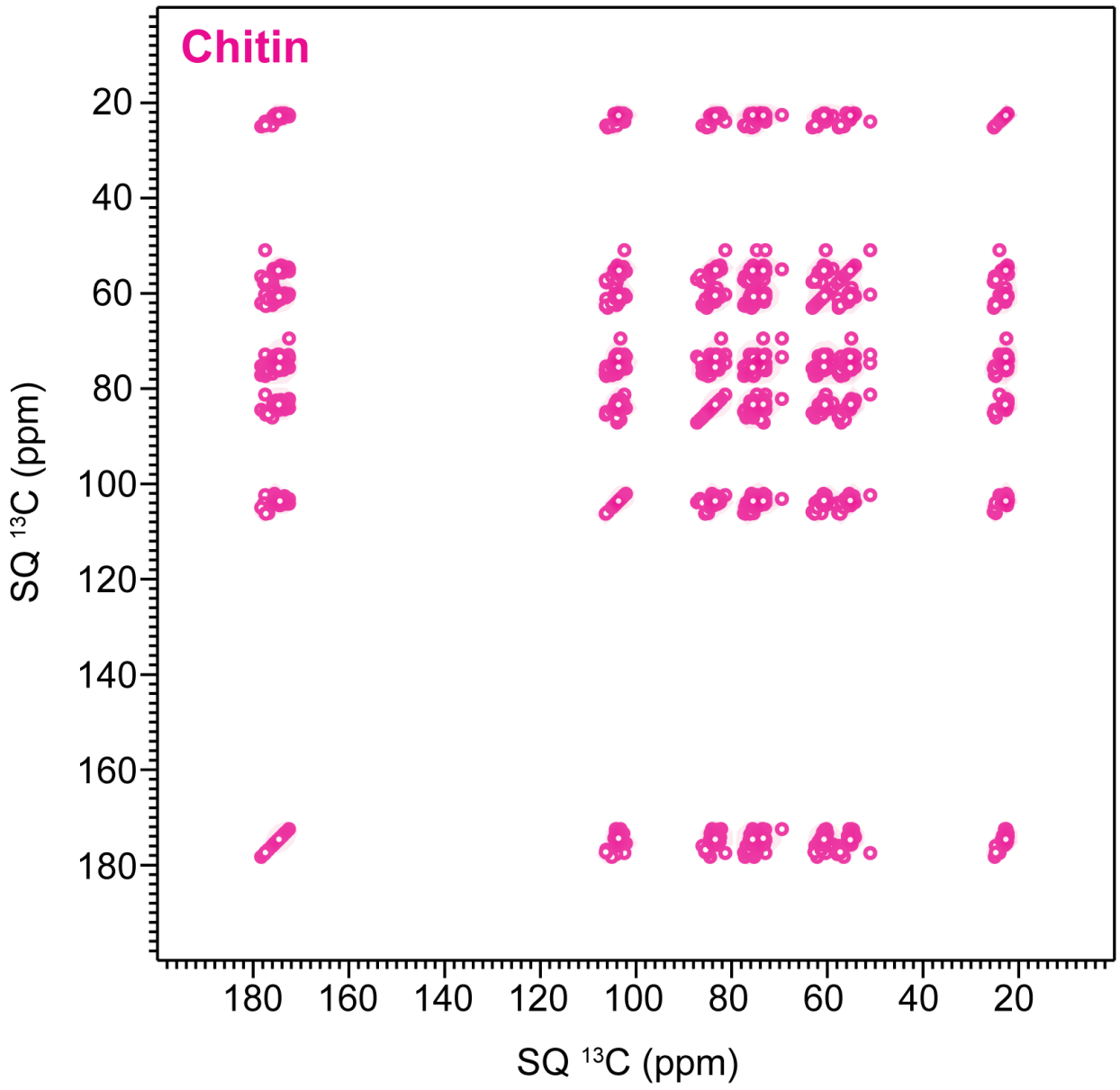

Figure S11.16 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi chitin.

