nature portfolio

Peer Review File

Emergence of Lignin-Carbohydrate Interactions During Plant Stem Maturation Visualized by Solid-State NMR

Corresponding Author: Professor Tuo Wang

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript entitled "Emergence of Lignin-Carbonhdyrate Interactions During Plant Stem Maturation Visualized by Solid-State NMR" studies the 13C labeled Arabidopsis mutant stems and discusses the role of the intimate mixture of S-lignin with xylan in the mechanical resistance of the tissue. The manuscript contains many high-quality NMR data, and some of their interpretation is very straightforward, such as the difference in quantity of S-lignin along the height of the stem. The main discussion uses the cross-peaks in the correlation spectra as a proxy for the "interaction" between lignin and polysaccharides, which is far more delicate and difficult to understand.
First, the term interaction assumes an energy stabilization between two molecule fragments. Here, the author uses it as a synonym for spatial proximity within 1 nm, which can lead to confusion when taken out of context. Another problem is the assignment of peaks due to the co-existence of many types of molecules. For example, in Fig. 1e(WT16A), there is a cross peak labeled S3/5-2/6 at 105-153 ppm (omega2-omega1), but 105 ppm coincides with the C1 of cellulose, and the same peak seems to be labeled S3/5-i4 in figure 4b. I do not know if the assignments are as clear as claimed using the arrows in Figure 4c. The polysaccharide region, 60-110 ppm, should also contain the propane carbons of lignin linked to one oxygen atom. The discussion about the difference in affinity between S-lignin and G-lignin is also delicate if the spatial distributions are not the same, with G-lignin being more distributed in the middle-lamella where the polysaccharide fraction is small, where simply statistically it should not find many carbohydrates in the proximity. Fig. S4 WT40 shows the different localization of S- and G-lignins. However, I cannot see for other pictures if they are co-localized or if the resolution does not allow the differentiation. The authors can express their opinion on this.
Other minor concerns are about the extent of generalization. For example, in lines 53-55, the authors mention S1, S2, and S3, but those are for the tracheid of gymnosperms, and I have not seen a report on angiosperm cell walls having these three layers. The existence of microfibrils (line 55) is also questionable, as it can be simply a fracture artifact.
There are mentions of "polar functionalities" (l 477) and electrostatic contacts (L85, 476, 494), do they refer to OH...O hydrogen bonding from OH groups of xylan to the lone pair of the ether oxygen? The authors probably have a clear idea, but the expression is very evasive. I cannot find clear NMR evidence of such electrostatic contact, but can authors pinpoint the data? The proximity of S-C3/C5 to xylose may support such interaction, but the distance linking C3-O...H-O-C(xyl) is quite far if one considers only the C-C correlation.
I conclusion, I recommend publication after taking into account the above concerns.

Reviewer #2

(Remarks to the Author)
This manuscript describes the solid-state NMR-based exploration of the major polymers of the plant secondary cell wall. I am not an NMR expert, so I cannot comment on many of the results, but I do have some comments on some of the other results presented.

Major comments

Starting on line 97, the manuscript states “The fah1-2 mutant has an undetectable transcript level of ferulate 5-hydroxylase (F5H), which hydroxylates coniferyl alcohol/coniferaldehyde (the G-lignin precursor) at its C-5 position to form S-monolignols, leading to defective S-lignin biosynthesis but elevated or unchanged G subunit content.” The authors cite references 37-40, but references 38 and 40 are not relevant or appropriate references to back up this statement. The authors

should instead cite these two references:
Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, et al. 1999. Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc. Natl. Acad. Sci. U S A 96: 8955-60
Humphreys JM, Hemm MR, Chapple C. 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc. Natl. Acad. Sci. U S A 96: 10045-50

On line 101-102, the authors state "…we managed to resolve the temporal functions of different lignin types…" What is a temporal function? Could the authors rephrase for clarity?
My major comments relate to the paragraphs extending from lines 122 through 182. I have two major concerns. First, there is very little new being reported here. Most of these observations have been published by others using different (e.g. thioacidolysis or DFRC vs NMR, or sometimes the same (see observations based on Maule staining beginning on line 127) techniques. So, from this perspective, I think that this manuscript could be substantially shortened without loss of impact. The Discussion itself highlights this issue in that it focuses almost exclusively on lignin-carbohydrate interactions (as it should, given the title of the manuscript). Few if any of these more duplicative results are mentioned. My second concern relates again to proper citations. The authors fail to cite published work that has previously shown results comparable to what the authors are reporting here.

Minor comments

On line 76, the authors state "This approach has led to five advances in our understanding…" Reading the text that follows, I really can't identify those five items. This isn't a super important point, it's just something confusing that caused me to stumble as I read the manuscript.

Three times in the paragraph beginning on line 184, the authors refer to lignin content "decreasing" across the various segments they analyzed. I think that this is a very counter-intuitive way of describing the results because it takes them out of their obvious developmental context. The segments being described, and for emphasis I will put them in the order C, B, and A, are taken along a developmental gradient of increasing age, and increasing lignin content. To describe lignin content as decreasing across samples A, B, and C seems to put more emphasis on how the authors did their analysis rather than on their biological context.

Line 186. Change mutant to mutation.

Reviewer #3

(Remarks to the Author)
Comments to the authors.

