



# Glycoscience@Synchrotron: Synchrotron radiation applied to structural glycoscience

Serge Pérez<sup>\*1</sup> and Daniele de Sanctis<sup>2</sup>

## Review

Open Access

### Address:

<sup>1</sup>Department of Molecular Pharmacochimistry, CNRS-University Grenoble Alpes, France and <sup>2</sup>ESRF – The European Synchrotron, Grenoble, France

### Email:

Serge Pérez\* - spsergeperez@gmail.com

\* Corresponding author

### Keywords:

antibodies; carbohydrate binding domains; cellulose; glycosaminoglycans; glycolipids; glycosyl hydrolases; glycosyl transferases; kinetic crystallography; lectins; polysaccharides; powder diffraction; small-angle X-ray scattering; starch; synchrotron radiation; transporters; X-ray crystallography

*Beilstein J. Org. Chem.* **2017**, *13*, 1145–1167.

doi:10.3762/bjoc.13.114

Received: 30 January 2017

Accepted: 17 May 2017

Published: 14 June 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

© 2017 Pérez and de Sanctis; licensee Beilstein-Institut.

License and terms: see end of document.

## Abstract

Synchrotron radiation is the most versatile way to explore biological materials in different states: monocrystalline, polycrystalline, solution, colloids and multiscale architectures. Steady improvements in instrumentation have made synchrotrons the most flexible intense X-ray source. The wide range of applications of synchrotron radiation is commensurate with the structural diversity and complexity of the molecules and macromolecules that form the collection of substrates investigated by glycoscience. The present review illustrates how synchrotron-based experiments have contributed to our understanding in the field of structural glycobiology. Structural characterization of protein–carbohydrate interactions of the families of most glycan-interacting proteins (including glycosyl transferases and hydrolases, lectins, antibodies and GAG-binding proteins) are presented. Examples concerned with glycolipids and colloids are also covered as well as some dealing with the structures and multiscale architectures of polysaccharides. Insights into the kinetics of catalytic events observed in the crystalline state are also presented as well as some aspects of structure determination of protein in solution.

## Introduction

Over the last decade, glycoscience has greatly benefited from the development of structural biology and the investigation of macromolecular structure and function relationships. Major contributions also came from considerable advances in high resolution NMR spectrometry and electron microscopy along

with the continuous evolution of synchrotron radiation and free electron laser light sources. Since its discovery, X-ray radiation has been an invaluable tool to investigate the structure of matter. The range of wavelengths, in the region of an angstrom, and energies, extending over electronic shell levels, make them

the perfect probe to study material at the atomic scale. Nevertheless, the low availability and versatility of sources had for a long time represented a limitation on the use of X-rays for scientific applications. A major breakthrough came from the advent of synchrotron science. Over the years, they became an indispensable resource in the exploration of matter, thanks to the continuous spectrum of emitted radiation, the extremely high flux and brightness. Those features allow a wide range of experiments, spanning virtually all branches of sciences and technological applications, particularly those akin to nanoscience. Developments in neutron sources have paralleled those of synchrotron sources. Figure 1 summarizes the differences and the complementarity of the information that can be gathered from analyses performed with the respective sources. The synergistic use of both sources becomes particularly relevant when accurate hydrogen details are necessary.