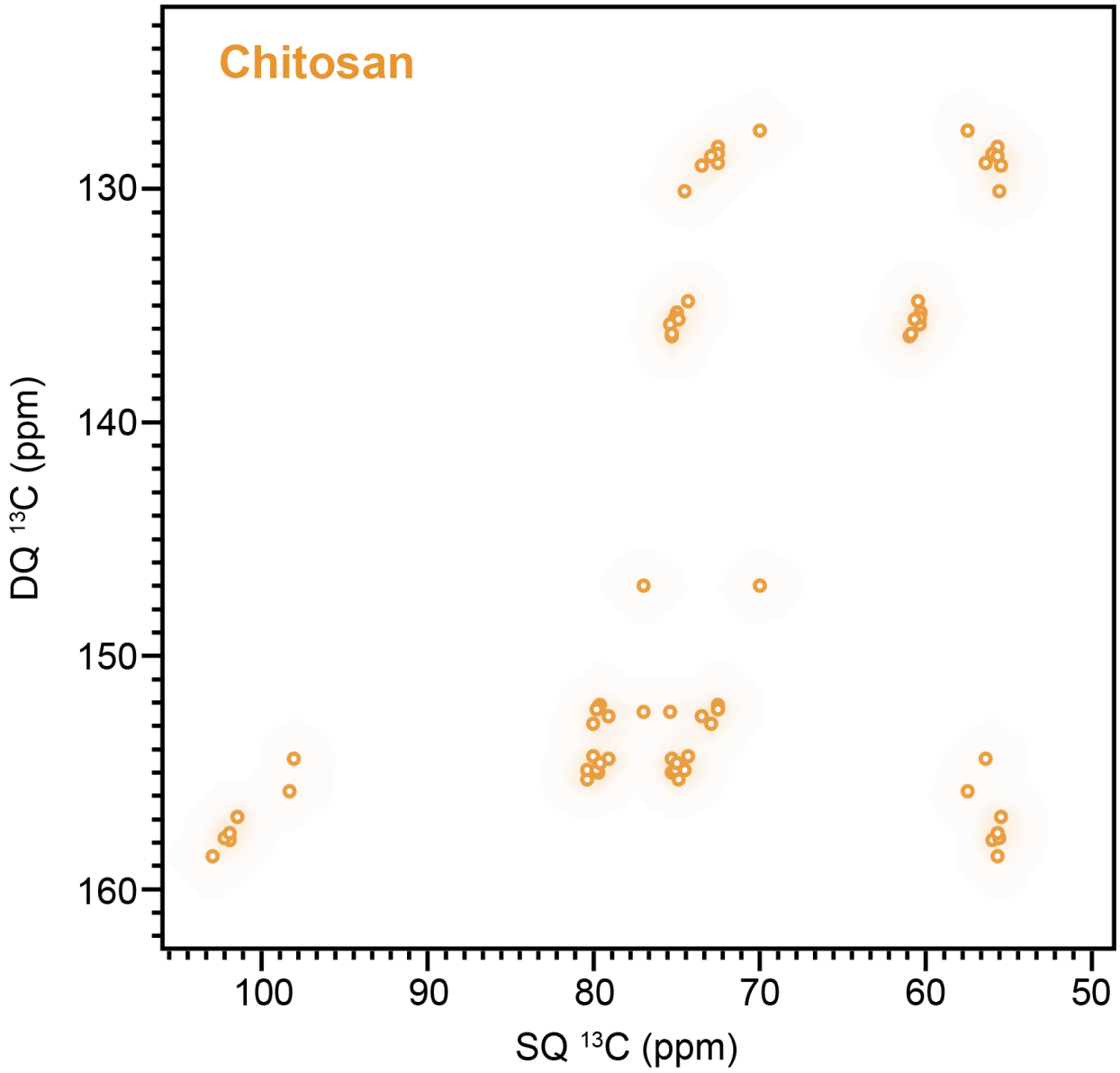

Figure S11.17 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi chitosan.