The manuscript entitled "Emergence of Lignin-Carbohydrate Interactions During Plant Stem Maturation Visualized by Solid-State NMR" by Xiao et. al. presents a well-executed study by using solid state NMR as the primary tool to probe the changes of lignin-carbohydrate interaction during plant stem maturation. The authors specifically investigate the potential roles of lignin subunits, S and G, in contributing to the mechanical property of the plant cell walls. Authors analyzed the inflorescence stem tissues at different heights(basal, middle, apical), which are harvested from the wild type and two lignin down-regulated mutants Arabidopsis (fah1, ref3, emphasize the different S/G ratios) with different ages. They found the relatively less lignin and lower S/G ratio deposited in the younger inflorescence stem tissue. The authors further demonstrate that S units form more stable interactions with cell wall polysaccharides compared to G units, as indicated by the mechanical differences observed between fah1 and ref3 mutants. Despite its reduced overall lignin content, ref3 maintains mechanical properties similar to the wild type, while fah1 exhibits a decrease in mechanical strength.
This work provides valuable insights into the role of lignin subunits in secondary cell wall architecture and their impact on plant mechanical properties. It adds important information in this challenging topic, which is certainly of interest to the plant cell wall community. Authors generated extensive NMR data in this work. The approach and methods are well developed. The data presentation is clear with high quality. The manuscript is overall well written.
While this study provides meaningful contributions, its novelty requires further clarification. The authors spent large amount of efforts to compare the lignin difference and S/G ratio over the stem tissues collected at different growth stages, which are well known (Rencoret et. al., 2010; Waliszewska et. al., 2019). The younger tissues analyzed, particularly the apical stem and 8-day-old samples, likely contain a substantial proportion of primary cell walls, which may confound the interpretation of lignin interactions. Additionally, the authors' prior work (Kirui et al., 2022, Fig. 3c) already demonstrated the dominant strong interaction of S units with polysaccharides compared to G units with polysaccharides in another dicot, which is one of the major findings in this work. The authors should explicitly differentiate this work from their previous studies and clarify its novel contributions.
On the other hand, authors' claim of direct link between the mechanical property difference observed on fah1 and ref3 mutants and their different S/G ratios seems weak if solely rely on the PDSD data, which only looks at the interactions of lignin subunits. As a complex heterogeneous polymer, lignin contains more than seven linkages, LCCs, and possible varying molecular sizes and cross linkings, which can all potentially contribute to the mechanical property changes observed. The manuscript does not provide sufficient evidence that changes in S/G ratio alone drive the observed mechanical differences. To strengthen this claim, the authors should consider including additional solution-state HSQC NMR and GPC analysis of extracted lignin to assess structural changes beyond S/G ratio differences. A discussion on whether mechanical differences arise from S and G unit proportions or broader lignin structural variations would be valuable.

Similar to the previous comment, authors focus their analysis only on the changes of lignin subunits between the mutants, but how do the other plant cell wall components response to the downregulated lignin content compare to those in the WT stem? Authors should consider to include monosaccharide analysis of the stem tissue they analyzed to monitor the composition change of the other wall components and briefly discuss how it may contribute to the mechanic difference observed, since increasing cellulose crystallinity is also observed by the NMR as stem ages.
Other minor comments as:
1. Figure 1b: The authors should clarify why the spectra were normalized at 72 ppm. Since primary cell wall content may vary with tissue age, additional explanation is needed on how spectra were analyzed.
2. Figure 1: The study only examines rigid lignin using cross-polarization (CP). Is there evidence of relatively mobile lignin? Can the authors provide an S/G/H ratio analysis for the mobile lignin fraction? Also, the sentences in lines 196-201 should be moved earlier in the text to improve logical flow.
3. Figure 1e: How were the 2D spectra normalized and compared? The intensity of the WT 16A spectrum appears slightly lower than the others—can the authors comment on this?
4. Figure 3e: In the Maule-staining images, the G and S distribution in ref3 does not appear as uniform as in WT. Could this uneven distribution contribute to its ability to maintain mechanical properties despite reduced lignin content?
5. Line 279, can author specify, in the ref3 16A sample in Fig 4a, where it shows the strong long-range interactions? Can authors please comment on how the 1s PDSD spectra were normalized and compared
6. Line 287, The authors state that S and G units interact with surface and internal cellulose chains in different patterns. Could they elaborate on how this conclusion was drawn from the 2D spectra?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript is significantly improved, but the term 'interaction' is often an overinterpretation. I like the expression in the abstract "molecular mixing pattern" as a more neutral and objective description. The authors measured the spatial distance through spin interactions; however, the proximity of the molecule does not necessarily imply a strong molecular interaction in the cell wall, contrary to what is observed in dilute systems, since the intermicrofibril distance is typically on the order of nanometers. In structural biology experiments, ligands are free to move in the solvent and attach to the protein only when the free energy of binding is significantly high. In the case of lignin-cellulose-hemicellulose, the system is dynamically blocked, and they are not free to move under ambient conditions, so the proximity does not necessarily imply interactions. I think replacing the term "interaction" with "spatial proximity" does not change the fundamental findings, but renders the description more accurate.

Mechanical properties are one of the essential aspects of the manuscript, but the data presentation is limited. I suppose the tensile measurement was performed on a humid sample; however, the state of the sample is not specified. Comparing the SS curve and the modulus, it appears that the authors are referring to the maximum tangent modulus, but this is not explicitly specified. If the SS curves were reproducible, it would show more distinct behaviors among the variants than the strength and modulus. Fah1 exhibits a higher initial rigidity than the other two, but doesn't display strain hardening, whereas ref3 exhibits strain hardening behavior. I would have loved to see the original data of the 6 SS curves for each type, but the "source data" did not contain the SS curves. I apologize for not raising this point in the first review round. I hope the authors consider these points.

Reviewer #2

(Remarks to the Author)
I am satisfied with the authors' responses to my initial review.

Reviewer #3

(Remarks to the Author)
I appreciate the authors' thoughtful responses to my previous comments. All of my concerns now have been fully addressed in the revised manuscript. I now fully support the publication of this manuscript in its current form.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am satisfied by the authors' response and corrections. However, I would like to remind the authors that if the center-to-center distance of microfibrils is on the order of a few nanometers and the microfibril diameter is also a few nanometers, the closest distance to any microfibril surface in the interstices is, on average, smaller than a nanometer. The manuscript can be published as is.

