Journal of Biomolecular NMR
https://doi.org/10.1007/s10858-020-00304-2

ARTICLE

CCMRD: a solid-state NMR database for complex carbohydrates

Xue Kang[1] · Wancheng Zhao[1] · Malitha C. Dickwella Widanage[1] · Alex Kirui[1] · Uluc Ozdenvar[1] · Tuo Wang[1]

Received: 24 July 2019 / Accepted: 27 January 2020
© Springer Nature B.V. 2020

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.

Keywords Carbohydrates · Polysaccharides · Database · Solid-state NMR

Abbreviations

CCMRD	Complex Carbohydrates Magnetic Resonance Database
ssNMR	Solid-state NMR

Introduction

Complex carbohydrates play central roles in many biological processes such as energy storage, structural building, and cell recognition (Himmel et al. 2007; Lowe and Marth 2003; Raman et al. 2005). These biomacromolecules also form the basis for novel biomaterials such as the scaffolds developed for tissue engineering and the carriers for drug delivery (Collins and Birkinshaw 2013; Kang et al. 2015). 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 (Fig. 1) (Albersheim et al. 2011; Aspinall 1983; Buchanan et al. 2000; Latge 2007; Lindberg 1990; Wang et al. 2016b). The structural complexity determines the physical and chemical properties and leads to the highly heterogeneous dynamics, hydration and intermolecular interactions of carbohydrates in biomaterials.

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 (Rytioja et al. 2014). The conventional methods typically need to hydrolyze the digestible polysaccharides to smaller, soluble segments (Sluiter et al. 2010; Wormald et al. 2002) and grind or ball-mill the indigestible portion into fine particles to enable structural analysis by solution-NMR and mass spectrometry (Bubb 2003; Cheng et al. 2013; Duus et al. 2000; Mansfield et al. 2012). The biochemical treatments could considerably perturb and even restructure the biomolecules. Due to these technical hurdles, our knowledge of the structure and dynamics of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10858-020-00304-2) contains supplementary material, which is available to authorized users.