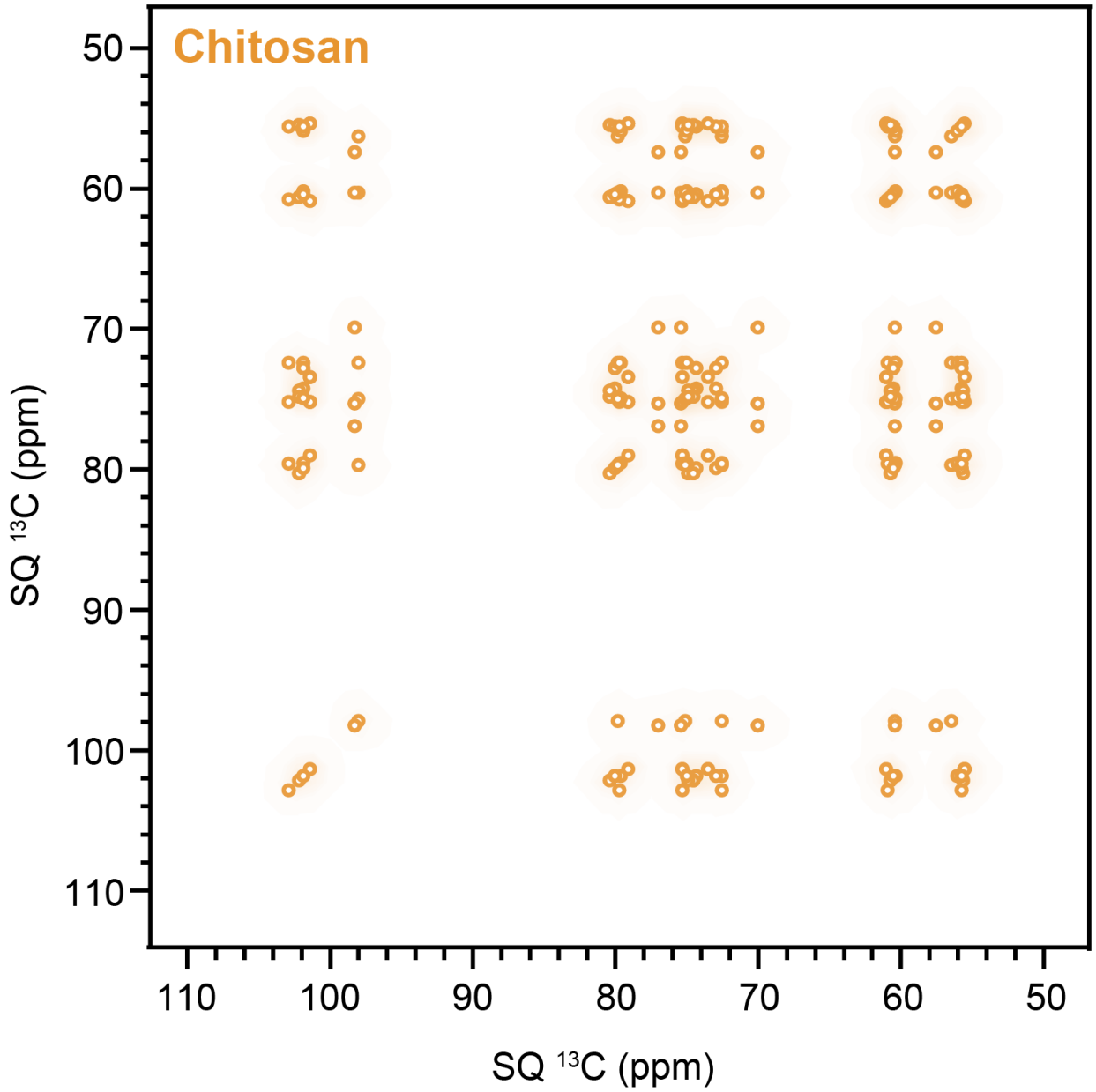

Figure S11.18 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi chitosan.

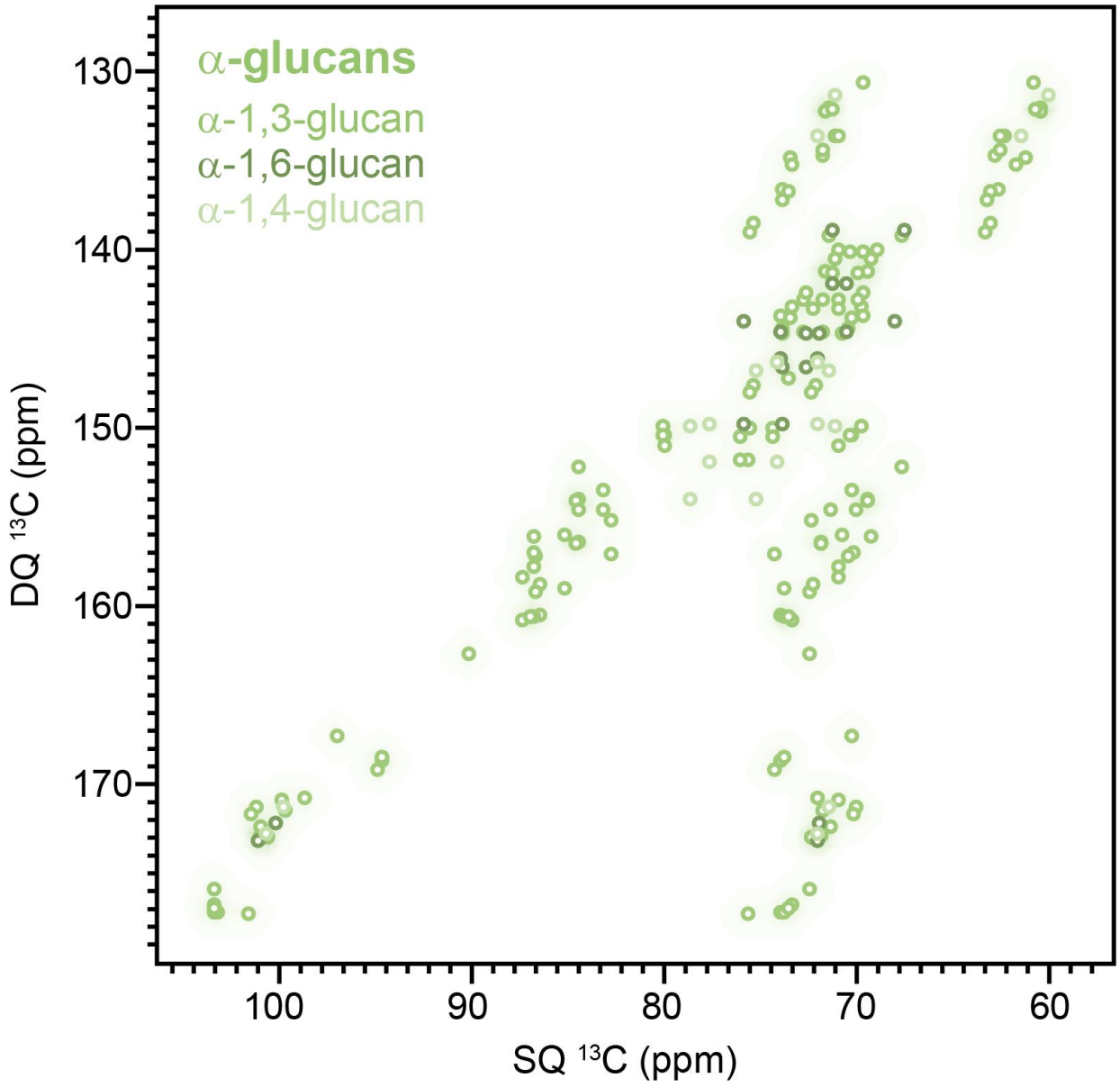

Figure S11.19 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi α-glucans.