Responses to Reviewers' Comments

Reviewer 1:

The manuscript entitled "Emergence of Lignin-Carbonhdyrate Interactions During Plant Stem Maturation Visualized by Solid-State NMR" studies the 13C labeled Arabidopsis mutant stems and discusses the role of the intimate mixture of S-lignin with xylan in the mechanical resistance of the tissue. The manuscript contains many high-quality NMR data, and some of their interpretation is very straightforward, such as the difference in quantity of S-lignin along the height of the stem. The main discussion uses the cross-peaks in the correlation spectra as a proxy for the "interaction" between lignin and polysaccharides, which is far more delicate and difficult to understand.

We would like to thank the reviewer for the positive comment regarding the data quality and part of the interpretation. We also appreciate the comments regarding the discussion of cross peaks and interactions, which we have now improved in the maintext as described in the detailed responses below.

First, the term interaction assumes an energy stabilization between two molecule fragments. Here, the author uses it as a synonym for spatial proximity within 1 nm, which can lead to confusion when taken out of context.

We fully agree with the reviewers that varying definitions of technical terms across research fields can cause confusion, and that unifying them is challenging. Intermolecular cross peaks observed in NMR spectroscopy are widely used as indicators of subnanometer-scale packing and commonly refer to intermolecular interactions in fields such as structural biology, as well as solution and solid-state NMR spectroscopy. Technically, the cross peaks observed in this study represent at least dipole-dipole interactions between different polymers, and biochemically, they represent averaged structure where two polymers are physically packed with close proximity, as pointed out by the reviewer, and these polymer packing interfaces are stabilized by numerous interactions happening between the closely positioned molecules.

To reduce ambiguity, we have now added clarifications wherever possible:

For example, in the first subsection of Results, when intermolecular interactions first showed up, we have now added a sentence "These cross peaks annotate sub-nanometer physical contacts between lignin and polysaccharides, stabilized by intermolecular interactions between these two types of polymers."

In the third paragraph of Introduction, we have better defined as "the non-covalent interactions and spatial proximities between carbohydrates and lignin."

At the beginning of the section "Evolution of lignin-carbohydrate packing in different segments of WT and mutant stems," where these intermolecular cross peaks are analyzed in detail, we added a

clarification as well: “A long mixing period (1.0 s) facilitated further polarization transfer, uncovering numerous long-range intermolecular cross peaks that represent subnanometer-scale spatial proximity of two carbon sites stabilized by interactions between the two polymers….”

In the same section, the captions of Figure 4, as well as the Methods section, we have also replaced “interactions” with “cross peaks” for some occurrence. We hope these changes can help improve clarity.

Another problem is the assignment of peaks due to the co-existence of many types of molecules. For example, in Fig. 1e (WT16A), there is a cross peak labeled S3/5-2/6 at 105-153 ppm (omega2-omega1), but 105 ppm coincides with the C1 of cellulose, and the same peak seems to be labeled S3/5-i4 in figure 4b. I do not know if the assignments are as clear as claimed using the arrows in Figure 4c. The polysaccharide region, 60-110 ppm, should also contain the propane carbons of lignin linked to one oxygen atom.

We fully agree with the reviewer that this represents one of the most difficult aspects of the analysis. Indeed, for each project, the evaluation of these cross-peaks alone has typically required 1-2 years of dedicated effort from multiple group members. Such analytical difficulties are also the reasons why rather than focusing solely on a few select peaks, we interpret both clearly resolved cross-peaks and broader statistical trends across many partially ambiguous peaks. This approach provides a more robust understanding of lignin-carbohydrate contacts. In response to the reviewer’s comments, we have now added **a paragraph in the Discussion** section to better explain the methods, acknowledge the current limitations of this technique and emphasize the need for continuous methodological improvement and developments to improve spectral resolution and interaction specificity.

Regarding the few cross peaks specifically mentioned by the reviewer:

Figure 1e shows three short mixing spectra using only 0.1 s DARR, which primarily reveal intramolecular cross-peaks within each molecule. This explains the presence of the S3/5-2/6 cross-peaks, where the S2/6 signal resonates at 102-105 ppm, consistent with both literature and our current observations.

The 105 ppm region overlaps with cellulose C1 and may appear in long-range correlation spectra such as the 1.0 s PDSD. As the reviewer rightly pointed out, such spectral overlaps can introduce ambiguity, this is why we need to use the 0.1 s DARR spectrum as a control and overlap it with the 1.0 s PDSD spectrum, which allowed us to focus on new cross peaks that are intermolecular.

In addition, there are unambiguous intermolecular cross peaks, for instance, the cross-peak between S3/5 and i4 at (152, 89 ppm) in Figure 4b is unambiguous; the 88-89 ppm region is characteristic of the C4 carbon in interior cellulose and does not overlap with lignin signals, including those from linkers.

We believe the reviewer was referring to the S3/5-Xn1^{3f} peak in Figure 4b, where signals from Xn1 (and possibly i1 and s1) indeed share similar chemical shifts with S2/6. However, as we discussed above, we rely on a comparative approach: peaks that appear only in the 1.0 s mixing spectrum and not in the 0.1 s spectrum are taken as indicative of intermolecular carbohydrate-lignin cross-peaks. This is based on the observation that carbohydrate signals tend to dominate these long-range spectra, often overshadowing weak lignin intramolecular correlations. In light of this, we have reviewed and adjusted the peak label positions in Fig. 4b and Supplementary Figs. 13-15 to ensure they point to the correct locations.

