

PolySac3DB: an annotated data base of 3 dimensional structures of polysaccharides

Sarkar and Pérez

DATABASE

Open Access

PolySac3DB: an annotated data base of 3 dimensional structures of polysaccharides

Anita Sarkar^{1,2} and Serge Pérez^{1,2,3*}

Abstract

Background: Polysaccharides are ubiquitously present in the living world. Their structural versatility makes them important and interesting components in numerous biological and technological processes ranging from structural stabilization to a variety of immunologically important molecular recognition events. The knowledge of polysaccharide three-dimensional (3D) structure is important in studying carbohydrate-mediated host-pathogen interactions, interactions with other bio-macromolecules, drug design and vaccine development as well as material science applications or production of bio-ethanol.

Description: PolySac3DB is an annotated database that contains the 3D structural information of 157 polysaccharide entries that have been collected from an extensive screening of scientific literature. They have been systematically organized using standard names in the field of carbohydrate research into 18 categories representing polysaccharide families. Structure-related information includes the saccharides making up the repeat unit(s) and their glycosidic linkages, the expanded 3D representation of the repeat unit, unit cell dimensions and space group, helix type, diffraction diagram(s) (when applicable), experimental and/or simulation methods used for structure description, link to the abstract of the publication, reference and the atomic coordinate files for visualization and download. The database is accompanied by a user-friendly graphical user interface (GUI). It features interactive displays of polysaccharide structures and customized search options for beginners and experts, respectively. The site also serves as an information portal for polysaccharide structure determination techniques. The web-interface also references external links where other carbohydrate-related resources are available.

Conclusion: PolySac3DB is established to maintain information on the detailed 3D structures of polysaccharides. All the data and features are available via the web-interface utilizing the search engine and can be accessed at <http://polysac3db.cermav.cnrs.fr>.

Keywords: Polysaccharides, Carbohydrates, Three-dimensional (3D) database, Graphical user interface (GUI), Atomic coordinates, Information portal

Background

Carbohydrates are an essential class of biological molecules. They are ubiquitous in the living world, occurring mostly as polysaccharides and oligosaccharides, frequently in the form of conjugates with other biomolecules like proteins (glycoproteins) or lipids (glycolipids). In comparison to the more vastly studied nucleic acids and proteins, carbohydrates have an information

carrying capacity of a much higher degree by virtue of the presence of a multiplicity of chiral centers, combinations of various glycosidic linkages and a large number of functional group modifications that might include acetylation, methylation, oxidation and sulfation, creating an even greater diversity out of the already numerous possible building blocks (monosaccharides) [1]. Polysaccharides (or carbohydrate polymers) are macromolecules made up of repeating monosaccharide units linked by glycosidic bonds. They are essential cellular constituents and their roles extend far beyond being mere energy stores (e.g. starch and glycogen) and structural support agents (e.g. cellulose and chitin). They partake in regulating cell wall plasticity (e.g. pectins,

* Correspondence: spsergeperez@gmail.com

¹Centre de Recherches sur les Macromolécules Végétales (CERMAV[®]) Centre National de la Recherche Scientifique, Grenoble Cedex 9 BP 53X, F-38041, France

²Université Joseph Fourier, Grenoble, France

Full list of author information is available at the end of the article

alginate and carrageenans), cell signaling, governing solution properties of some physiological fluids and participating in the structural build-up of the intercellular matrix (e.g. glycosaminoglycans), eliciting immune responses, cancer progression and as an anti-coagulating agent for the prevention of blood clots (e.g. heparin). Polysaccharides are useful in tissue engineering and repair, wound healing and drug delivery systems, biofuels, biodegradable fibers and bio-composites, due to their generally non-toxic and biodegradable properties and being a renewable resource. Polysaccharides are frequently found on the cell surface of single-celled or multicellular organisms [2] and in the extra-cellular matrix of eukaryotes [3] and are involved in host-pathogen recognition events.

Polysaccharides range in structure from linear to highly-branched. They are often quite heterogeneous, containing slight modifications of their repeating units. This high degree of complexity and inherent micro-heterogeneity of polysaccharide structures make them very difficult macromolecules to handle and explore experimentally. Their structure determines their properties and consequently their function. To understand the molecular basis of the native arrangements of polysaccharides and relating their properties and functions to their structures, the different levels of their structural organization must be determined. As with other macromolecules, the elucidation of the primary structure (implying the sequence of monomeric units with the respective glycosidic linkages) is a pre-requisite. Depending on their primary structures and biosynthesis, polysaccharides may have single or multiple chains in characteristic helical forms that define their secondary structure. Energetically favored interactions between chains of well-defined secondary structures may result in ordered organizations, referred to as tertiary structures. A higher level of organization involving further associations between these well-structured entities results in quaternary structures.

The most important method for the 3D structural characterization of crystalline (or semi-crystalline) polysaccharides is X-ray fiber diffraction. In contrast to other bio-macromolecules, the diffraction data that can be obtained from fibrillar samples of polysaccharides are not of a sufficient quantity and quality to provide enough experimental information to resolve the crystallographic structure unambiguously. A modeling technique must be used which allows the calculation of diffraction intensities from various models for comparison with the observed intensities. In general, the proposed models provide an accuracy of the final atomic coordinates within a few tenths of an Angstrom; typically the R values are around 0.20 in a majority of reported polysaccharide structures. Nuclear magnetic

resonance (NMR) is used in assessing 3D features of polysaccharides in solution which adopt a more or less coiled structure, that fluctuate between local and overall conformations. These features can be translated into 3D structures with the use of molecular modeling. Indeed, molecular modeling has also become an essential component, not only as a complementary technique to be used in the elucidation of 3D crystalline structures, but also as a powerful tool in studying packing of polysaccharides that enables building of models, studying chain-chain interactions [4] and calculating energies. These molecular modeling techniques can be used to construct structures starting from the content of the crystallographic unit cell to much larger macromolecular assemblies offering a unique possibility to visualize morphological features which are in many cases, the relevant level of structural organization with respect to functions or properties of carbohydrate polymers.

3D structures provide information that is indispensable in many aspects of molecular interaction studies. The unification of the resources on carbohydrate polymers, combination of the various tools developed in this field and the free and simple availability of the results that have been generated are necessary for the advancement of polysaccharide science. Bioinformatics has played a role in unifying the resources and information available in genetics and proteomics. Similarly, glycoinformatics has a crucial role to play in the field of carbohydrates. Although a large amount of 3D information regarding the structure of polysaccharides has accumulated over time, the effort to collect, curate and disseminate this data electronically and freely to the scientific community has been feeble. PDB [5] contains very low number of polysaccharide entries, though some coordinates are available, and the Cambridge Structural Database [6] is not a free resource. The only similar contribution, with respect to polysaccharides, has been in the form of a book chapter published in 1997, wherein all the atomic coordinates of polysaccharide structures established by X-ray fiber diffraction have been reported and categorized [7]. A similar effort has been made for celluloses and cellulose derivatives in a book devoted to the structures of this important polysaccharide [8].