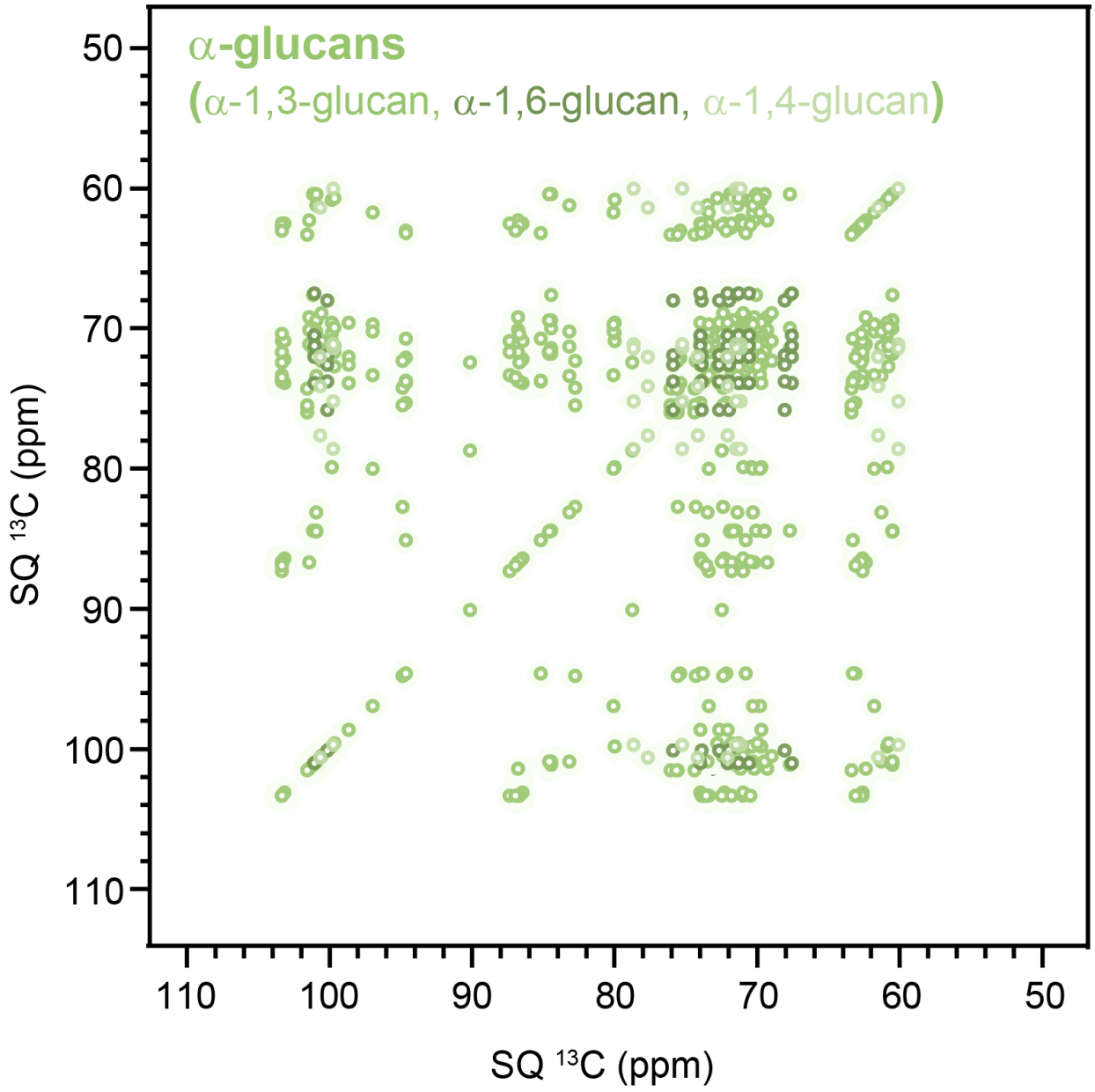

Figure S11.20 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi α-glucans.