✉ Xue Kang
kangxuepku@gmail.com

✉ Tuo Wang
tuowang@lsu.edu

[1] Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803, USA

Published online: 03 March 2020

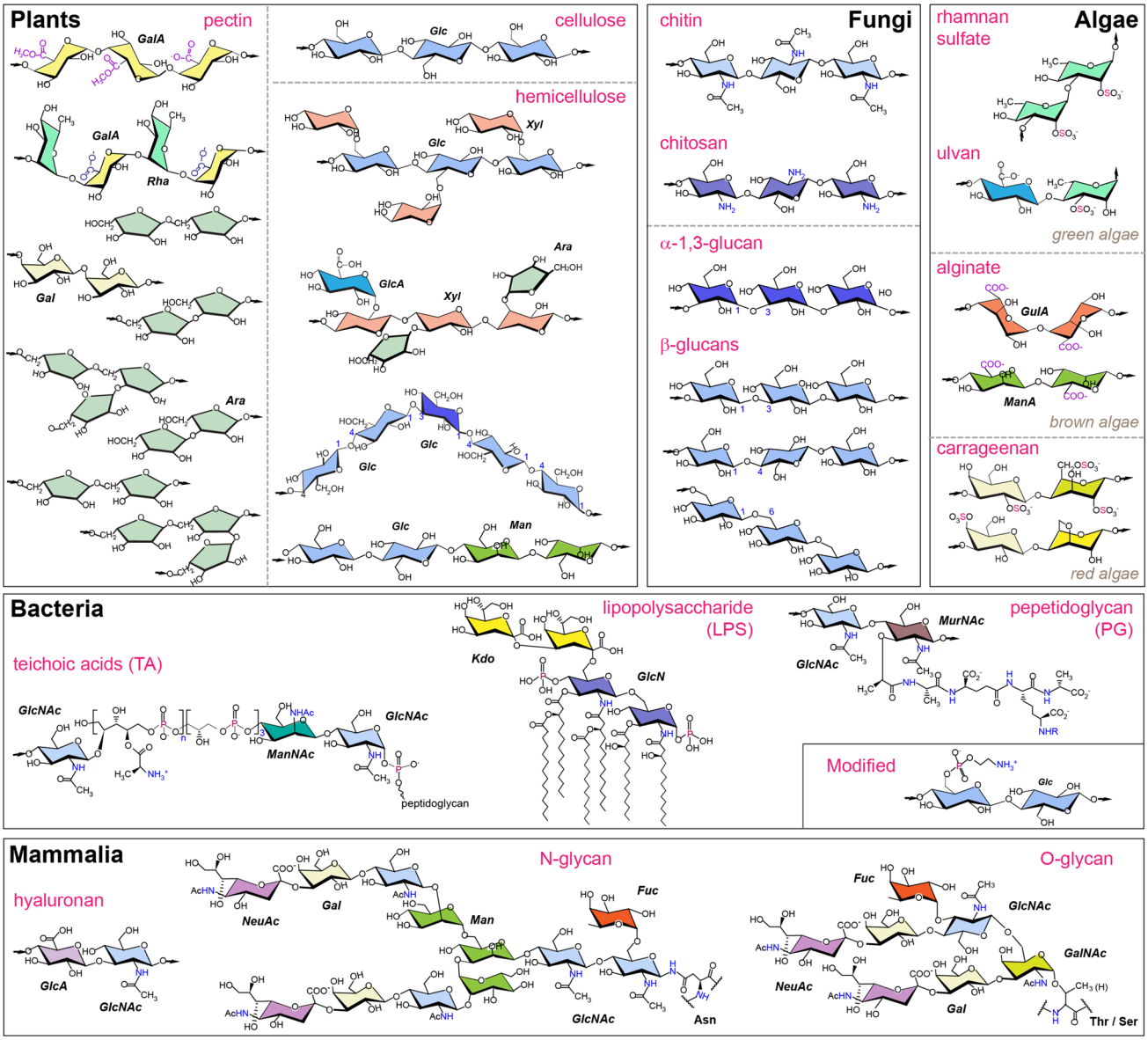

Fig. 1 Representative structure of complex carbohydrates from different origins. Most of these polysaccharides have been characterized by solid-state NMR. The structures of polysaccharides in their cellular environments are often more complicated and irregular

carbohydrates remains inadequate, which has impeded the development of carbohydrate-based biorenewable energy biomaterials (Duus et al. 2000).

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 (Atalla and Vanderhart 1984; Newman et al. 1996). 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 (Arnold et al. 2018, 2015; Bardet et al. 1997; Cadars et al. 2005; Chatterjee et al. 2015; Grantham et al. 2017; Kang et al. 2018, 2019; Kern et al. 2010, 2008; Kim et al. 2015; Laguri et al. 2018; Phyo et al. 2017; Renault et al. 2012; Simmons et al. 2016; Takahashi et al. 2013;

Terrett et al. 2019; Thongsomboon et al. 2018; Wang et al. 2015, 2016c; Werby and Cegelski 2018; White et al. 2014). With the readily improving resolution from the improvement and development of ultrahigh-field magnets (Gan et al. 2017), the sensitivity enhancement from dynamic nuclear polarization (DNP) and solid-state CryoProbe (Rossini et al. 2013; Saliba et al. 2017; Sergeyev et al. 2017; Smith and Long 2016; Takahashi et al. 2012), and the assistance from various NMR methods such as ultrafast spinning and proton detection (Andreas et al. 2015; Phyo and Hong 2019; Struppe et al. 2017) as well as paramagnetic and spectral editing techniques (Jaroniec 2015; Schmidt-Rohr et al. 2012; Wang et al. 2016a), 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.

Methods

Database assembly

The CCMRD system is constructed with a three-tier architecture: server, client and database. The data records are stored in MySql database (Greenspan and Bulger 2001). 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 (Fig. 2a). The organizational schema is shown in Fig. S1.

Server/web interface

The server/web interface was programmed using PHP/Laravel framework (Stauffer 2016). The website is currently hosted on Amazon Web Services. The web server