Here we report the construction of an annotated polysaccharide 3D structural database called PolySac3DB which provides details of experimental and modeled structures of polysaccharides.

Construction and content

Construction

PolySac3DB is a web-based, platform-independent, manually curated database of polysaccharide 3D structures. It currently runs on an Apache web server [9]

hosted at Centre de Recherches sur les Macromolécules Végétales (CERMAV) with the application program Hypertext Preprocessor (PHP) [10]. It has been developed based on a combination of three layers. The underlying layer is the MySQL database system [11], a relational database management system [MySQL 5.1.41 (Community Server) with PBXT engine 1.0.09-rc] that stores all the structural information along with the respective publications in the back-end and provides the facility to link two or more tables in the database. The intermediate layer is an Apache-PHP application [Apache 2.x; PHP 5.3.1] that receives the query from the user and connects to the database to fetch data from the upper layer, which comprises populated HTML and PHP pages, to the web browser client. The PHP and Java scripts are embedded in the HTML web pages for this effect and are used as application programs for integrating the back-end (MySQL database) to the web pages (HTML). Apache is used as the web server for building the interface between the web browser and the application programs. HTML and PHP have been used to build the web interface.

Content

Data sources –screening, conversion and information extraction

In order to collect structural information about the constituent members of the various polysaccharide families, an extensive screening of literature was performed. This yielded 84 publications that supplied records of the atomic coordinates of polysaccharide (unit) structures established using various structure determination techniques as well as molecular modeling, predominantly containing diffraction data. Enough information could be extracted from these publications to fit the minimum information criteria set for this database and thereafter a total of 157 polysaccharide structures were incorporated into PolySac3DB. The classification of the polysaccharide structures into families is presented in Table 1. The information was manually extracted and curated before incorporation into the repository. The publications provided atomic coordinates within the asymmetric unit of the cell content available as fractional, Cartesian or cylindrical polar coordinates. The available data was converted to either fractional or Cartesian coordinates to generate the atomic coordinate files in standardized representations of PDB (Protein Data Bank) [5] or Mol2 (SYBYL) [12] formats. The files were generated using an in-house PHP script called PDBGenerator, developed for the construction of this database, which can convert fractional and cylindrical/polar coordinates to PDB format. Besides, SYBYL, PyMol, Mercury and Polys were also used to generate helical/expanded forms of the unit cell structures [12,14,15]. The afore-mentioned formats

were chosen to provide a broad readability by various visualization programs as well as to expedite comparisons of glycan with nucleic acid and protein structure as well as computer simulation of their interactions. Application of the symmetry operators of the space groups was done to generate the atomic content of the unit cell and extend them to larger structures. Where symmetry operator information was unavailable, models were generated (wherever possible) to offer a representation of the expanded forms assumed by the polysaccharides. The 3D structures of the repeat units and the packing structures were split into two separate tables on the relational database, respectively. The work flow is described in Figure 1.

The extracted data also included information about carbohydrate composition, glycosidic linkages as well as space group, unit cell dimensions (a , b , c , α , β and γ), the type of helix (that the polysaccharide chains form) which is made available via the 'Expert Mode'. The experimental methods used in structure determination of the respective polysaccharide, the link to the abstract and the reference to the publication cited were also extracted. Particular attention was given to the recording of the available diffraction patterns, which are indeed the original experimental data from which the 3D structures were established. In the present version of PolySac3DB more than 120 diffractograms have been collected; they form a unique collection of information that have been generated over almost half a century of structural research in the area of carbohydrate polymers.

Data storage

Efficiency of data storage and management are the hallmarks of a fully-functional database. At present the database comprises four tables stored within the relational database working in the back-end of PolySac3DB developed using MySQL which provides the facility of linking/relating two or more tables in a database. The important tables within the database are 'strucdata', 'images', 'polysac3dview' and 'polysac3d-dwnld' that incorporate information regarding the experimental or modeled structures and other information extracted from the publications, the diffraction data and the figure legends and the atomic coordinates of the 3D structures for viewing and download, respectively. The tables are linked via a unique key to maintain non-redundancy in PolySac3DB. Subsequent tables can easily be added and logically connected to the existing relational database to accommodate more data about the polysaccharide structures that would be deemed relevant in the future.

Utility and Discussion

PolySac3DB search: Navigation and retrieval

The links to access various utilities and the search engine are provided on the left panel of the website via

Table 1 Polysaccharide Families and their constituent members present in PolySac3DB

No.	Family Name	Polysaccharide Member	Reference
1	Agaroses	Agarose (single)	[16]
		Agarose (double)	[17]
		Agarose Molecular Models	[18]
2	Alginates	Poly- α -L-Guluronic Acid	[19]
		Poly- β -D-Mannuronic Acid	[20]
		Alginate Molecular Models	[21,22]
3	Amyloses & Starch	A Starch	[23,24]
		Starch Nanocrystals	[25]
		Amylopectins	[26]
		B Starch	[27]
		Amylose DMSO	[28]
		Amylose KOH	[29]
		Amylose Triacetate	[30]
		Amylose tri-O-ethyl (TEA3)	[31]
		Amylose V	[32]
		Amylose V propanol complex	[33]
4	Bacterial Polysaccharides	Dextran (high T polymorph)	[34]
		Dextran (low T polymorph)	[35]
		Exo-polysaccharide (<i>Burkholderia cepacia</i>)	[36]
		α (2-8)-linked Sialic Acid Polysaccharide	[37]
		M41 Capsular Polysaccharide (<i>E. coli</i>)	[38]
		O-antigenic polysaccharide (<i>E. coli</i> 1303)	Manuscript in preparation
		O-antigenic polysaccharide (<i>E. coli</i> O5ab)	Manuscript in preparation
		O-antigenic polysaccharide (<i>E. coli</i> O5ac)	Manuscript in preparation
		O-antigenic polysaccharide (<i>E. coli</i> O65)	Manuscript in preparation
		Capsular Polysaccharide (<i>Rhizobium trifolii</i>)	[39]
		Gellan Native K	[40]
		Gellan K	[41]
		Gellan Li	[42]
		RMDP17	[43]
		Welan (Ca)	[44]
Xanthan	[45]		
5	Carrageenans	Iota Carrageenan	[46]
		Iota Carrageenan (Na salt)	[47]
		Kappa Carrageenan	[48]
6	Celluloses	Cellulose I α	[49]
		Cellulose I β	[50]
		Cellulose I triacetate	[51]
		Cellulose II	[52]
		Cellulose II hydrate	[53]
		Cellulose II hydrazine	[54]
		Cellulose II triacetate	[55]
		Cellulose III ₁	[56]
		Cellulose IV ₁	[57]
		Cellulose microfibrils	[58]

Table 1 Polysaccharide Families and their constituent members present in PolySac3DB (Continued)