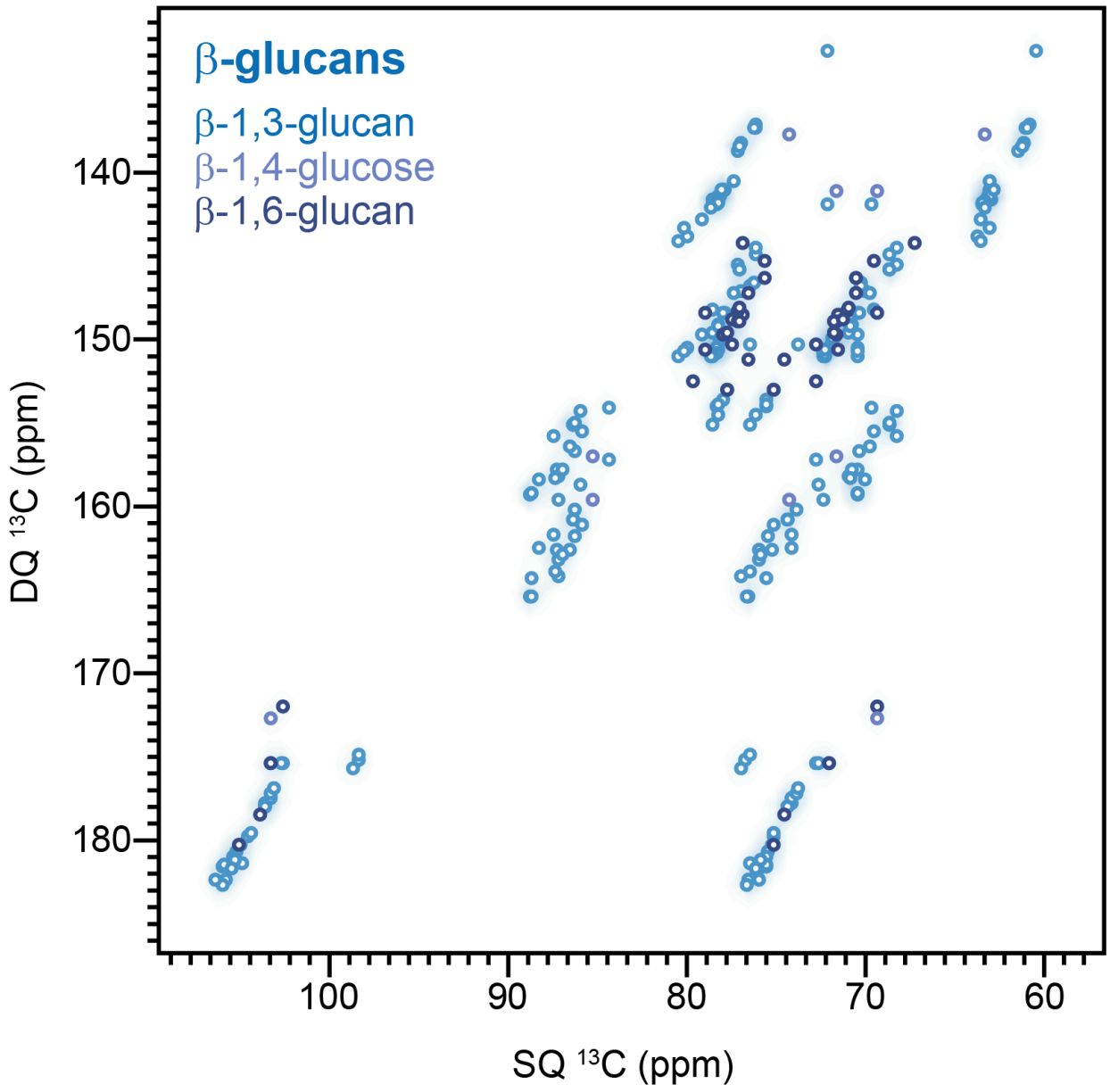

Figure S11.21 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi β-glucans.

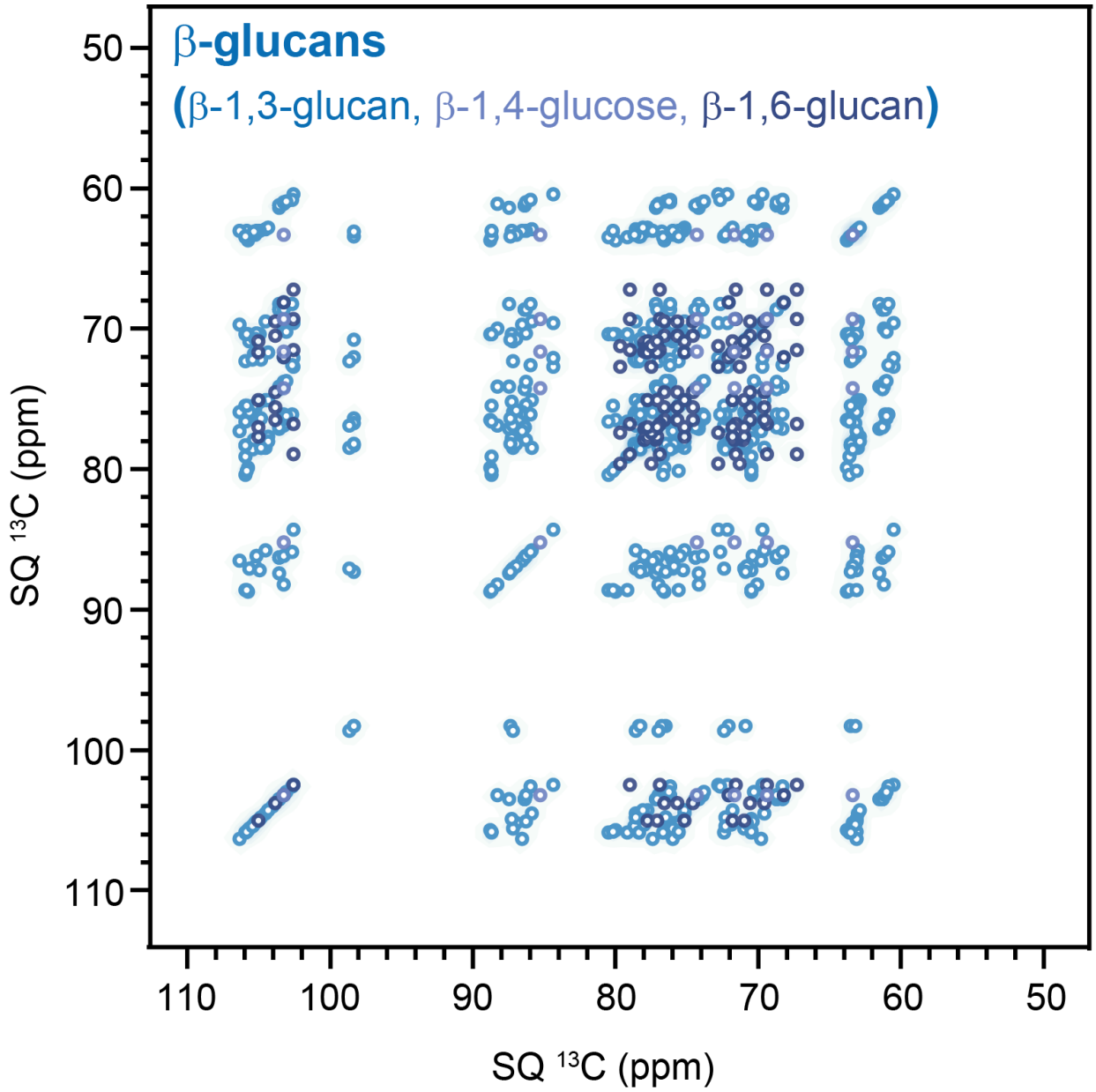

Figure S11.22 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi β-glucans.