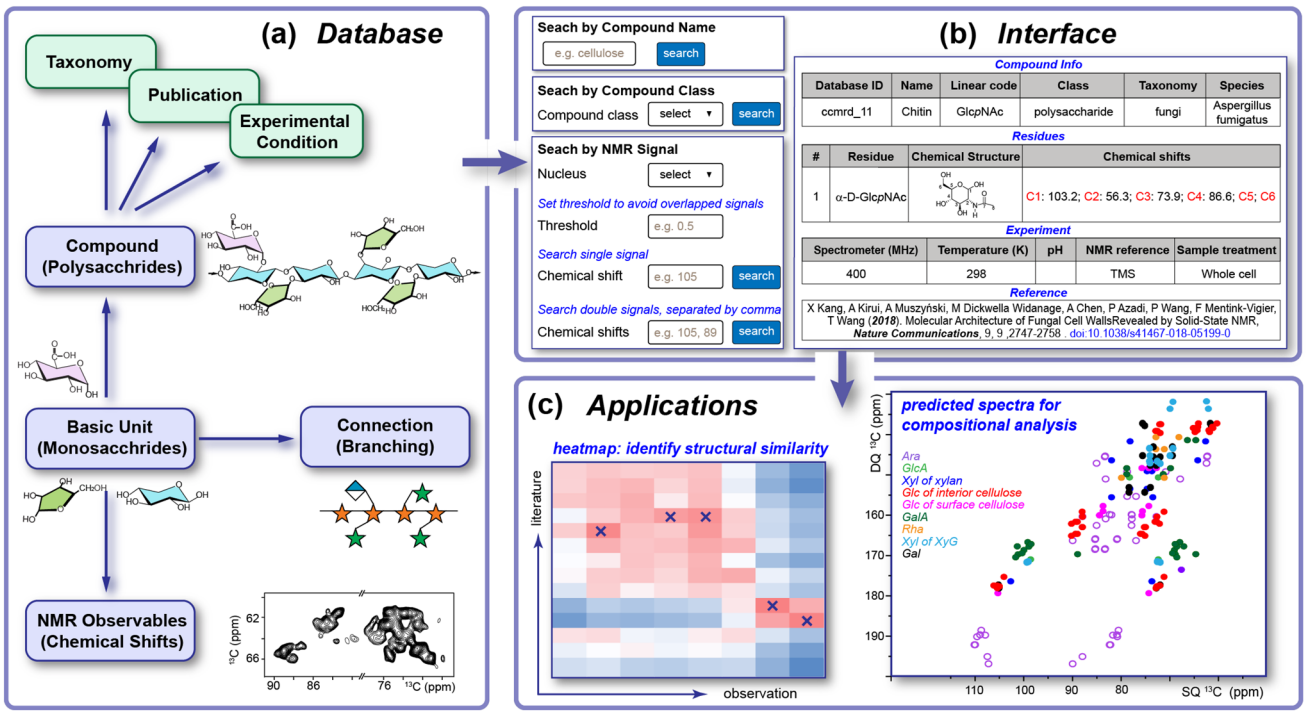

Fig. 2 Overview of the carbohydrate database. **a** Structure of the database. **b** Search interface (left) and a representative entry (right). More detailed instructions for the search and data deposition are in Figs. S2, Fig. 3. **c** For possible applications, a cellulose ^{13}C chemical shift heatmap (left) and a synthesized spectrum of 40 conformers from 9 plant carbohydrates (right) are shown as examples

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.

Results and discussion

To date, CCMRD stores the ssNMR chemical shifts of 435 complex carbohydrates from various organisms (Fig. 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 (Kang et al. 2018; Wang et al. 2014). 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 (Fig. 3b).

Data search

The web interface provides three options for data search (Fig. S2):

(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 (Fig. 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.

(2) Search by polysaccharide names: the users can conduct a search using a trivial name (Fig. S2). 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.

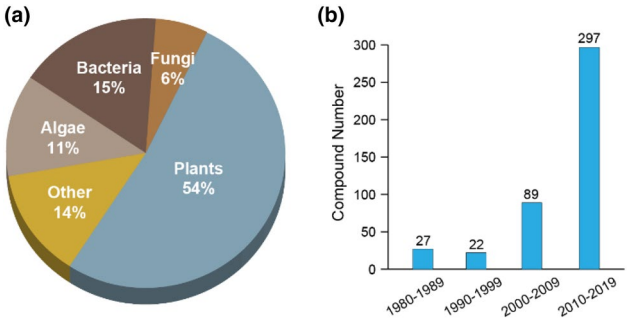

Fig. 3 Statistics of compounds in CCMRD. The data are categorized by **a** taxonomical context, **b** years

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 (Neelamegham et al. 2019; Thieker et al. 2016).

Data deposition

The web server of CCMRD also provides a portal for data submission (Fig. 4 and Fig. S3). 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