A similar rationale applies to the 60-110 ppm region, where lignin linker carbons may theoretically appear. While these lignin aliphatic signals fall within this range, they are generally too weak to be clearly detected in solid-state NMR and are often masked by stronger carbohydrate signals. Additionally, the cross-sections extracted from the 2D spectra exhibit characteristic carbohydrate patterns, further confirming the dominance of carbohydrate signals in these interaction regions.

The discussion about the difference in affinity between S-lignin and G-lignin is also delicate if the spatial distributions are not the same, with G-lignin being more distributed in the middle-lamella where the polysaccharide fraction is small, where simply statistically it should not find many carbohydrates in the proximity. Fig. S4 WT40 shows the different localization of S- and G-lignins. However, I cannot see for other pictures if they are co-localized or if the resolution does not allow the differentiation. The authors can express their opinion on this.

We have now updated **Supplementary Figure 4**, which displays the Maule staining images, to include zoomed-in views for each stem height analyzed. As in the previous version, the WT 40A samples show a distinct G lignin signal localized in the middle lamella region between adjacent cells. In the 34A stems, a green signal indicative of G lignin can also be observed between neighboring cells, while the 22A stems exhibit only faint G lignin presence, limited to small regions at some tricellular junctions. In younger (shorter) stems, the current staining and imaging protocol lacks the resolution necessary to determine whether spatial separation of lignin subunits occurs at these earlier developmental stages. Due to the low abundance of this specific lignin type, the corresponding fluorescence signal is dim under laser scanning confocal microscopy and may not be detectable. Additionally, the lateral resolution limit of this imaging technique imposes an additional constraint, making it difficult to resolve the spatial distribution of lignin in the thinner cell walls of younger stems.

Other minor concerns are about the extent of generalization. For example, in lines 53-55, the authors mention S1, S2, and S3, but those are for the tracheid of gymnosperms, and I have not seen a report on angiosperm cell walls having these three layers. The existence of microfibrils (line 55) is also questionable, as it can be simply a fracture artifact.

Thank you for the advice. We have now removed the two sentences regarding cellulose layers and macrofibrils. Now the introduction of secondary cell wall formation is directly followed by the

introduction of lignification, which is the central focus of this study. This revision has helped to make the Introduction more concise and focused.

There are mentions of "polar functionalities" (l 477) and electrostatic contacts (L85, 476, 494), do they refer to OH...O hydrogen bonding from OH groups of xylan to the lone pair of the ether oxygen? The authors probably have a clear idea, but the expression is very evasive. I cannot find clear NMR evidence of such electrostatic contact, but can authors pinpoint the data? The proximity of S-C3/C5 to xylose may support such interaction, but the distance linking C3-O...H-O-C(xyl) is quite far if one considers only the C-C correlation.

Apologies for not making this clear in the previous version. While we refer to the hydroxyl groups in xylan as polar functional groups, xylan also contains substantial acetylation at the O2 and O3 positions. The polar carbonyl group (C=O) in these acetyl moieties, extending outward from the xylose backbone, may also facilitate close-range polar interactions with lignin. In lignin, both hydroxyl and methoxy substituents also contribute to polarity.

To improve accessibility for a broader readership, we have replaced “electrostatic interactions” with “non-covalent interactions” in the maintext. These contacts discussed here encompass a broad range of non-covalent interactions, including dipole-dipole interactions, London dispersion forces, hydrogen bonds, and other interactions. Although we do not have direct NMR evidence confirming the precise nature of these non-covalent interactions, we infer the involvement of methoxy groups in lignin based on differences in lignin-carbohydrate interactions observed among various mutants, and in previous reports, different plants. These mutants primarily differ in their lignin composition, namely, the relative abundance of G- versus S-type units, and thus, the number of methoxy groups.

Regarding potential hydrogen bonds, certain lignin-carbohydrate cross-peaks are evident in the 0.1 s mixing spectrum (e.g., Fig. 4b, signals between S3/5 and C4/C3), corresponding to internuclear distances of up to ~6 Å between interacting ^{13}C atoms. Considering typical geometries, the actual distances between lignin C3/5-O···H-O-C(xylan) moieties could fall within the distance range, or potentially closer for other polar interactions.

We have now clarified this point in the Discussion section.

I conclusion, I recommend publication after taking into account the above concerns.

We appreciate the reviewer’s support for the publication of this study, as well as the valuable guidance on improving the clarity regarding the cross peaks and interactions.

Reviewer 2:

This manuscript describes the solid-state NMR-based exploration of the major polymers of the plant secondary cell wall. I am not an NMR expert, so I cannot comment on many of the results, but I do have some comments on some of the other results presented.

Thanks. We found the comments very helpful, and we have addressed them below and we believe the changes have helped to improve the connection of our results with literature.

Major comments

Starting on line 97, the manuscript states "The fah1-2 mutant has an undetectable transcript level of ferulate 5-hydroxylase (F5H), which hydroxylates coniferyl alcohol/coniferaldehyde (the G-lignin precursor) at its C-5 position to form S-monolignols, leading to defective S-lignin biosynthesis but elevated or unchanged G subunit content." The authors cite references 37-40, but references 38 and 40 are not relevant or appropriate references to back up this statement. The authors should instead cite these two references:

Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, et al. 1999. Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc. Natl. Acad. Sci. U S A 96: 8955-60

Humphreys JM, Hemm MR, Chapple C. 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc. Natl. Acad. Sci. U S A 96: 10045-50

We have now included the papers as new references. We have also removed the original reference #40 Zhang et al. We kept the previous reference #38 Gou et al., a study by one of the coauthors, in which the lignin content change of the *fah1-2* and *ref3-3* were reported.

On line 101-102, the authors state "…we managed to resolve the temporal functions of different lignin types…" What is a temporal function? Could the authors rephrase for clarity?