7	Chitins & Chitosans	Chitin I (Chitin β)	[59]
		Chitin II (Chitin α)	[60]
		Chitosan (anhydrous)	[61]
		Chitosan (high T Polymorph)	[62]
8	Curdians	Curdlan I (Native)	[63]
		Curdlan II	[64]
		Curdlan III	[65]
9	GAGs	Chondroitin (unsulphated)	[66]
		Chondroitin 4-sulphate Ca	[67]
		Chondroitin 4-sulphate K	[68]
		Chondroitin 4-sulphate Na	[69]
		Dermatan 4-sulphate Na (allomorphs I, II, III)	[70]
		Hyaluronate I & II Sodium	[71]
		Hyaluronate III Sodium	[72]
		Hyaluronate I Potassium	[73]
		Hyaluronate II Potassium	[74]
		Hyaluronate III Potassium	[75]
		Hyaluronate Calcium	[76]
		Hyaluronic acid	[77]
		Heparin (dp 12) Heparin (dp 18, 24, 30, 36)	[78,79]
		Keratan-6-sulphate	[80]
10	Galactoglucans	Galactoglucan	[81]
11	Galactomannans	Galactomannan	[82]
12	Glucomannans	Konjac glucomannan	[83]
13	Mannans	Mannan I	[84]
		Mannan II	[85]
		α -D-1,3-Mannan	[86]
		Mannan dihydrate	[87]
14	Pectins	Pectic Acid	[88]
		Calcium Pectate	[89]
		Sodium Pectate	[88]
		Polyuronides Molecular Models	[21,22]
		Arabinan	[90]
		Arabino-Galactan Type I	[90]
		Arabino-Galactan Type II	[90]
		RG-I	[90]
		RG-II	[91]
15	Scleroglucans	Scleroglucan	[92]
16	Xylans	Xylan (β -1,3)	[93]
		Xylan (β -1,4)	[94,95]
17	Nigeran	Nigeran	[96]
18	Others	Inulin hemihydrate	[97]
		Inulin monohydrate	[97]
		α -D-glucan	[98]
		α -1,3-glucan triacetate	[99]