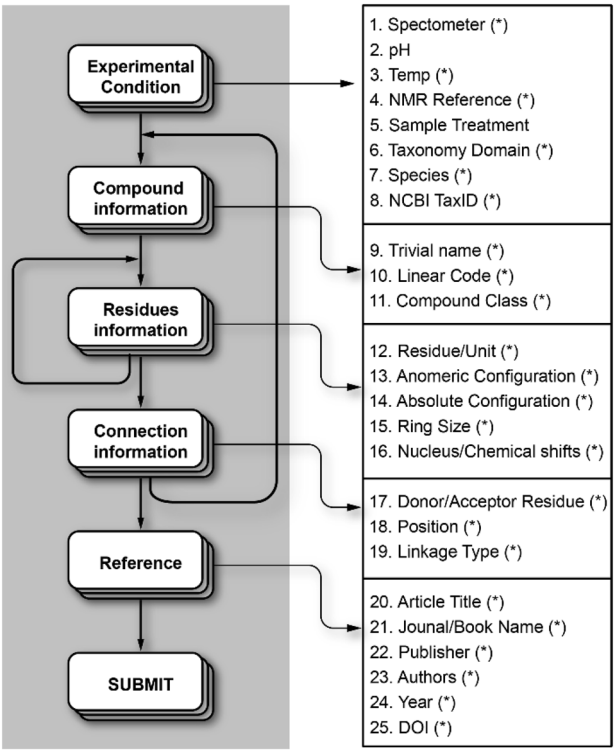

Fig. 4 Flowchart of data deposition. The entries listed here will be needed for depositing data into CCMRD. The step-by-step user guide is shown in Fig. S3

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.
(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 (Fig. 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 (McNaught 1996). 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 the "add new compound" function 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.

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 (Wang et al. 2016b). 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.

Conclusion and outlook

CCMRD presents a timely implementation to the existing databases of carbohydrates, such as Bacterial Carbohydrate Structure DataBase (BCSDB) (Toukach and Egorova 2015) and the GLYCOSCIENCES.de portal (Böhm et al. 2018), 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 (Kirui et al. 2019) and artificial spectra (Fig. 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.

Acknowledgement This work is supported by the National Science Foundation NSF MCB-1942665 and OIA-1833040.

Author contributions XK and TW designed the database and wrote the database; XK, WZ, AK, MCDW and UO indexed the data; XK and WZ deposited the data and programmed the database and interface.

Compliance with ethical standards

Conflict of interest The authors declare no competing financial interest.

References

Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A (2011) Plant cell walls. Garland Science, New York

Andreas LB, Le Marchand T, Jaudzems K, Pintacuda G (2015) High-resolution proton-detected NMR of proteins at very fast MAS. J Magn Reson 253:36–49

Arnold AA, Genard B, Zito F, Tremblay R, Warschawski DE, Marcotte I (2015) Identification of lipid and saccharide constituents of whole microalgal cells by ^{13}C solid-state NMR. Biochim Biophys Acta 1848:369–377

Arnold AA, Bourgouin JP, Genard B, Warschawski DE, Tremblay R, Marcotte I (2018) Whole cell solid-state NMR study of *Chlamydomonas reinhardtii* microalgae. J Biol NMR 70:123–131

Aspinall GO (1983) The polysaccharides, 1st edn. Academic Press, New York

Atalla RH, Vanderhart DL (1984) Native cellulose: a composite of two distinct crystalline forms. Science 223:283–285

Bardet M, Emsley L, Vincendon M (1997) Two-dimensional spin-exchange solid-state NMR studies of ^{13}C-enriched wood. Solid State Nucl Mag 8:25–32

Böhm M, Bohne-Lang A, Frank M, Loss A, Rojas-Macias MA, Lütteke T (2018) Glycosciences. DB: an annotated data collection linking glycomics and proteomics data (2018 update). Nucleic Acids Res 47:D1195–D1201

Bubb WA (2003) NMR spectroscopy in the study of carbohydrates: characterizing the structural complexity. Concept Magn Reson A 19:1–19

Buchanan BB, Gruissem W, Jones RL (2000) Biochemistry and molecular biology of plants. American Society of Plant Physiologists, Rockville

Cadars S, Lesage A, Emsley L (2005) Chemical shift correlations in disordered solids. J Am Chem Soc 127:4466–4476

Chatterjee S, Prados-Rosales R, Itin B, Casadevall A, Stark RE (2015) Solid-state NMR reveals the carbon-based molecular architecture of cryptococcus neoformans fungal eumelanins in the cell wall. J Biol Chem 290:13779–13790

Cheng K, Sorek H, Zimmermann H, Wemmer DE, Pauly M (2013) Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d6/1-ethyl-3-methylimidazolium acetate solvent. Anal Chem 85:3213–3221

Collins MN, Birkinshaw C (2013) Hyaluronic acid based scaffolds for tissue engineering: a review. Carbohydr Polym 92:1262–1279

Duus JØ, Gotfredsen CH, Bock K (2000) Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. Chem Rev 100:4589–4614

Gan ZH et al (2017) NMR spectroscopy up to 35.2 T using a series-connected hybrid magnet. J Magn Reson 284:125–136

Grantham NJ et al (2017) An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. Nat Plants 3:859–865

Greenspan J, Bulger B (2001) MySQL/PHP database applications. Wiley, New York

Himmel ME, Ding S-Y, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science 315:804–807