Apologies for the ambiguity. We have now revised the sentence to: "…we managed to identify how different types of lignin function during the formation of secondary cell walls, and the distinct carbohydrates that interact with them as binding partners."

My major comments relate to the paragraphs extending from lines 122 through 182. I have two major concerns. First, there is very little new being reported here. Most of these observations have been published by others using different (e.g. thioacidolysis or DFRC vs NMR, or sometimes the same (see observations based on Maule staining beginning on line 127) techniques. So, from this perspective, I think that this manuscript could be substantially shortened without loss of impact. The Discussion itself highlights this issue in that it focuses almost exclusively on lignin-carbohydrate interactions (as it should, given the title of the manuscript). Few if any of these more duplicative results are mentioned. My second concern relates again to proper citations. The authors fail to cite

published work that has previously shown results comparable to what the authors are reporting here.

We agree with the reviewer that the results in the first subsection of the Results are largely consistent with prior literature. This section serves as a foundational step, allowing readers to familiarize themselves with solid-state NMR spectra and analysis before addressing more complex findings. We need this section because it is based on the same plant samples at different developmental stages, which paves the way for subsequent analysis of lignin–carbohydrate interactions. This is also the first time when such compositional changes have been observed directly within intact cell wall samples, without extraction or dissolution procedures being applied.

Following the reviewer's advice, we have made the description in the second paragraph of the first subsection of the Results section more concise. The third paragraph in this section is fully new as it probes lignin-carbohydrate interactions in the developing stem; therefore, we kept this paragraph.

We thank the reviewer for highlighting the importance of citing studies that use complementary techniques. In the same paragraph, we have now added a clarification: "These solid-state NMR results observations of intact cell walls align with previous studies using chemical and spectroscopic methods—such as derivatization followed by reductive cleavage, thioacidolysis, and solution NMR—performed on extracted or dissolved cell walls[30,31,33, 37-41]." The references added are listed below:

Derivatization followed by reductive cleavage (DFRC):
Schilmiller et al. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. *Plant J.* (2009) 60, 771-82.

Meyer et al. Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* (1998) 95, 6619-23.

Thioacidolysis (TA): Patten et al. Probing native lignin macromolecular configuration in *Arabidopsis thaliana* in specific cell wall types: further insights into limited substrate degeneracy and assembly of the lignins of ref8, fah 1-2 and C4H::F5H lines. *Mol. Biosyst.* (2010) 6, 499-515.

Solution NMR, TA, and DFRC: J M Marita, J Ralph, R D Hatfield, C Chapple. NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate 5-hydroxylase. *Proc. Natl. Acad. Sci. U. S. A.* (1999) 96, 12328-32.

Rencoret et al. Lignin Composition and Structure in Young versus Adult Eucalyptus globulus Plants. Plant Physiol. 155, 667-682 (2011).

In addition, two papers Gou et al. from the coauthors' group C.-J. Liu have also been added here.

For the maule staining, we have already included related references where this technique was used to observed G and S-lignin distribution: Pradhan Mitra, P. & Loqué, D (2014). Yamashita et al. (2016)

We hope these references help us better connect with the existing literature. We fully recognize that the field, particularly because of the interdisciplinary nature of this study, may include many other important works. We would greatly appreciate any additional suggestions for references we may have missed.

Minor comments

On line 76, the authors state "This approach has led to five advances in our understanding…" Reading the text that follows, I really can't identify those five items. This isn't a super important point, it's just something confusing that caused me to stumble as I read the manuscript.

We have now adjusted the writing by clearly mentioning the five conceptual findings with "first", "second", etc. to help the readers.

Three times in the paragraph beginning on line 184, the authors refer to lignin content "decreasing" across the various segments they analyzed. I think that this is a very counter-intuitive way of describing the results because it takes them out of their obvious developmental context. The segments being described, and for emphasis I will put them in the order C, B, and A, are taken along a developmental gradient of increasing age, and increasing lignin content. To describe lignin content as decreasing across samples A, B, and C seems to put more emphasis on how the authors did their analysis rather than on their biological context.

Following the reviewer's advice, we have now rephrased this sentence as "…where lignin content progressively increased in segments C, B, and A, along a developmental gradient of increasing age."

Line 186. Change mutant to mutation.

Thanks. Changed.

Reviewer 3:

The manuscript entitled “Emergence of Lignin-Carbohydrate Interactions During Plant Stem Maturation Visualized by Solid-State NMR” by Xiao et. al. presents a well-executed study by using solid state NMR as the primary tool to probe the changes of lignin-carbohydrate interaction during plant stem maturation. The authors specifically investigate the potential roles of lignin subunits, S and G, in contributing to the mechanical property of the plant cell walls. Authors analyzed the inflorescence stem tissues at different heights(basal, middle, apical), which are harvested from the wild type and two lignin down-regulated mutants Arabidopsis (fah1, ref3, emphasize the different S/G ratios) with different ages. They found the relatively less lignin and lower S/G ratio deposited in the younger inflorescence stem tissue. The authors further demonstrate that S units form more stable interactions with cell wall polysaccharides compared to G units, as indicated by the mechanical differences observed between fah1 and ref3 mutants. Despite its reduced overall lignin content, ref3 maintains mechanical properties similar to the wild type, while fah1 exhibits a decrease in mechanical strength.

This work provides valuable insights into the role of lignin subunits in secondary cell wall architecture and their impact on plant mechanical properties. It adds important information in this challenging topic, which is certainly of interest to the plant cell wall community. Authors generated extensive NMR data in this work. The approach and methods are well developed. The data presentation is clear with high quality. The manuscript is overall well written.