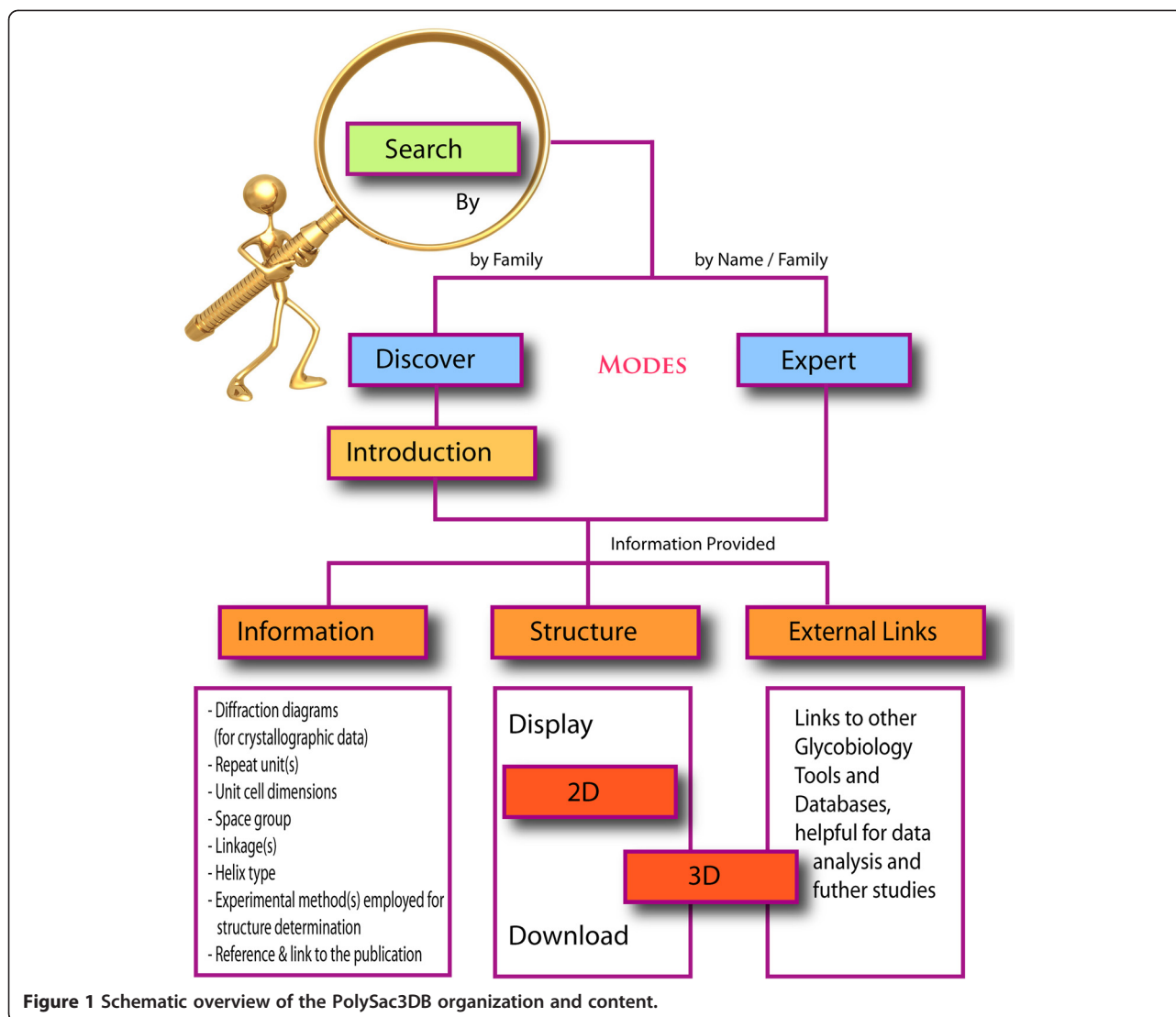


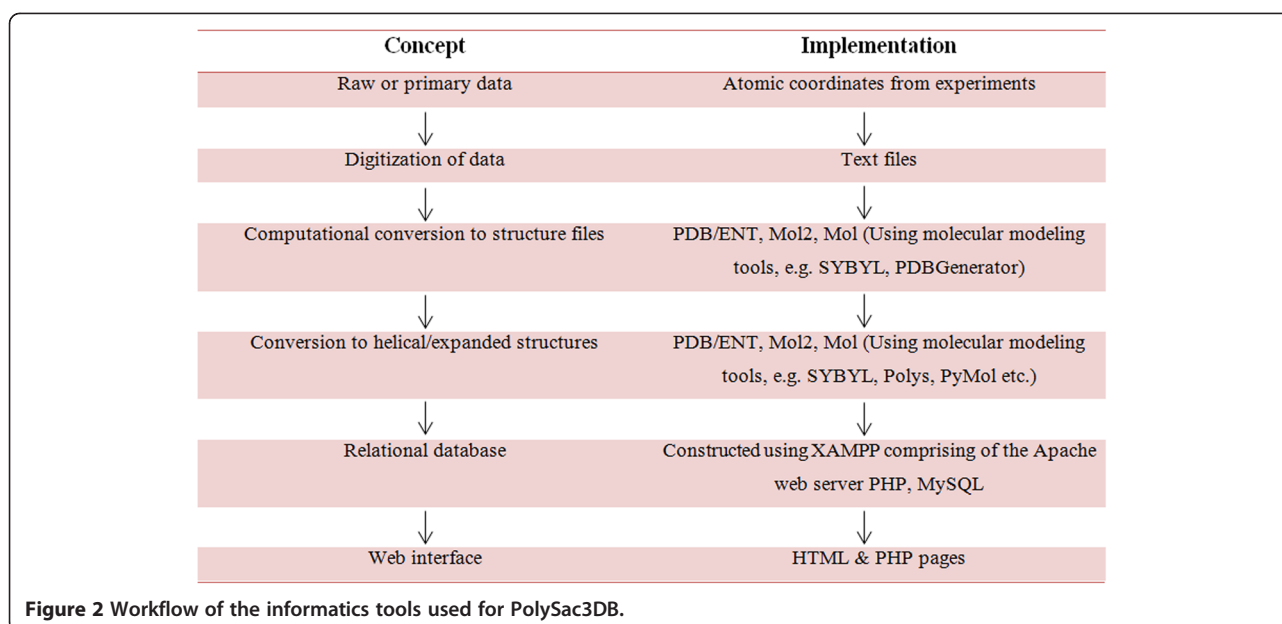
Figure 1 Schematic overview of the PolySac3DB organization and content.

which the data content of the repository can be browsed and retrieved by the user. The 'User Guide' describes each search parameter and its output with detailed examples. A 'Discover Mode' is available that provides background information about the entry/family (mainly regarding occurrence, biosynthesis, property and function). The two-dimensional representations of the polysaccharide repeating unit have been constructed and made available through the 'Discover Mode' in PolySac3DB to aid users to find a familiar representation of the glycan. Information about the nature of the helical structure and all other information can be retrieved upon querying through the 'Expert Mode'.

Data access

Data retrieval and usability are the primary goals set by the developers of an effective database. An interactive

front-end was designed for PolySac3DB with HTML pages and server side scripts that extract data from the tables on the relational database for user-queries on 'Search' and display the retrieved information in a coherent manner. PolySac3DB is equipped with a user-friendly GUI for quick and easy access to the required data. The interface provides the user with options to search by 'Name' or 'Family' of the polysaccharide. This GUI was tested on different versions of four web browser clients (Google Chrome, Mozilla Firefox, Safari and Internet Explorer) with which it performed efficiently. The 'User Guide' gives a detailed description of the content and searchable options within the repository. A schematic over-view is provided in Figure 2. PolySac3DB also provides an overview of the polysaccharide structure determination methods, acting as an information portal on how X-ray, neutron and electron



diffraction as well as molecular modeling are applied to polysaccharides. A list of references is provided on the site on a separate web-page incorporating all the publications from which the atomic coordinates of the structures in the database have been derived, besides proper referencing on the individual 'Expert' pages. In an effort to assimilate other relevant resources for sugars, 'External Links' are provided that empowers the user to explore more online glycoinformatics resources.

PolySac3DB output

The bulk of the structure information for the polysaccharide entries is made available via the 'Expert Mode'. 3D structures can be viewed over the website via the Jmol application [100]. Jmol is an interactive web browser applet that is an open-source, cross-platform 3D Java visualizing tool for visualizing chemical and molecular structures. It provides high-performance 3D rendering with standard available hardware. Downloading the atomic coordinates for further independent use is of course another option provided via the expert mode. The GUI has been designed to retrieve, interpret and display the related information about each entry stored in the back-end on the four tables of the relational database and display it interactively to the user.

Data arrangement followed data collection and fields were set up under which the data would be categorized in the database. Since the majority of experimental structures in our dataset contained entries from crystallography, the data fields were defined upon these guidelines.

Beyond the unit cell contents

Besides providing essential structural information, the 3D crystallographic data on polysaccharides open the way to further insights into other strata of structural organization. The following describes some examples of such extensions. In the case of celluloses, the availability of an accurate description of the crystalline structures of the two allomorphs (cellulose I α [49] and cellulose I β [50]) has provided new insights into the crystalline morphology of the native celluloses. These models were used to generate different ordered atomic surfaces, and evaluate their occurrence along with their respective features. Full atomic models of the crystalline morphology and surfaces of a micro-fibril of cellulose made up of 36 cellulose chains could be conceptualized [58]. Such a model was built as a part of the present database as shown in Figure 3.

In the case of starch, full atomic models of a nano-crystal containing 300 double helical segments in full crystallographic register have been constructed as a part of the work on PolySac3DB. They explain the morphology of these macromolecular assemblies as revealed by transmission electron microscopy [25]. Figure 4 displays the different levels of structural organizations of starch as represented in the various structures present in PolySac3DB.

Cases occur where the quality of the experimental data are far from being sufficient to establish a non-ambiguous model of the 3D arrangement. For example, extensive molecular modeling has provided insights about the way chain-pairing occurs, being mediated by Ca²⁺ interactions in alginates and pectins [22].