Jaroniec CP (2015) Structural studies of proteins by paramagnetic solid-state NMR spectroscopy. J Magn Reson 253:50–59

Kang B, Opatz T, Landfester K, Wurm FR (2015) Carbohydrate nanocarriers in biomedical applications: functionalization and construction. Chem Soc Rev 44:8301–8325

Kang X et al (2018) Molecular architecture of fungal cell walls revealed by solid-state NMR. Nat Commun 9:2747

Kang X, Kirui A, Widanage MCD, Mentink-Vigier F, Cosgrove DJ, Wang T (2019) Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. Nat Commun 10:347

Kern T et al (2008) Toward the characterization of peptidoglycan structure and protein-peptidoglycan interactions by solid-state NMR spectroscopy. J Am Chem Soc 130:5618–5619

Kern T et al (2010) Dynamics characterization of fully hydrated bacterial cell walls by solid-state NMR: evidence for cooperative binding of metal ions. J Am Chem Soc 132:10911–10919

Kim SJ, Chang J, Singh M (2015) Peptidoglycan architecture of Gram-positive bacteria by solid-state NMR. BBA-Biomembranes 1848:350–362

Kirui A, Ling Z, Kang X, Dickwella Widanage MC, Mentink-Vigier F, French AD, Wang T (2019) Atomic resolution of cotton cellulose structure enabled by dynamic nuclear polarization solid-state NMR. Cellulose 26:329–339

Laguri C et al (2018) Solid state NMR studies of intact lipopolysaccharide endotoxin. ACS Chem Biol 13:2106–2113

Latge JP (2007) The cell wall: a carbohydrate armour for the fungal cell. Mol Microbiol 66:279–290

Lindberg B (1990) Components of bacterial polysaccharides. Adv Carbohydr Chem Biochem 48:279–318

Lowe JB, Marth JD (2003) A genetic approach to mammalian glycan function. Annu Rev Biochem 72:643–691

Mansfield SD, Kim H, Lu FC, Ralph J (2012) Whole plant cell wall characterization using solution-state 2D NMR. Nat Protoc 7:1579–1589

McNaught AD (1996) Nomenclature of carbohydrates (IUPAC Recommendations 1996). Pure Appl Chem 68:1919–2008

Neelamegham S et al (2019) Updates to the symbol nomenclature for glycans guidelines. Glycobiology 29:620–624

Newman RH, Davies LM, Harris PJ (1996) Solid-state 13C nuclear magnetic resonance characterization of cellulose in the cell walls of *Arabidopsis thaliana* leaves. Plant Physiol 111:475–485

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

Phyo P, Wang T, Xiao C, Anderson CT, Hong M (2017) 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 18:2937–2950

Raman R, Raguram S, Venkataraman G, Paulson JC, Sasisekharan R (2005) Glycomics: an integrated systems approach to structure-function relationships of glycans. Nat Methods 2:817

Renault M, Tommassen-van Boxtel R, Bos MP, Post JA, Tommassen J, Baldus M (2012) Cellular solid-state nuclear magnetic resonance spectroscopy. Proc Natl Acad Sci USA 109:4863–4868

Rossini AJ, Zagdoun A, Lelli M, Lesage A, Coperet C, Emsley L (2013) Dynamic nuclear polarization surface enhanced NMR spectroscopy. Acc Chem Res 46:1942–1951

Rytioja J, Hildén K, Yuzon J, Hatakka A, de Vries RP, Mäkelä MR (2014) Plant-polysaccharide-degrading enzymes from basidiomycetes. Microbiol Mol Biol Rev 78:614–649

Saliba EP et al (2017) Electron decoupling with dynamic nuclear polarization in rotating Solids. J Am Chem Soc 139:6310–6313

Schmidt-Rohr K, Fritzsching KJ, Liao SY, Hong M (2012) Spectral editing of two-dimensional magic-angle-spinning solid-state NMR spectra for protein resonance assignment and structure determination. J Biomol NMR 54:343–353

Sergeyev IV, Itin B, Rogawski R, Day LA, McDermott AE (2017) Efficient assignment and NMR analysis of an intact virus using sequential side-chain correlations and DNP sensitization. Proc Natl Acad Sci USA 114:5171–5176

Simmons TJ et al (2016) Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. Nat Commun 7:13902

Sluiter JB, Ruiz RO, Scarlata CJ, Sluiter AD, Templeton DW (2010) Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. J Agric Food Chem 58:9043–9053

Smith AN, Long JR (2016) Dynamic nuclear polarization as an enabling technology for solid state nuclear magnetic resonance spectroscopy. Anal Chem 88:122–132