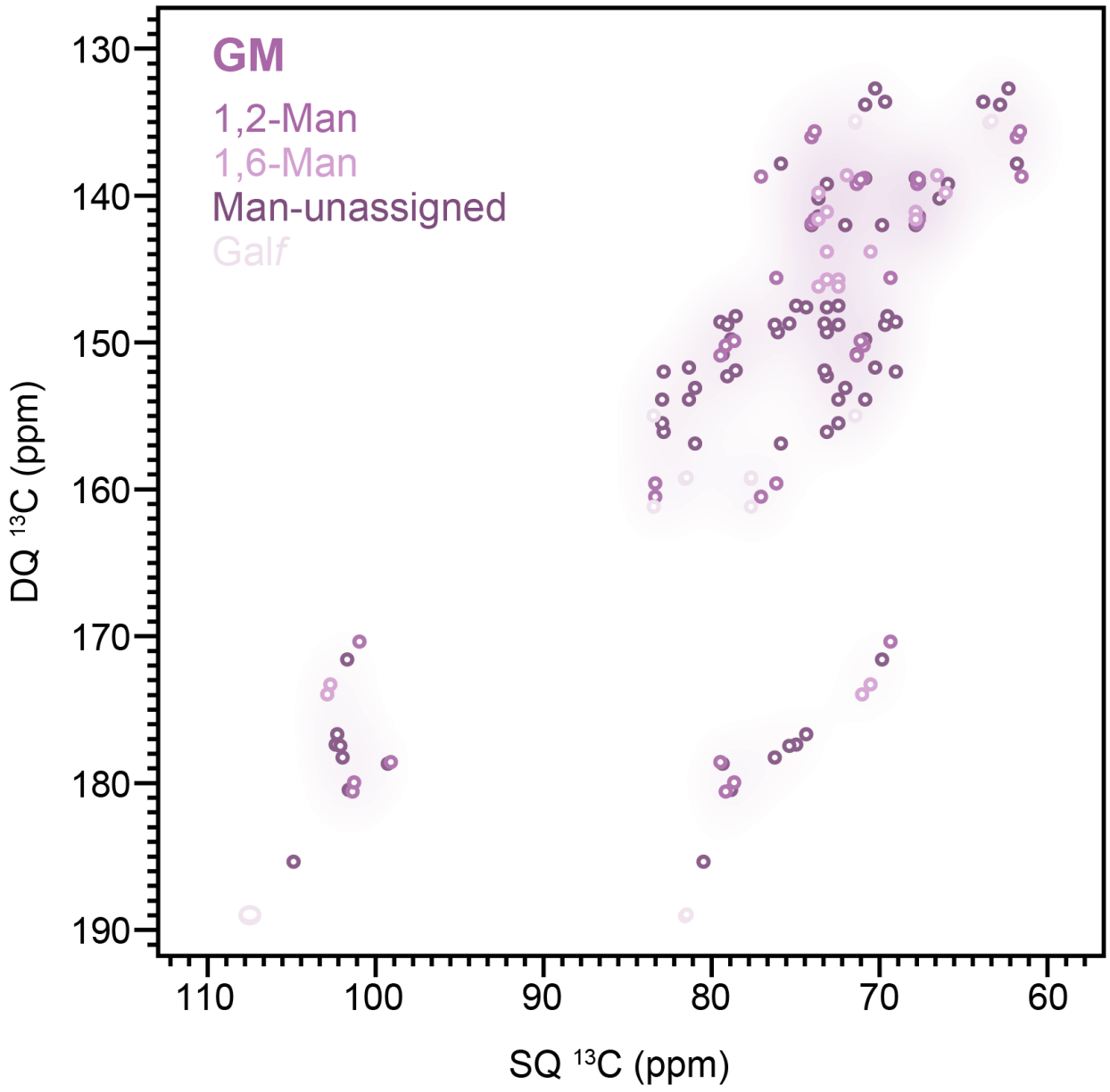

Figure S11.23 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi galactomannan (GM).