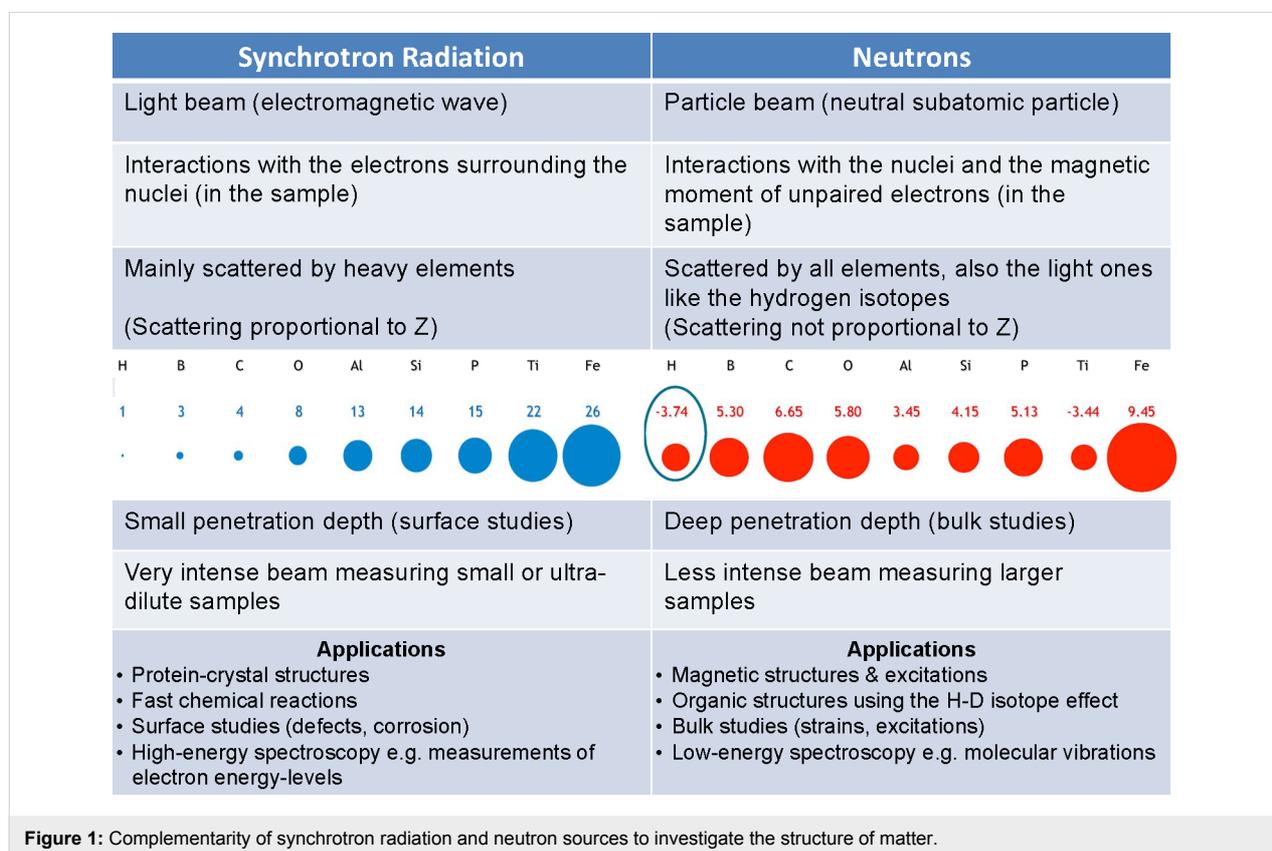
Structural glycobiology gained recognition with the elucidation of glycosyl hydrolases mechanism by X-ray crystallography, but the scope of applications in glycobiology is much broader: it encompasses the range of glycan containing (macro)-molecules and their conjugates. The present article reviews the application of synchrotron radiation to some key areas of glycoscience potentially of interest to the growing number of non-specialist users.

Structural characterization of protein–carbohydrate interactions are covered as well as some involving glycolipids and colloids and the structure and architecture of polysaccharides. Insights into the kinetics of catalytic events occurring in the crystalline state are also described as well as some aspects of the determination of structure of proteins in solution.