Stauffer M (2016) Laravel: up and running: a framework for building modern PHP apps. O'Reilly Media Inc, Newton

Struppe J et al (2017) 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 87:117–125

Takahashi H, Lee D, Dubois L, Bardet M, Hediger S, De Paepe G (2012) 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 Edit 51:11766–11769

Takahashi H, Ayala I, Bardet M, De Paepe G, Simorre JP, Hediger S (2013) Solid-state NMR on bacterial cells: selective cell wall signal enhancement and resolution improvement using dynamic nuclear polarization. J Am Chem Soc 135:5105–5110

Terrett OM et al (2019) Molecular architecture of softwood revealed by solid-state NMR. Nat Commun 10:4978

Thieker DF, Hadden JA, Schulten K, Woods RJ (2016) 3D implementation of the symbol nomenclature for graphical representation of glycans. Glycobiology 26:786–787

Thongsomboon W, Serra DO, Possling A, Hadjineophytou C, Hengge R, Cegelski L (2018) Phosphoethanolamine cellulose: a naturally produced chemically modified cellulose. Science 359:334–338

Toukach PV, Egorova KS (2015) Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts. Nucleic Acids Res 44:D1229–D1236

Wang T, Salazar A, Zabotina OA, Hong M (2014) Structure and dynamics of brachypodium primary cell wall polysaccharides from two-dimensional ^{13}C solid-state nuclear magnetic resonance spectroscopy. Biochemistry 53:2840–2854

Wang T, Park YB, Cosgrove DJ, Hong M (2015) Cellulose-pectin spatial contacts are inherent to never-dried arabidopsis primary cell walls: evidence from solid-state nuclear magnetic resonance. Plant Physiol 168:871–884

Wang T, Chen Y, Tabuchi A, Cosgrove DJ, Hong M (2016a) The target of β-expansin EXPB1 in maize cell walls from binding and solid-state NMR studies. Plant Physiol 172:2107–2119

Wang T, Phyo P, Hong M (2016b) Multidimensional solid-state NMR spectroscopy of plant cell walls. Solid State Nucl Magn Reson 78:56–63

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

Werby SH, Cegelski L (2018) Spectral comparisons of mammalian cells and intact organelles by solid-state NMR. J Struc Biol 206:49–54

White PB, Wang T, Park YB, Cosgrove DJ, Hong M (2014) Water-polysaccharide interactions in the primary cell wall of *Arabidopsis thaliana* from polarization transfer solid-state NMR. J Am Chem Soc 136:10399–10409

Wormald MR, Petrescu AJ, Pao Y-L, Glithero A, Elliott T, Dwek RA (2002) Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. Chem Rev 102:371–386

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supporting Information

CCMRD: A Solid-State NMR Database for Complex Carbohydrates

Xue Kang*, Wancheng Zhao, Malitha C. Dickwella Widanage, Alex Kirui, Uluc Ozdenvar, and Tuo Wang*

Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803

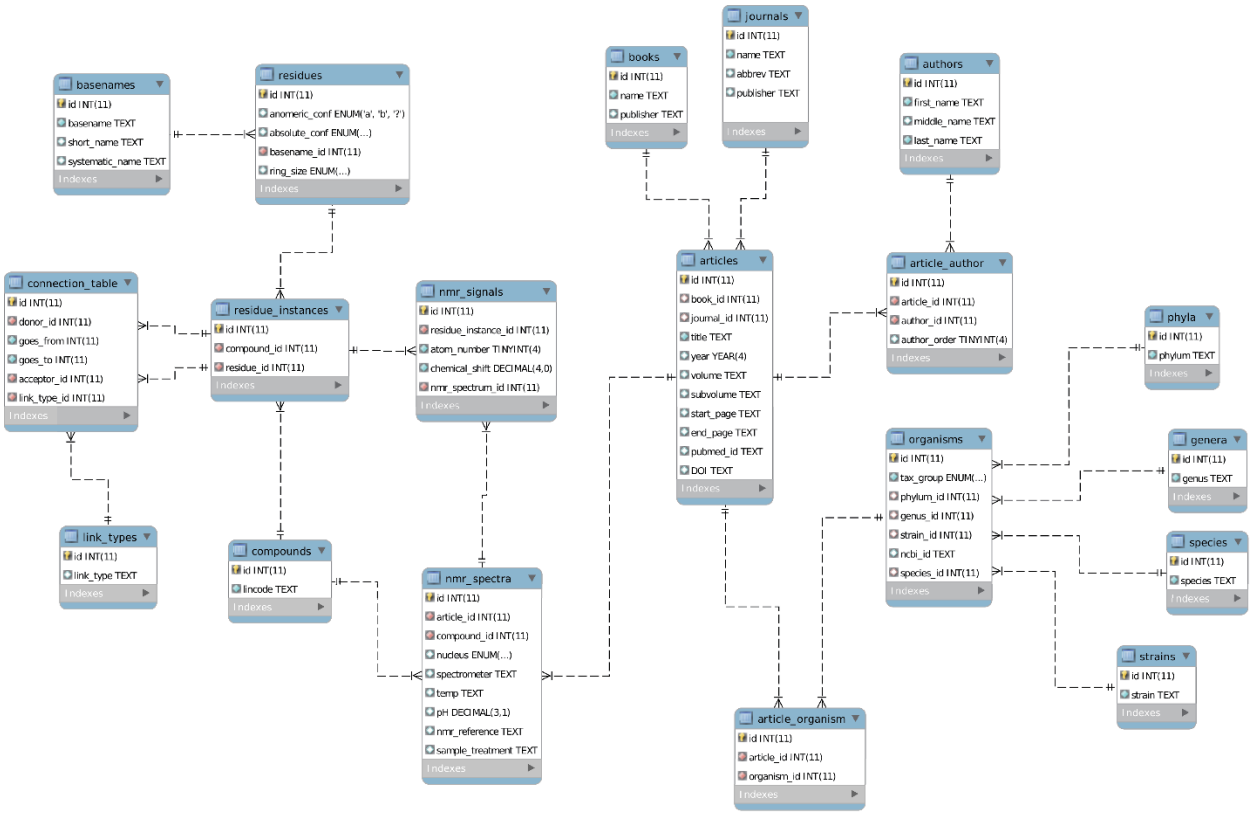

Figure S1. Organizational schema of CCMRD. This figure shows the primary organization and how the different entries (NMR signals, compounds, publications, species, etc.) are connected.