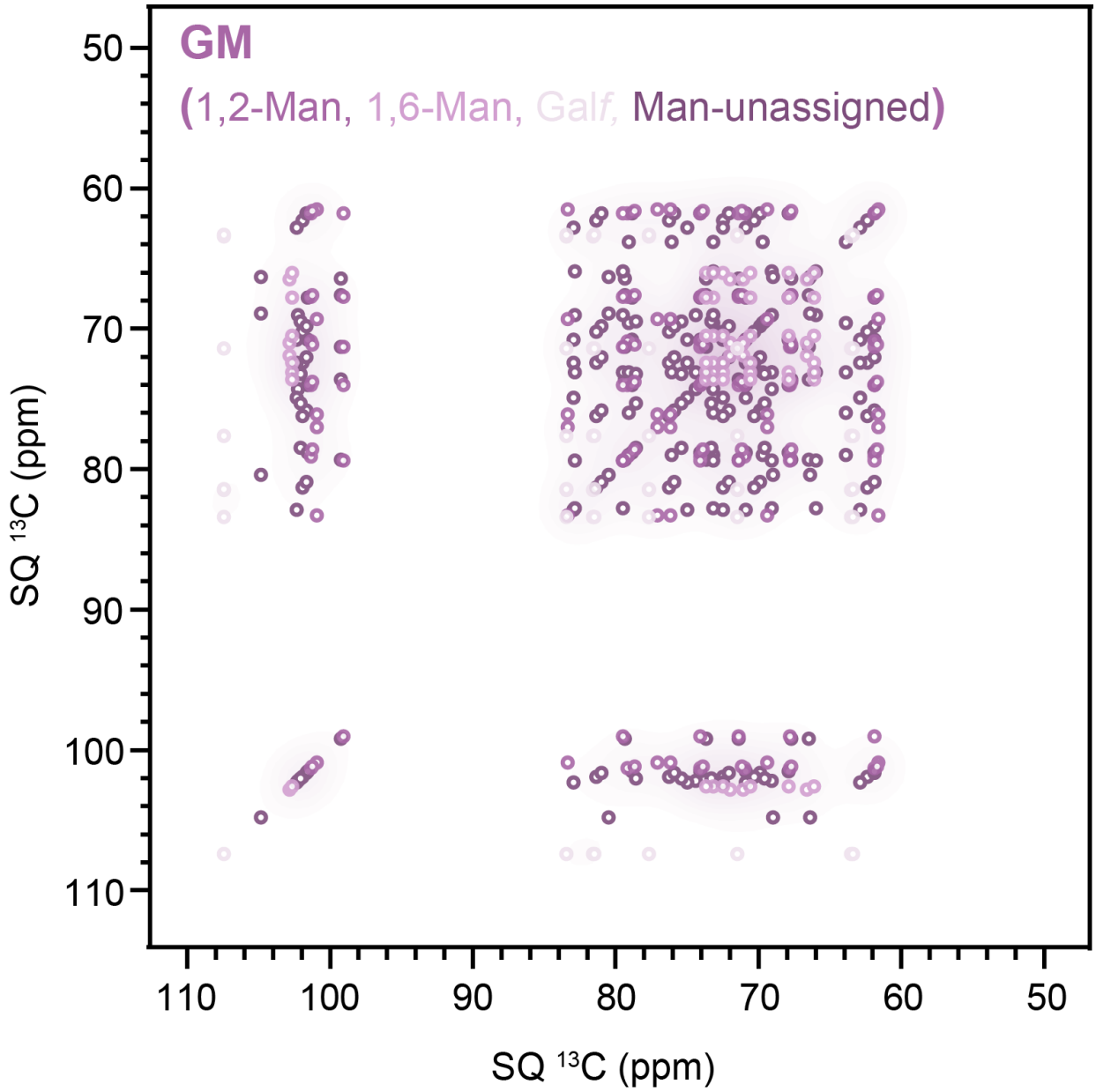

Figure S11.24 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi galactomannan (GM).