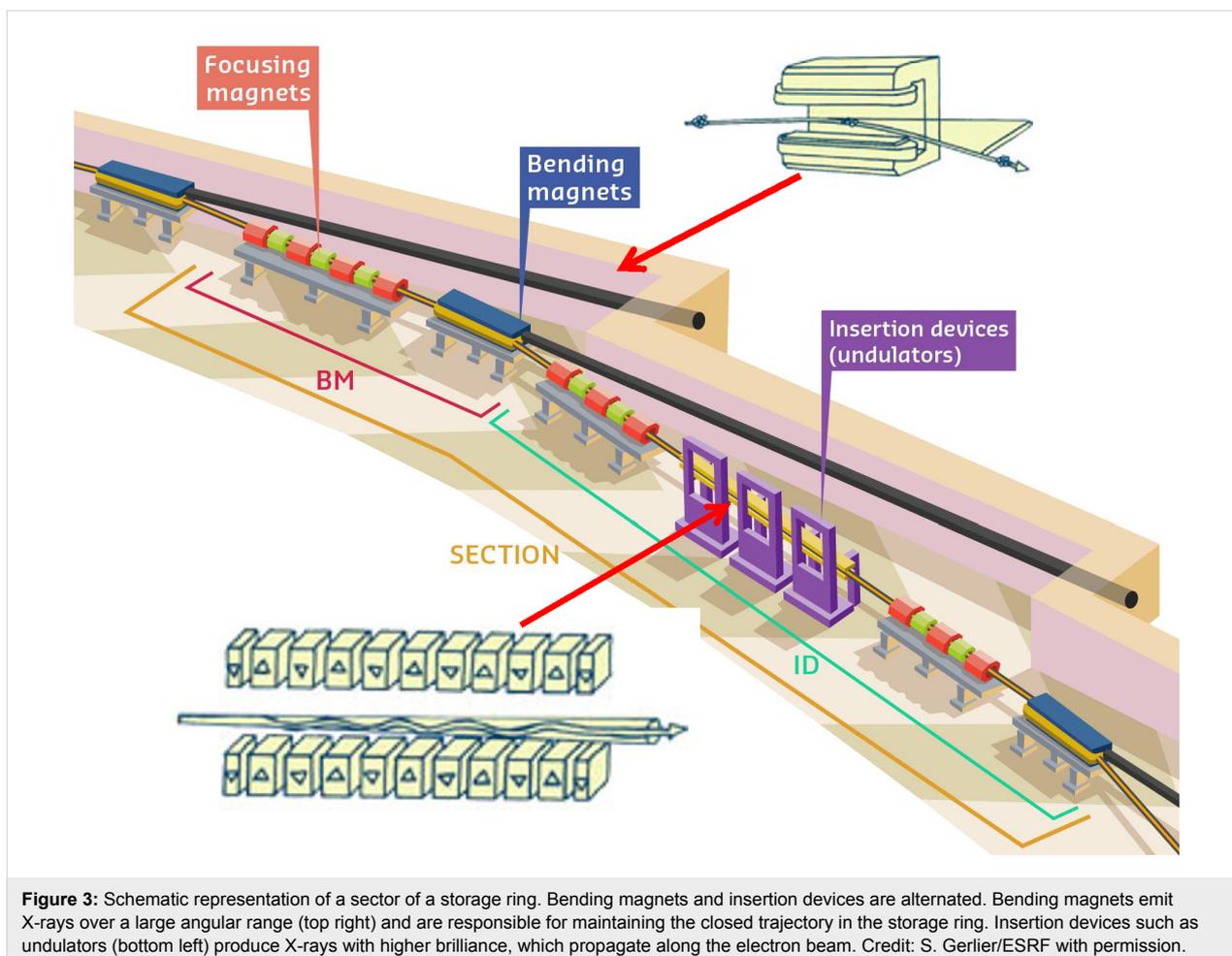
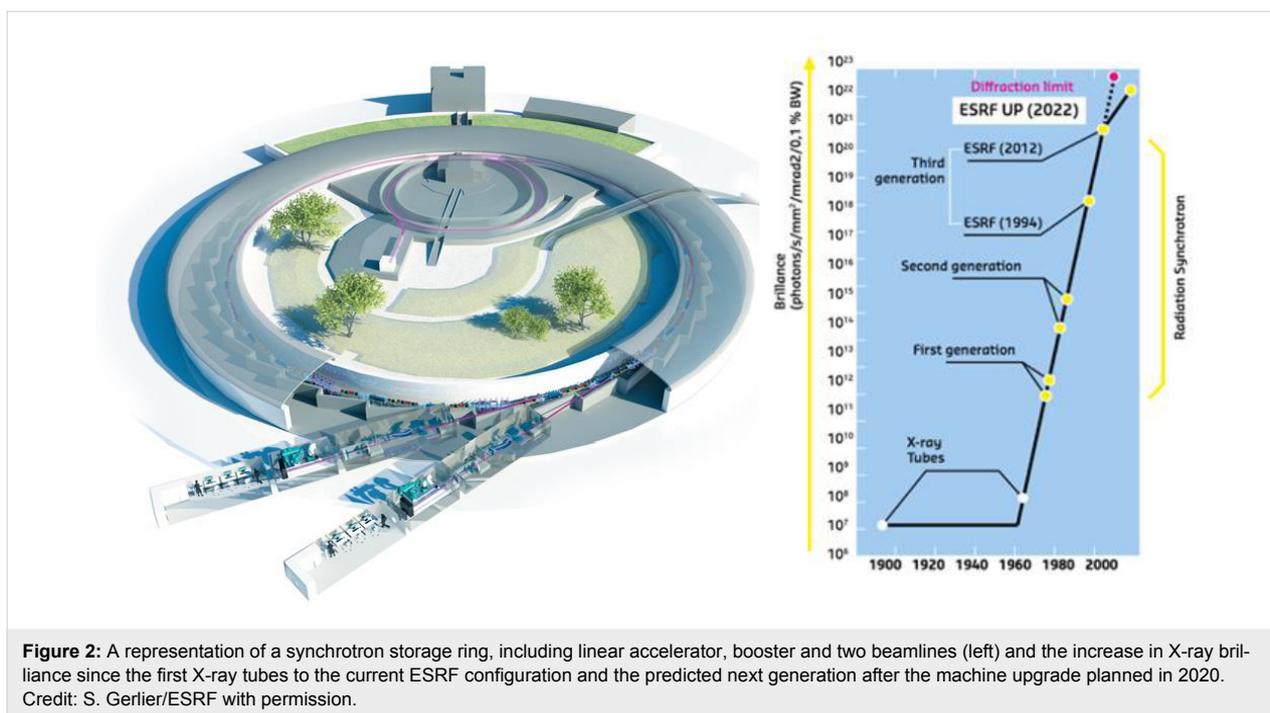
## Review