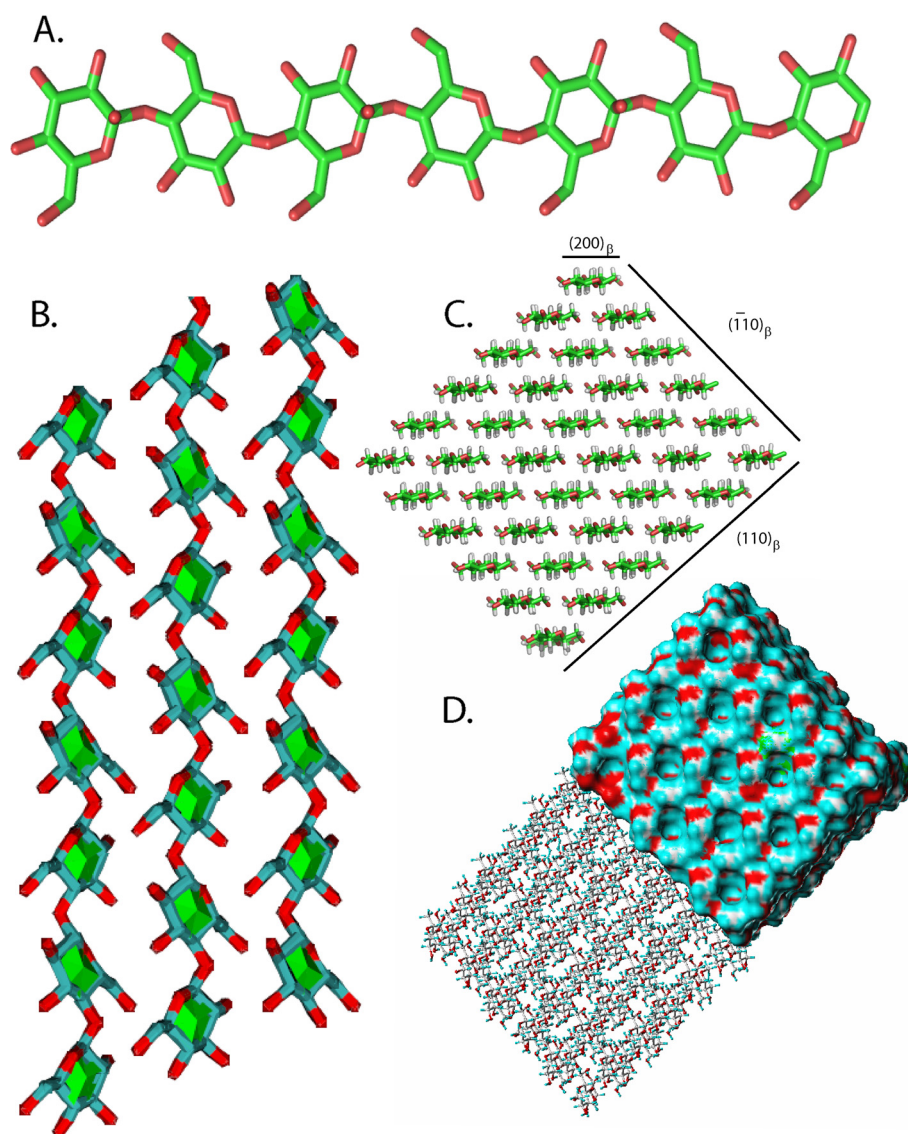


Figure 3 Cellulose chain conformation and morphology. (A) Crystalline conformations of the cellulose chain in the 1β allomorph showing the disordered orientation of hydroxylic hydrogen atoms. (B) Relative orientation of cellulose chains of native cellulose 1β . (C) Molecular model of the microfibril of cellulose projected along the fibril axis along with the indexing of the surfaces. (D) Computer representation of the crystalline morphology and surfaces of the microfibril of cellulose made up of 36 cellulose chains.

Conclusion

The aim of the present work was to provide an organization of all polysaccharide atomic coordinates in one single database serving as a unifying repository and to categorize them in a logical fashion for the user to access the required data using pre-customized searching techniques. The search period covered by the present investigation is about 50 years, during which these structural models have been proposed in carbohydrate research. In view of the crucial role played by molecular modeling techniques, it was important to preserve, organize and distribute the macromolecular models

developed. Their extensions to higher level of structures may expand our knowledge from the molecular to the microscopic level and help scrutinizing the several levels of structural organization of polysaccharides that underline their remarkable functions and properties. At a time when more and more carbohydrates and especially polysaccharides are being called to the fore for their increased use in a plethora of areas as diversified as tissue engineering and repair, wound healing, drug delivery systems, biofuels, bio-degradable fibers and bio-composites due to their generally non-toxic and bio-degradable properties and being a renewable resource,

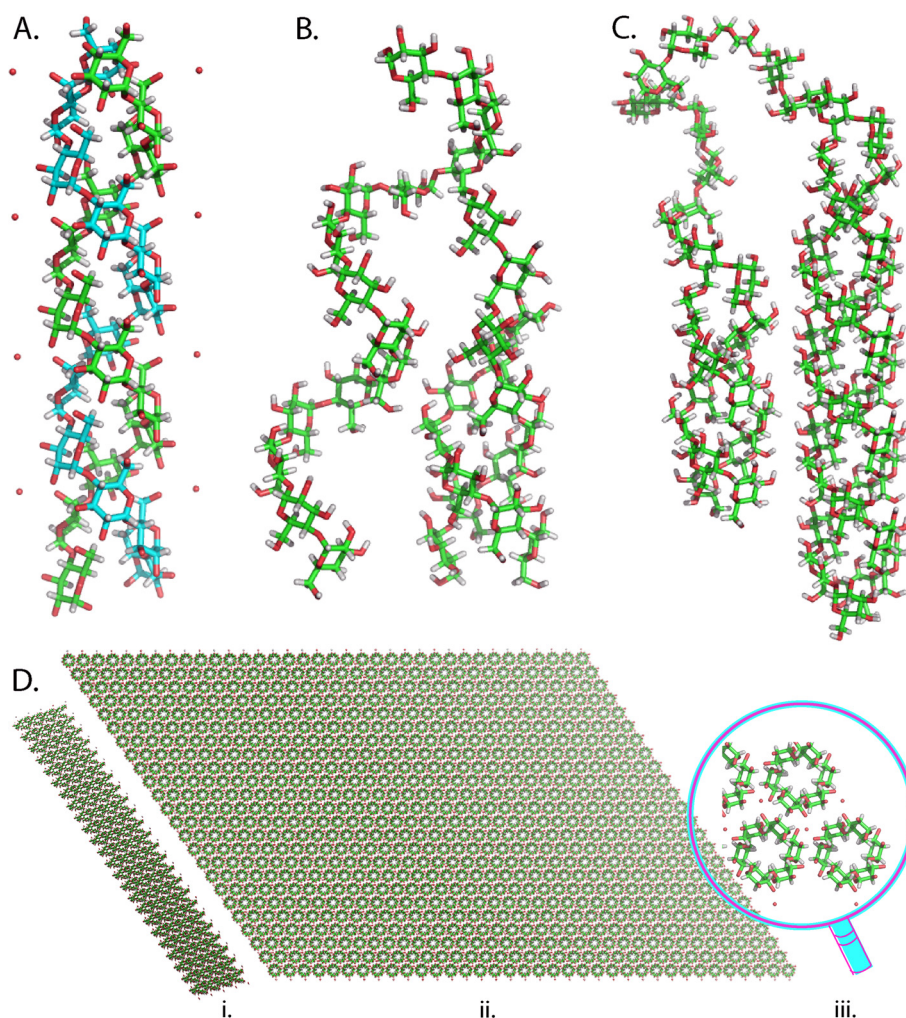


Figure 4 Different levels of structural organization in starch. (A) Representation of the left-handed single chains that are parallel stranded in A-starch double helix. (B) and (C) Representations of the double helix of crystalline starch after modeling the branching point between the strands. (D) Computer representation of an ideal platelet nanocrystal showing (i) width of the platelet with the tilt angle of the double helical component, (ii) composition of the platelet and (iii) the enlarged view of the constituent repeating unit.

PolySac3DB shall be an asset to the community for probing further into the behavior of this class of biological macromolecules.