We deeply appreciate the encouraging comments on the significance of this study. Below, we summarize the changes made to better highlight the novelty, strengthen connections to the literature, and incorporate new solution and solid-state NMR analyses comparing lignin and carbohydrate structures in WT and mutant samples.

While this study provides meaningful contributions, its novelty requires further clarification. The authors spent large amount of efforts to compare the lignin difference and S/G ratio over the stem tissues collected at different growth stages, which are well known (Rencoret et. al., 2010; Waliszewska et. al., 2019). The younger tissues analyzed, particularly the apical stem and 8-day-old samples, likely contain a substantial proportion of primary cell walls, which may confound the interpretation of lignin interactions. Additionally, the authors’ prior work (Kirui et al., 2022, Fig. 3c) already demonstrated the dominant strong interaction of S units with polysaccharides compared to G units with polysaccharides in another dicot, which is one of the major findings in this work. The authors should explicitly differentiate this work from their previous studies and clarify its novel contributions.

We apologize for not having clearly highlighted the novelty of our study earlier. We have restructured the third and fourth paragraph of the Discussion section to emphasize the novelty of this study. The two references have also been included in the revised manuscript.

Regarding the analysis on lignin composition on different samples and growth stages, this is a prerequisite step for further analyses such as the lignin-carbohydrate interactions for different samples and stages. In addition, this has not been characterized specifically for *Arabidopsis* by solid-state NMR.

Technically, this study presents three improvements over previous solid-state NMR studies. First, the lignin chemical shift map in the solid state (Fig. 3b) provides valuable information for the research community. Second, the use of signature spectral patterns to cross-compare xylan-lignin and pectin-lignin interactions across multiple samples (Fig. 6b) introduces a new protocol. Third, the large-scale statistical heatmap analysis of type-based lignin-carbohydrate interactions (Fig. 5) has not been applied to plant samples before. Together, these advancements aim to moderately increase the bandwidth and throughput of solid-state NMR analysis.

Conceptually, we have deciphered the distinct functions of different lignin units and various carbohydrate components, and we provide insights into the developmental stages at which these functions occur. This analysis required the use of a consistent plant system, and the *Arabidopsis* inflorescence stem proved to be the most suitable model. This contrasts with previous solid-state NMR studies that examined a broad range of plant systems without focusing on developmental timing. We now show that older stems are not only richer in S-lignin, a finding reported previously in some species using other techniques, but also exhibit tighter associations between lignin and carbohydrates. For the first time, we report *in muro* and at the molecular level specific interactions between methylated pectin and G-lignin during early-stage lignification. Additionally, we demonstrate that the pattern of lignin-carbohydrate mixing, rather than lignin content alone, plays a more critical role in determining lignocellulose properties. Notably, this is also the first observation that S-lignin alone can efficiently stabilize lignin-carbohydrate interactions

In addition, a unique contribution of this study is the confirmation of the use of *Arabidopsis* inflorescence stems as a model system for solid-state NMR investigations of lignification. This system allows analysis along the axis gradient and across developmental stages and growth periods, with the additional advantage of easy isotopic enrichment. It opens new opportunities to investigate a wide range of mutant lines that affect the biosynthesis of lignin and various polysaccharides, enabling assessment of their impact on lignin-carbohydrate interactions and mechanical properties.

On the other hand, authors' claim of direct link between the mechanical property difference observed on fah1 and ref3 mutants and their different S/G ratios seems weak if solely rely on the PDSD data, which only looks at the interactions of lignin subunits. As a complex heterogeneous polymer, lignin contains more than seven linkages, LCCs, and possible varying molecular sizes and cross linkings, which can all potentially contribute to the mechanical property changes observed. The manuscript does not provide sufficient evidence that changes in S/G ratio alone drive the observed mechanical differences. To strengthen this claim, the authors should consider including additional solution-state HSQC NMR and GPC analysis of extracted lignin to assess structural changes beyond S/G ratio

differences. A discussion on whether mechanical differences arise from S and G unit proportions or broader lignin structural variations would be valuable.

Following the reviewer's suggestion, we processed and dissolved the same materials previously used for solid-state NMR measurements to obtain solution-state NMR data. The results are now presented in two new figures: **Supplementary Figures 11 and 12**. The experimental details are provided as a new **Supplementary Table 6**, as well as a new section in the Methods. We have also added a new paragraph at the end of the section "Mapping the lignin composition and carbohydrate contacts in WT and mutant stems" to describe these results and to propose lignin linkage patterns as a potential contributor to cell wall architecture and mechanics, though further studies are needed to clarify this connection.

Briefly, solution NMR results allowed us to observe complex lignin linkage patterns across the samples. Resinol linkages appeared at similar levels across all samples. Phenylcoumaran linkers were reduced in the S-lignin-rich *ref3* mutant compared to *fah1* and WT. The β-aryl ether was the dominant linkage in all cases. While the Aα region showed similar patterns, variations in the Aβ(G) signals were noted in *fah1* and *ref3*, though further resolution is beyond the scope of this work. Consistent with S/G ratio and total lignin content data, Aβ(S) was absent in *fah1* but present in WT and *ref3*. Notably, although overall Aβ signal intensity was lower in *ref3*, the Aβ(S) signal was relatively stronger than Aβ(G) in this mutant compared to WT. Additionally, lignin endgroups, such as cinnamyl alcohol (Iα and Iβ) and G′, were prominent in the G-rich WT and *fah1* samples but absent in *ref3*, suggesting that S-rich lignin forms more extended, interconnected chains, while G-rich lignin tends to form shorter, cluster-like polymers with more endgroups.