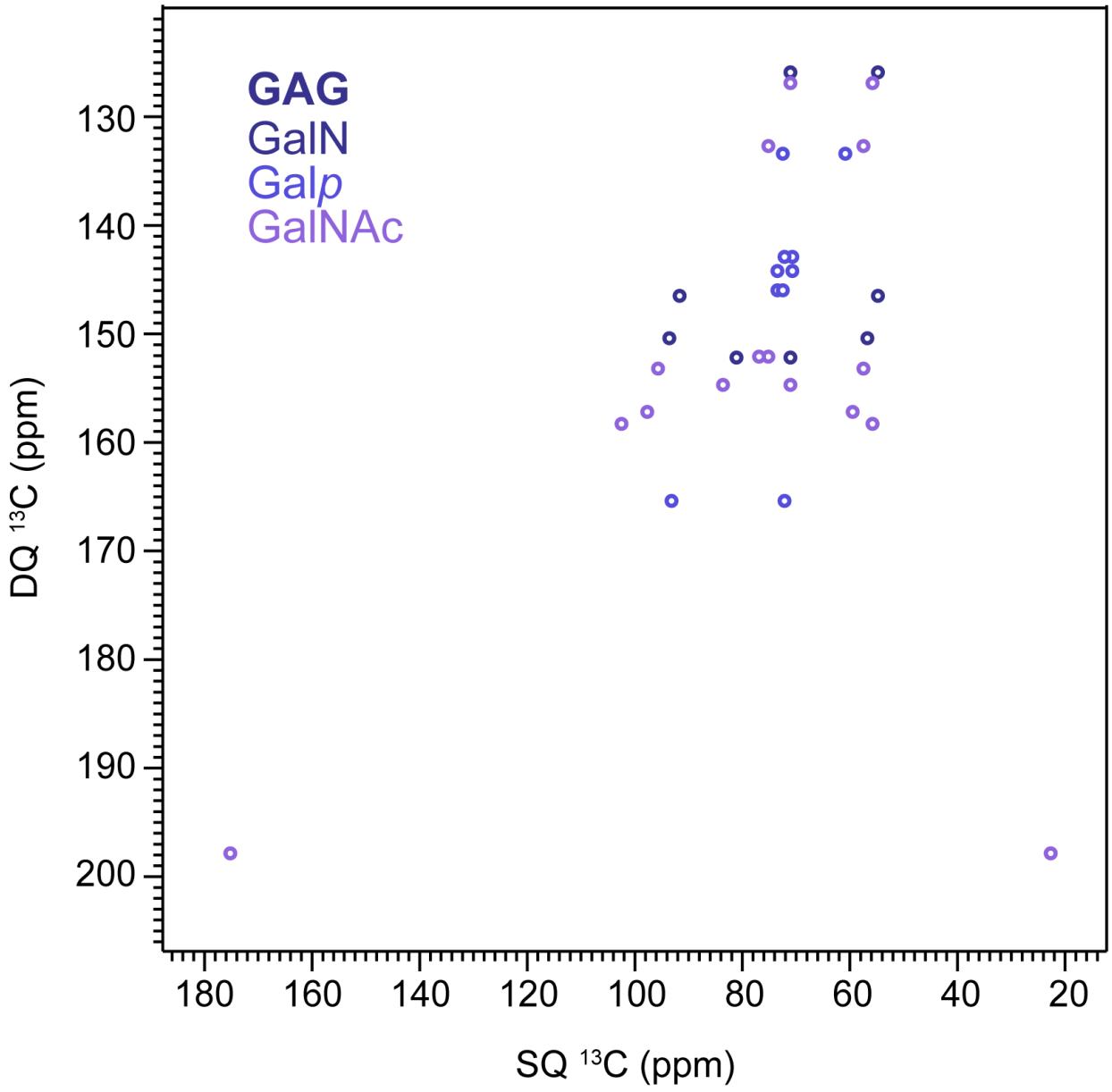

Figure S11.25 Projected 2D ^{13}C INADEQUATE correlation spectra of fungi galactosaminogalactan (GAG).

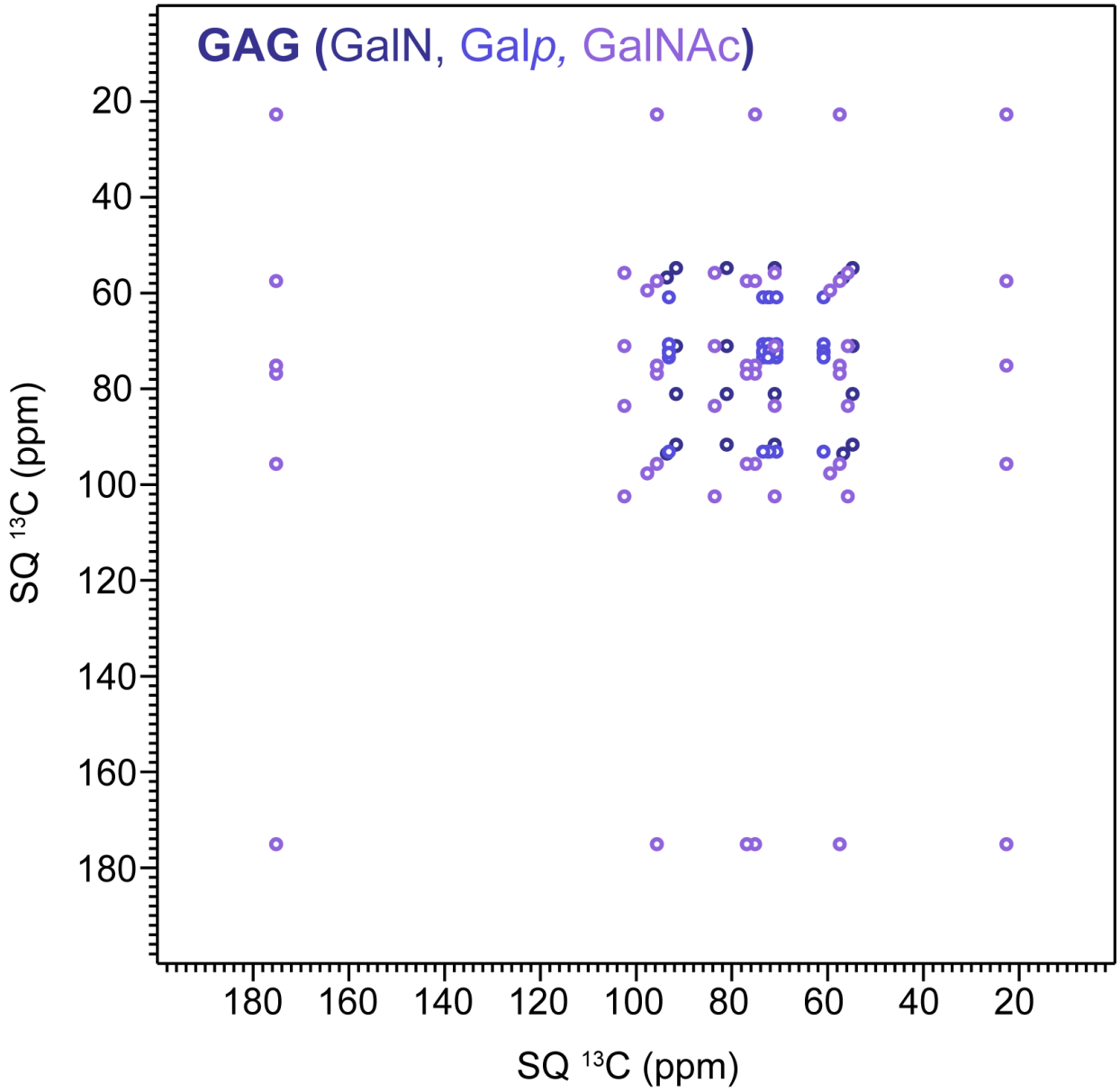

Figure S11.26 Projected 2D ^{13}C SQ-SQ correlation spectra of fungi galactosaminogalactan (GAG).