Availability and requirements

The database PolySac3DB is now available at <http://poly-sac3db.cermav.cnrs.fr> under the glycoscience portal : glyco3d.cermav.cnrs.fr

Abbreviations

3D: Three-dimensional; GUI: Graphical user interface; NMR: Nuclear magnetic resonance; PDB: Protein data bank.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Corresponding author SP designed the framework for the project and wrote the detailed 'Discover' notes. AS designed the method,

developed the MySQL database, web interface and related PHP scripts and wrote the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

The research leading to this publication has received funding from the European Commission's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 215536 and from the Agence Nationale de la Recherche under the "Genomic and Plant Biotechnology" Action throughout the "Wall-Array" project. We are grateful to the Marie Curie Initial Training Network as part of the FP7 People Programme for training and funding. We acknowledge the help of Cyril Bras for testing the webpages for the search engine, Chen Pan for helping to obtain a few extended structures in this database and Alexandre Zakhm for his contribution to PDBGenerator.

Author details

¹Centre de Recherches sur les Macromolécules Végétales (CERMAV*) Centre National de la Recherche Scientifique, Grenoble Cedex 9 BP 53X, F-38041, France. ²Université Joseph Fourier, Grenoble, France. ³European Synchrotron Radiation Facility, (ESRF), Grenoble, France.

Received: 10 May 2012 Accepted: 29 October 2012
Published: 14 November 2012

References

- Gabius H-J: In *Biological Information Transfer Beyond the Genetic Code: The Sugar Code*. In *"The Codes of Life", Biosemiotics*. Edited by Barbiori M, Hoffmeyer J. The Netherlands: Springer; 2008:223–246.
- Gabius HJ: Cell surface glycans: The why and how of their functionality as biochemical signals in lectin-mediated information transfer. *Crit Rev Immunol* 2006, **26**(1):43–79.
- Iozzo RV: Matrix Proteoglycans: From molecular design to cellular function. *Annu Rev Biochem* 1998, **67**(1):609–652.
- Pérez S, Imberty A, Scaringe Raymond P: Modeling of Interactions of Polysaccharide Chains. In *Computer Modeling of Carbohydrate Molecules. ACS Symposium Series 430*. Edited by French AD, Brady JW.: American Chemical Society; 1990:281–299.
- Berman H, Henrick K, Nakamura H: Announcing the worldwide Protein Data Bank. *Nat Struct Biol* 2003, **10**(12):980.
- Allen F: The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr B* 2002, **58**(3 Part 1):380–388.
- Chandrasekaran R: Molecular architecture of polysaccharide helices in oriented fibers. *Adv Carbohydr Chem Biochem* 1997, **52**:311–439.
- Zugenmaier P: *Crystalline Cellulose and Derivatives: Characterization and Structures*. Berlin Heidelberg: Springer; 2008.
- Apache Web Server. <http://www.apache.org/>.
- PHP: *Hypertext Preprocessor*. <http://www.php.net/>.
- Vaswani V: *MySQL: The Complete Reference*. 1st edition. Emeryville, Cal., USA: McGraw-Hill Osborne Media; 2003.
- TRIPOS: SYBYL-X 1.3, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
- DeLano, WL: *The PyMol Molecular Visualization System on an open source foundation and Shrodinger LLC*. www.pymol.org
- Engelsen SB, Cros S, Mackie W, Pérez S: A molecular builder for carbohydrates: Application to polysaccharides and complex carbohydrates. *Biopolymers* 1996, **39**(3):417–433.
- Macrae CF, Bruno IJ, Chisholm JA, Edgington PR, McCabe P, Pidcock E, Rodriguez-Monge L, Taylor R, Van de Streek J, Wood PA: Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J Appl Crystallogr* 2008, **41**(2):466–470.
- Foord SA, Atkins EDY: New X-ray diffraction results from agarose: Extended single helix structures and implications for gelation mechanism. *Biopolymers* 1989, **28**(8):1345–1365.
- Arnott S, Fulmer A, Scott WE, Dea IC, Moorhouse R, Rees DA: The agarose double helix and its function in agarose gel structure. *J Mol Biol* 1974, **90**(2):269–284.
- Kouwijzer M, Pérez S: Molecular modeling of agarose helices, leading to the prediction of crystalline allomorphs. *Biopolymers* 1998, **46**(1):11–29.
- Atkins EDT, Nieduszynski IA, Mackie W, Parker KD, Smolko EE: Structural components of alginic acid. II. The crystalline structure of poly- α -L-guluronic acid. Results of X-ray diffraction and polarized infrared studies. *Biopolymers* 1973, **12**(8):1879–1887.
- Atkins EDT, Nieduszynski IA, Mackie W, Parker KD, Smolko EE: Structural components of alginic acid. I. The crystalline structure of poly- β -D-mannuronic acid. Results of X-ray diffraction and polarized infrared studies. *Biopolymers* 1973, **12**(8):1865–1878.
- Braccini I, Grasso RP, Pérez S: Conformational and configurational features of acidic polysaccharides and their interactions with calcium ions: a molecular modeling investigation. *Carbohydr Res* 1999, **317**(1–4):119–130.
- Braccini I, Pérez S: Molecular basis of Ca^{2+} -induced gelation in alginates and pectins: the egg-box model revisited. *Biomacromolecules* 2001, **2**(4):1089–1096.