Similar to the previous comment, authors focus their analysis only on the changes of lignin subunits between the mutants, but how do the other plant cell wall components response to the downregulated lignin content compare to those in the WT stem? Authors should consider to include monosaccharide analysis of the stem tissue they analyzed to monitor the composition change of the other wall components and briefly discuss how it may contribute to the mechanic difference observed, since increasing cellulose crystallinity is also observed by the NMR as stem ages.

We have added an additional analysis of the carbohydrate content in the WT and two mutant samples, now presented as a new **Supplementary Figure 8**. The carbohydrate regions show only subtle differences, including minor changes in primary cell wall components such as the xyloglucan xylose sidechains (X) and certain galacturonic acid (GA) peaks. Small variations are also observed in the xylan (Xn) C4 region, but these are relatively minor. In contrast, the lignin region exhibits dramatic changes. The original solid-state NMR samples were unpacked and subsequently destroyed during preparation of the new solution-state NMR data, as detailed in our previous response, and we do not have more plant material for chemical compositional analysis. However, the solid-state NMR spectra of the WT and mutant samples already demonstrated overall consistent carbohydrate patterns.

Other minor comments as:

1. Figure 1b: The authors should clarify why the spectra were normalized at 72 ppm. Since primary cell wall content may vary with tissue age, additional explanation is needed on how spectra were analyzed.

The 72 ppm is the highest signal of the carbohydrate region, containing contributions from all (primary and secondary) cell wall polysaccharides. It is used to represent the overall carbohydrate content for normalization purposes. As can be seen from Fig. 1b, the variations in carbohydrate signals in the WT08A, WT12A and WT16A samples, are much less than 5% overall, but the lignin changes are in the range of 25-50% for these three samples. We have now clarified this point in the captions of Figure 1: "the highest signal of the carbohydrate region, which contains contributions from all cell wall polysaccharides." We have also clarified it in the maintext: "Spectra were normalized to the highest 72-ppm peak, which contained contributions from all cell wall polysaccharides. This normalization resulted in comparable carbohydrate signal intensities across samples, while lignin intensities varied significantly."

2. Figure 1: The study only examines rigid lignin using cross-polarization (CP). Is there evidence of relatively mobile lignin? Can the authors provide an S/G/H ratio analysis for the mobile lignin fraction? Also, the sentences in lines 196-201 should be moved earlier in the text to improve logical flow.

Thanks for the insightful comment. We have now added a new panel as Supplementary Fig. 3b, showing the analysis of the mobile lignin using 2 s DP spectra. We have also modified the caption of that figure to reflect this change. In the maintext, we have added a new clarification: "A simliar trend was also observed in the mobile fraction where the S content increased from 19-20% in WT08A and WT12A to 26% in WT16A (**Supplementary Fig. 3b**)." The H lignin content is very low, therefore were not analyzed.

For the sentences in original lines 196-201 mentioning the quantitative information of lignin content, they come after the CP results because the analyses of quantitative DP spectra are complicated and require the inclusion of prefactors based on CP analysis. We have now clarified this point in the maintext: "The analyses of DP spectra were complicated by the dominance of protein and lipid signals in this DP spectra and required comparison to the CP results with respect to the unambiguous lignin signals at 140-156 ppm as outlined in the Methods section."

3. Figure 1e: How were the 2D spectra normalized and compared? The intensity of the WT 16A spectrum appears slightly lower than the others—can the authors comment on this?

We have now replotted the WT16A spectrum in **Fig. 1e** to better match those of the other samples, and it does not change the observations. The reviewer is correct that WT16A was previously plotted with fewer contours, but even with this situation, the S-lignin signal is uniquely strong in this sample. The contours in WT12A look higher also because it mainly has only G-lignin so the signals are well-centered at each peak, rather than more dispersed as both G and S as in WT16A.

We have also clarified in the figure captions that the spectra were plotted so that the base contour levels of the diagonal peaks are comparable across all panels. This figure is not intended for quantification (which is presented later in Fig. 5), but rather to illustrate the relative abundance of different lignin units within each sample.

4. Figure 3e: In the Maule-staining images, the G and S distribution in ref3 does not appear as uniform as in WT. Could this uneven distribution contribute to its ability to maintain mechanical properties despite reduced lignin content?

Since the fiber cells (not vascular bundles of vessels cells) should be primarily responsible for the mechanical properties of stem tissues, and in both WT and *ref3* the fibers cells primarily contain S lignin, it is unlikely that these observed differences in distribution contribute to the mechanical differences.

5. Line 279, can author specify, in the ref3 16A sample in Fig 4a, where it shows the strong long-range interactions? Can authors please comment on how the 1s PDSD spectra were normalized and compared

For **Figure 4a**, it is not normalized (see the responses to question 3). It is only a rough comparison guided by similar base contour of the diagonal peaks followed by identical increment multiplification factors used in all three spectra. The strong long-range interactions are mainly referring to the region of 70-110 ppm in the direct dimension, and we have now clarified it in the maintext: "…*ref3*-16A showed fewer intramolecular cross peaks (in the region of 110-155 ppm) in the 0.1 s spectrum, likely due to its compromised lignin synthesis, but displayed strong long-range correlations (in the region of 70-110 ppm) in the 1.0 s spectrum."

In the captions of Fig. 4a, we have now clarified that the spectra were plotted so that the base contour levels of diagonal peaks are comparable. This section provides a qualitative comparison of the spectra. The intensity analysis and quantification are in in the later section related to Figure 5, where the peaks were analyzed with respect to the total spectral integral within each single cross section.

6. Line 287, The authors state that S and G units interact with surface and internal cellulose chains in different patterns. Could they elaborate on how this conclusion was drawn from the 2D spectra?