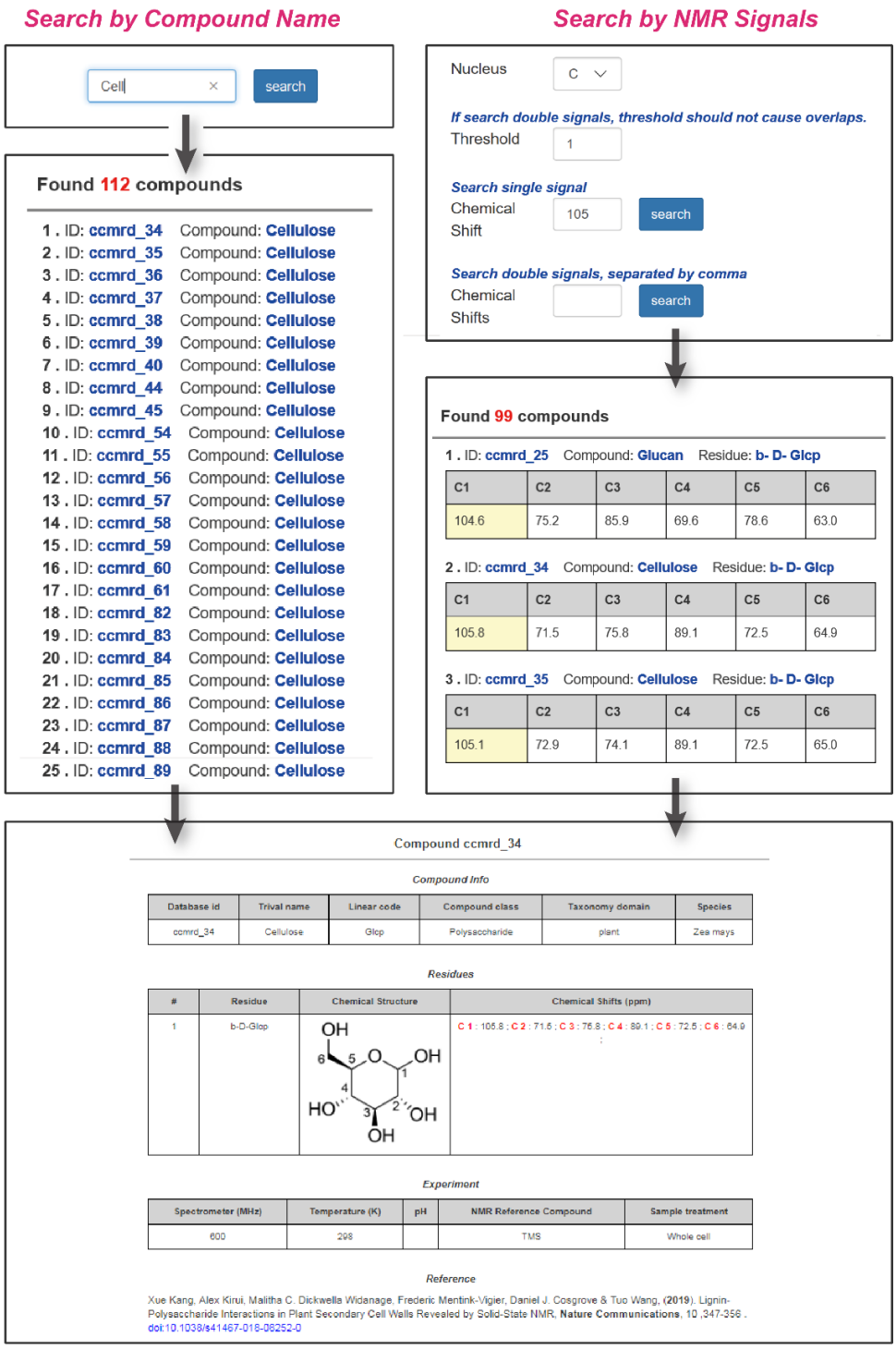

Figure S2. Data Search by compound name and NMR signals. Part or the full name of the polysaccharides is needed for searching by the compound name. In the example, 112 entries have been listed for searching cellulose. Two opinions are available for data searching by the NMR signals: a single NMR chemical 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.

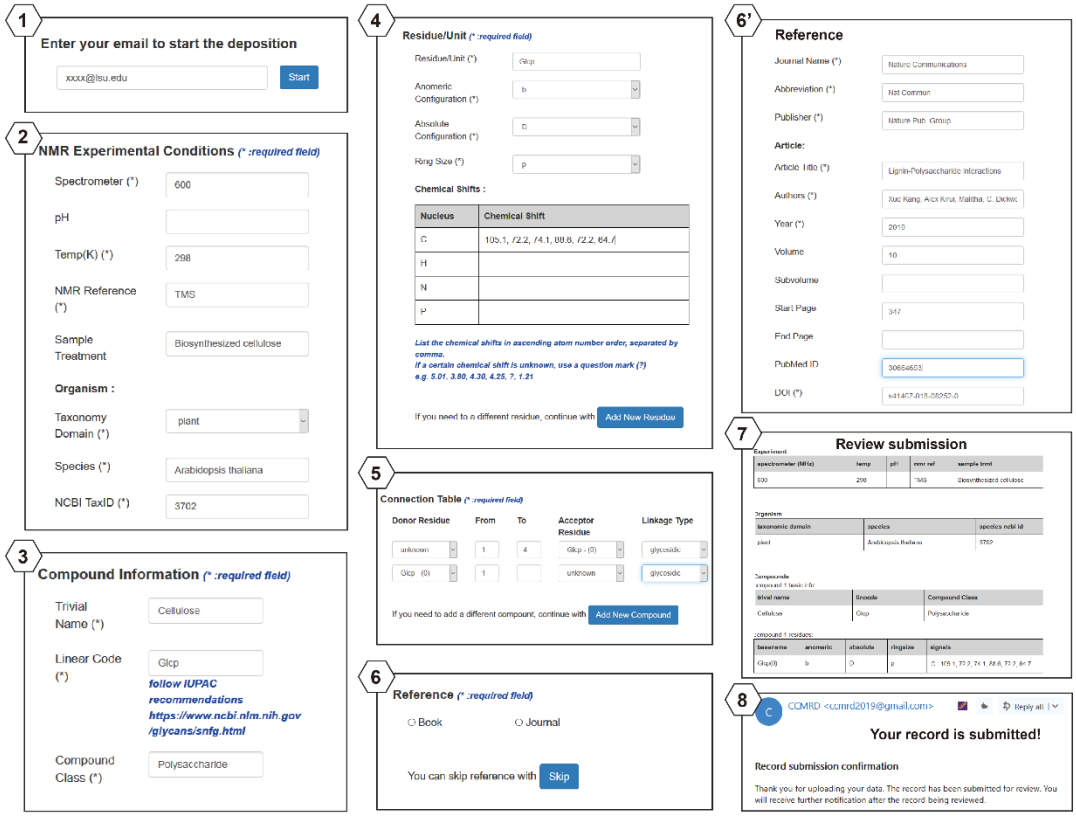

Figure S3. User deposition for CCMRD. The flow chart, interface, and an example of step-by-step deposition are given. The user deposition requires approval by the administrator before the record is indexed by the database.