- Imberty A, Chanzy H, Pérez S, Buleon A, Tran V: The double-helical nature of the crystalline part of A-starch. *J Mol Biol* 1988, **201**(2):365–378.
- Popov D, Buleon A, Burghammer M, Chanzy H, Montesanti N, Putaux JL, Potocki-Veronese G, Riekel C: Crystal Structure of A-amylose: A Revisit from Synchrotron Microdiffraction Analysis of Single Crystals. *Macromolecules* 2009, **42**(4):1167–1174.
- Pérez S, Bertoft E: The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch-Starke* 2010, **62**(8):389–420.
- O'Sullivan AC, Pérez S: The relationship between internal chain length of amylopectin and crystallinity in starch. *Biopolymers* 1999, **50**(4):381–390.
- Imberty A, Pérez S: A revisit to the three-dimensional structure of B-type starch. *Biopolymers* 1988, **27**(8):1205–1221.
- Winter WT, Sarko A: Crystal and molecular structure of the amylose-DMSO complex. *Biopolymers* 1974, **13**(7):1461–1482.
- Sarko A, Biloski A: Crystal structure of the KOH-amylose complex. *Carbohydr Res* 1980, **79**(1):11–21.
- Sarko A, Marchessault RH: Crystalline structure of amylose triacetate I. Stereochemical approach. *J Am Chem Soc* 1967, **89**(25):6454–6462.
- Bluhm TL, Zugenmaier P: The crystal and molecular structure of tri-*o*-ethylamylose (TEA 3). *Carbohydr Res* 1979, **68**(1):15–21.
- Winter WT, Sarko A: Crystal and molecular structure of V-anhydrous amylose. *Biopolymers* 1974, **13**(7):1447–1460.
- Nishiyama Y, Mazeau K, Morin M, Cardoso MB, Chanzy H, Putaux J-L: Molecular and crystal structure of 7-fold V-amylose complexed with 2-propanol. *Macromolecules* 2010, **43**(20):8628–8636.
- Guizard C, Chanzy H, Sarko A: Molecular and crystal structure of dextrans: a combined electron and X-ray diffraction study. I. The anhydrous, high-temperature polymorph. *Macromolecules* 1984, **17**(1):100–107.
- Guizard C, Chanzy H, Sarko A: The molecular and crystal structure of dextrans: a combined electron and X-ray diffraction study. II. A low temperature, hydrated polymorph. *J Mol Biol* 1985, **183**(3):397–408.
- Strino F, Nahmany A, Rosen J, Kemp GJL, Sá-correia I, Nyholm P-G: Conformation of the exopolysaccharide of *Burkholderia cepacia* predicted with molecular mechanics (MM3) using genetic algorithm search. *Carbohydr Res* 2005, **340**(5):1019–1024.
- Brisson JR, Baumann H, Imberty A, Pérez S, Jennings HJ: Helical epitope of the group B meningococcal alpha(2-8)-linked sialic acid polysaccharide. *Biochemistry* 1992, **31**(21):4996–5004.
- Moorhouse R, Winter WT, Arnott S, Bayer ME: Conformation and molecular organization in fibers of the capsular polysaccharide from *Escherichia coli* M41 mutant. *J Mol Biol* 1977, **109**(3):373–391.
- Lee EJ, Chandrasekaran R: The "pseudo double-helical" structure of the gel-forming capsular polysaccharide from *Rhizobium trifolii*. *Carbohydr Res* 1992, **231**:171–183.
- Chandrasekaran R, Radha A, Thailambal VG: Roles of potassium ions, acetyl and L-glyceryl groups in native gellan double helix: an X-ray study. *Carbohydr Res* 1992, **224**:1–17.
- Chandrasekaran R, Puigjaner LC, Joyce KL, Arnott S: Cation interactions in gellan: An X-ray study of the potassium salt. *Carbohydr Res* 1988, **181**:23–40.
- Chandrasekaran R, Millane RP, Arnott S, Atkins EDT: The crystal structure of gellan. *Carbohydr Res* 1988, **175**(1):1–15.
- Bian W, Chandrasekaran R, Rinaudo M: Molecular structure of the rhamnan-like exocellular polysaccharide RMPD17 from *Sphingomonas paucimobilis*. *Carbohydr Res* 2002, **337**(1):45–56.
- Chandrasekaran R, Radha A, Lee EJ: Structural roles of calcium ions and side chains in welan: an X-ray study. *Carbohydr Res* 1994, **252**:183–207.
- Moorhouse R, Walkinshaw MD, Arnott S: Xanthan Gum - Molecular Conformation and Interactions. In *Extracellular Microbial Polysaccharides*. WASHINGTON, D. C.: American Chemical Society; 1977:90–102.
- Arnott S, Scott WE, Rees DA, McNab CG: Iota-carrageenan: molecular structure and packing of polysaccharide double helices in oriented fibres of divalent cation salts. *J Mol Biol* 1974, **90**(2):253–267.
- Janaswamy S, Chandrasekaran R: Three-dimensional structure of the sodium salt of iota-carrageenan. *Carbohydr Res* 2001, **335**(3):181–194.
- Millane RP, Chandrasekaran R, Arnott S, Dea ICM: The molecular structure of kappa-carrageenan and comparison with iota-carrageenan. *Carbohydr Res* 1988, **182**(1):1–17.
- Nishiyama Y, Sugiyama J, Chanzy H, Langan P: Crystal Structure and Hydrogen Bonding System in Cellulose Ia from Synchrotron X-ray and Neutron Fiber Diffraction. *J Am Chem Soc* 2003, **125**(47):14300–14306.
- Nishiyama Y, Langan P, Chanzy H: Crystal Structure and Hydrogen-Bonding System in Cellulose Ib from Synchrotron X-ray and Neutron Fiber Diffraction. *J Am Chem Soc* 2002, **124**(31):9074–9082.
- Stipanovic AJ, Sarko A: Molecular and crystal structure of cellulose triacetate I: A parallel chain structure. *Polymer* 1978, **19**(1):3–8.
- Langan P, Nishiyama Y, Chanzy H: X-ray structure of mercerized cellulose II at 1 Å resolution. *Biomacromolecules* 2001, **2**(2):410–416.