We have now elaborated this point by rewriting this section: "S and G units also exhibited interactions with cellulose. For example, both units exhibited cross peaks with the surface chains of cellulose such as G1, S1/4-s4 and S3/5-s4 and G3/4-s4 cross peaks (Fig. 4b). Meanwhile, interactions between S/G units and the interior chains of cellulose were detected, as evidenced by strong cross-peaks such as S3/5-i4 (152, 89 ppm) and G3/4-i4 (145, 89 ppm). These cross-peaks reflect the tight physical packing within the surface-proximal interior layer of cellulose microfibrils, with

components colocalized within a nanometer scale. A relatively weak but unambiguous cross-peak of S1/4-i4 (136, 87 ppm) was also observed, but only for the S unit, likely arising from relayed polarization transfer from S-lignin to either deeply embedded, inaccessible interior cellulose chains or interior chains with structural alterations due to their interactions with hemicellulose. Therefore, both S and G interacted with the surface chains of cellulose extensively but exhibited distinct patterns of interactions with internal chains."

Response to Reviewers' Comments

Reviewer #1

The manuscript is significantly improved, but the term 'interaction' is often an overinterpretation. I like the expression in the abstract "molecular mixing pattern" as a more neutral and objective description. The authors measured the spatial distance through spin interactions; however, the proximity of the molecule does not necessarily imply a strong molecular interaction in the cell wall, contrary to what is observed in dilute systems, since the intermicrofibril distance is typically on the order of nanometers. In structural biology experiments, ligands are free to move in the solvent and attach to the protein only when the free energy of binding is significantly high. In the case of lignin-cellulose-hemicellulose, the system is dynamically blocked, and they are not free to move under ambient conditions, so the proximity does not necessarily imply interactions. I think replacing the term "interaction" with "spatial proximity" does not change the fundamental findings, but renders the description more accurate.

We would like to thank the reviewer for the thoughtful advice. Regarding the distance, all the ^{13}C-^{13}C cross peaks detected refer to sub-nanometer physical contacts (typically below 5-7 angstrom), as stated throughout the manuscript. Therefore, they typically imply interfaces stabilized by numerous intermolecular interactions. Meanwhile, these observations are not related to the interfibrillar distance, which is significantly larger.

Following the advice, we have now replaced "interactions" with more specific terms such as "spatial proximities," "packing," "cross peaks," "contacts," or "correlations" when describing NMR observations in the Abstract and Results, depending on the intended meaning in each context. These edits included 24 changes. However, we would like to retain at least some instances of "interactions" in the manuscript, as it is accepted terminology in NMR structural biology and does not represent a novel usage introduced by this study. With the added explanation from the last round of revision and the inclusion of these alternative terms to reduce repetition, we believe the concept is now straightforward for readers to grasp.

We hope that the revisions have improved the clarity of the manuscript and made it more accessible to broader readership.

Mechanical properties are one of the essential aspects of the manuscript, but the data presentation is limited. I suppose the tensile measurement was performed on a humid sample; however, the state of the sample is not specified. Comparing the SS curve and the modulus, it appears that the authors are referring to the maximum tangent modulus, but this is not explicitly specified. If the SS curves were reproducible, it would show more distinct behaviors among the variants than the strength and modulus. Fah1 exhibits a higher initial rigidity than the other two, but doesn't display strain hardening, whereas ref3 exhibits strain hardening behavior. I would have loved to see the original data of the 6 SS curves for each type, but the "source data" did not contain the SS curves. I apologize for not raising this point in the first review round. I hope the authors consider these points.

We thank the reviewer for these comments on the mechanical data presented in the manuscript. The stem samples were frozen and then thawed immediately prior to the mechanical tests (see Methods description). The hydration state of the samples after thawing was maintained, i.e. the samples were never dried before testing. The samples were tested in ambient humidity but maintained inherent hydration from

the freeze-thaw procedure. The tests were conducted in a short time frame so drying during stretching was not a concern.

The reviewer is correct that the data in Fig. 2h are tangential modulus values of the stress-strain curve (averages of local derivatives in the 4.5-5% strain interval) for each stem tested.

We have now added the source data to the Source Data file for the individual stress-strain curves, each representing one of the eight stems tested per type, prior to averaging.

As the reviewer noted, reproducible stress-strain curves (with little variability between biological replicates) may have allowed for more mechanical insights to be drawn out of the data. The averaged stress-strain curves, presented in Fig. 2f, show that there was considerably variability between the eight biological replicates (stems) tested for each genotype (the SEM error bars for each genotype overlap considerably throughout the stress-strain curve). It is for this reason that the description of the mechanical data points was limited, as to not overextend the conclusions that may be drawn from this analysis. What can be concluded from the biological variability and overlapping data points is that despite the considerable differences in lignin content between these three samples, there are no significant differences between the tensile mechanical properties. This conclusion is in line with prior reports that showed that cellulose microfibril angle affected the tensile mechanics of woody tissues but that lignin content did not substantially contribute to the tensile mechanical properties (e.g. references cited in this manuscript: Gindl, W. and Teischinger, A. 2002 & Özparpucu, M. et al. 2019).

Reviewer #2

I am satisfied with the authors' responses to my initial review.

Thanks a lot!

Reviewer #3

I appreciate the authors' thoughtful responses to my previous comments. All of my concerns now have been fully addressed in the revised manuscript. I now fully support the publication of this manuscript in its current form.

Thank you so much for your support to the publication of this manuscript.

Response to Reviewers' Comments

Reviewer #1

I am satisfied by the authors' response and corrections. However, I would like to remind the authors that if the center-to-center distance of microfibrils is on the order of a few nanometers and the microfibril diameter is also a few nanometers, the closest distance to any microfibril surface in the interstices is, on average, smaller than a nanometer. The manuscript can be published as is.

Response: We would like to thank the reviewer for the information regarding the interfibrillar distance. The information is really helpful. We appreciate the support from the reviewer for the publication of this manuscript.