53. David ML, John B: Structure of cellulose II hydrate. *Biopolymers* 1981, **20**(10):2165–2179.
54. David ML, John B, Litt MH: Structure of a cellulose II-hydrazine complex. *Biopolymers* 1983, **22**(5):1383–1399.
55. Roche E, Chanzy H, Boudeulle M, Marchessault RH, Sundararajan P: Three-dimensional crystalline structure of cellulose triacetate II. *Macromolecules* 1978, **11**(1):86–94.
56. Wada M, Chanzy H, Nishiyama Y, Langan P: Cellulose III₁ Crystal Structure and Hydrogen Bonding by Synchrotron X-ray and Neutron Fiber Diffraction. *Macromolecules* 2004, **37**(23):8548–8555.
57. Gardiner ES, Sarko A: Packing analysis of carbohydrates and polysaccharides .16. The crystal structures of cellulose-IV₁ and cellulose-IV₂. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* 1985, **63**(1):173–180.
58. Pérez S, Samain D: Structure and Engineering of Celluloses. In *Advances in carbohydrate chemistry and biochemistry*. 64th edition. Edited by Derek H: Academic Press; 2010:25–116.
59. Gardner KH, Blackwell J: Refinement of the structure of beta-chitin. *Biopolymers* 1975, **14**(8):1581–1595.
60. Carlstrom D: The Crystal Structure of alpha-chitin (Poly-N-Acetyl-D-Glucosamine). *J Cell Biol* 1957, **3**(5):669–683.
61. Yui T, Imada K, Okuyama K, Obata Y, Suzuki K, Ogawa K: Molecular and Crystal-Structure of the Anhydrous Form of Chitosan. *Macromolecules* 1994, **27**(26):7601–7605.
62. Mazeau K, Winter WT, Chanzy H: Molecular and crystal structure of a high-temperature polymorph of chitosan from electron diffraction data. *Macromolecules* 1994, **27**(26):7606–7612.
63. Okuyama K, Otsubo A, Fukuzawa Y, Ozawa M, Harada T, Kasai N: Single-helical structure of native curdlan and its aggregation state. *J Carbohydr Chem* 1991, **10**(4):645–656.
64. Chuah CT, Sarko A, Deslandes Y, Marchessault RH: Packing analysis of carbohydrates and polysaccharides. Part 14. Triple-helical crystalline structure of curdlan and paramylon hydrates. *Macromolecules* 1983, **16**(8):1375–1382.
65. Deslandes Y, Marchessault RH, Sarko A: Triple-helical Structure of (1,3)-β-D-glucan. *Macromolecules* 1980, **13**(6):1466–1471.
66. Sattelle BM, Shakeri J, Roberts IS, Almond A: A 3D-structural model of unsulfated chondroitin from high-field NMR: 4-sulfation has little effect on backbone conformation. *Carbohydr Res* 2010, **345**(2):291–302.
67. Cael JJ, Winter WT, Arnott S: Calcium chondroitin 4-sulfate: molecular conformation and organization of polysaccharide chains in a proteoglycan. *J Mol Biol* 1978, **125**(1):21–42.
68. Millane RP, Mitra AK, Arnott S: Chondroitin 4-sulfate: Comparison of the structures of the potassium and sodium salts. *J Mol Biol* 1983, **169**(4):903–920.
69. Winter WT, Arnott S, Isaac DH, Atkins ED: Chondroitin 4-sulfate: the structure of a sulfated glycosaminoglycan. *J Mol Biol* 1978, **125**(1):1–19.
70. Mitra AK, Arnott S, Atkins ED, Isaac DH: Dermatan sulfate: molecular conformations and interactions in the condensed state. *J Mol Biol* 1983, **169**(4):873–901.
71. Guss JM, Hukins DW, Smith PJ, Winter WT, Arnott S: Hyaluronic acid: molecular conformations and interactions in two sodium salts. *J Mol Biol* 1975, **95**(3):359–384.
72. Winter WT, Smith PJ, Arnott S: Hyaluronic acid: structure of a fully extended 3-fold helical sodium salt and comparison with the less extended 4-fold helical forms. *J Mol Biol* 1975, **99**(2):219–235.
73. Mitra AK, Raghunathan S, Sheehan JK, Arnott S: Hyaluronic acid: molecular conformations and interactions in the orthorhombic and tetragonal forms containing sinuous chains. *J Mol Biol* 1983, **169**(4):829–859.
74. Mitra AK, Arnott S, Sheehan JK: Hyaluronic acid: molecular conformation and interactions in the tetragonal form of the potassium salt containing extended chains. *J Mol Biol* 1983, **169**(4):813–827.
75. Arnott S, Mitra AK, Raghunathan S: Hyaluronic acid double helix. *J Mol Biol* 1983, **169**(4):861–872.
76. Winter WT, Arnott S: Hyaluronic acid: the role of divalent cations in conformation and packing. *J Mol Biol* 1977, **117**(3):761–784.
77. Haxaire K, Braccini I, Milas M, Rinaudo M, Pérez S: Conformational behavior of hyaluronan in relation to its physical properties as probed by molecular modeling. *Glycobiology* 2000, **10**(6):587–594.
78. Khan S, Gor J, Mulloy B, Perkins SJ: Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J Mol Biol* 2010, **395**(3):504–521.
79. Mulloy B, Forster MJ, Jones C, Davies DB: N.M.R. and molecular-modelling studies of the solution conformation of heparin. *Biochem J* 1993, **293**(Pt 3):849–858.
80. Arnott S, Gus JM, Hukins DW, Dea IC, Rees DA: Conformation of keratan sulphate. *J Mol Biol* 1974, **88**(1):175–184.
81. Chandrasekaran R, Lee EJ, Thailambal VG, Zevenhuizen LPTM: Molecular architecture of a galactoglucan from *Rhizobium meliloti*. *Carbohydr Res* 1994, **261**(2):279–295.
82. Chandrasekaran R, Radha A, Okuyama K: Morphology of galactomannans: an X-ray structure analysis of guaran. *Carbohydr Res* 1998, **306**(1–2):243–255.
83. Yui T, Ogawa K, Sarko A: Molecular and crystal structure of konjac glucomannan in the mannan II polymorphic form. *Carbohydr Res* 1992, **229**(1):41–55.
84. Chanzy H, Pérez S, Miller DP, Paradossi G, Winter WT: An electron diffraction study of the mannan I crystal and molecular structure. *Macromolecules* 1987, **20**(10):2407–2413.
85. Millane RP, Hendrixson TL: Crystal structures of mannan and glucomannans. *Carbohydr Polym* 1994, **25**(4):245–251.
86. Yui T, Ogawa K, Sarko A: Molecular and crystal structure of the regenerated form of (1 → 3)-α-D-mannan. *Carbohydr Res* 1992, **229**(1):57–74.
87. Ogawa K, Miyanishi T, Yui T, Hara C, Kiho T, Ukai S, Sarko A: X-Ray diffraction study on (1 → 3)-α-D-mannan dihydrate. *Carbohydr Res* 1986, **148**(1):115–120.
88. Walkinshaw MD, Arnott S: Conformation and interactions of pectins. I. X-ray diffraction analyses of sodium pectate in neutral and acidified forms. *J Mol Biol* 1981, **153**(4):1055–1073.
89. Walkinshaw MD, Arnott S: Conformations and interactions of pectins. II. Models for junction zones in pectinic acid and calcium pectate gels. *J Mol Biol* 1981, **153**(4):1075–1085.
90. Engelsen SB, Cros S, Mackie W, Pérez S: A molecular builder for carbohydrates: Application to polysaccharides and complex carbohydrates. *Biopolymers* 1996, **39**(3):417–433.
91. Pérez S, Rodríguez-Carvajal MA, Doco T: A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. *Biochimie* 2003, **85**(1–2):109–121.
92. Bluhm TL, Deslandes Y, Marchessault RH, Pérez S, Rinaudo M: Solid-state and solution conformation of scleroglucan. *Carbohydr Res* 1982, **100**(1):117–130.
93. Atkins EDT, Parker KD: The helical structure of a β-D-1,3-xylan. *Journal of Polymer Science Part C: Polymer Symposia* 1969, **28**(1):69–81.
94. Nieduszynski I, Marchessault RH: Structure of β-D-(1,4′)-Xylan Hydrate. *Nature* 1971, **232**(5305):46–47.
95. Nieduszynski IA, Marchessault RH: Structure of β-D-(1,4′)-xylan hydrate. *Biopolymers* 1972, **11**(7):1335–1344.
96. Pérez S, Roux M, Revol JF, Marchessault RH: Dehydration of nigeran crystals: crystal structure and morphological aspects. *J Mol Biol* 1979, **129**(1):113–133.
97. André I, Mazeau K, Tvaroska I, Putaux JL, Winter WT, Taravel FR, Chanzy H: Molecular and crystal structures of inulin from electron diffraction data. *Macromolecules* 1996, **29**(13):4626–4635.
98. Ogawa K, Okamura K, Sarko A: Molecular and crystal structure of the regenerated form of (1 → 3)-α-D-glucan. *Int J Biol Macromol* 1981, **3**(1):31–36.
99. Yui T, Sarko A: Molecular and crystal structure of (1,3)-α-D-glucan triacetate. *Int J Biol Macromol* 1992, **14**(2):87–96.
100. Herraez A: Biomolecules in the computer: Jmol to the rescue. *Biochem Mol Biol Educ* 2006, **34**(4):255–261.

doi:10.1186/1471-2105-13-302
Cite this article as: Sarkar and Pérez: PolySac3DB: an annotated data base of 3 dimensional structures of polysaccharides. *BMC Bioinformatics* 2012 **13**:302.