### Synchrotron radiation

Synchrotrons are particle accelerators in which charged particles circulate along a closed path. Storage rings are a particular kind of synchrotron in which the charged particles, usually electrons, are accelerated to speeds close to  $c$ , the speed of light, and kept orbiting at constant energy (Figure 2). In practice, the terms synchrotron and storage ring are often used interchangeably. The application of magnetic fields induces curvature in the trajectories of the particles, which lose energy by emitting electromagnetic radiation, known as synchrotron light. The electrons are forced to deviate from a straight trajectory either by bending magnets that present a constant dipolar magnetic field and ensure the closing of the orbit, or by insertion devices, such as undulators. Undulators are a much more efficient way to produce X-ray beams and force electrons along an oscillating path in the horizontal plane (Figure 3). In this manner, the X-ray emitted at one oscillation is in phase with the



**Figure 1:** Complementarity of synchrotron radiation and neutron sources to investigate the structure of matter.



radiation from the following oscillations, resulting in an intrinsic higher brilliance. The improvements in insertion devices have made storage rings the most versatile intense X-ray sources, and many storage rings have been constructed around the world and planning for the construction and commissioning of a new generation of storage rings is under way [1].

In the quantum mechanics wave-particle duality, X-rays produced by a synchrotron can be regarded as a linearly polarized electromagnetic plane wave or as photons with energy given by Planck's law. An X-ray photon that interacts with an atom can either be scattered or absorbed. Scattering that occurs with the same momentum (where there is no change in wavelength between scattered and incident waves) is called elastic or Thomson scattering. This is not generally the case, as an incident photon can transfer part of its energy to the electron and is scattered at a lower frequency by a phenomenon known as Compton scattering.

Photoelectric absorption occurs instead when an atom absorbs an X-ray photon. The excess energy is transferred to an electron, which is expelled and the atom is ionized. When the incident photon has an energy above the atomic K shell (so called K-edge energy), it expels an electron from the inner shell and creates a hole, which is eventually filled by an electron decaying from an outer shell. The difference in energy is emitted as a photon of energy equal to the difference of the two atomic shells. This effect is known as X-ray fluorescence and the photon energy provides a unique fingerprint of that atom. Moreover, modulation in absorption around the edge reflects the local structure of the material [2]. Photoelectric absorption, besides depending on the energy, varies with the Z atomic number (approximately proportional to  $Z^4$ ). This phenomenon produces the contrast that is used in X-ray imaging techniques.

The synchrotron light spectrum is polychromatic, with a spectral bandwidth that depends on the type and configuration of the sources. Many experiments require a narrower bandwidth, or a monochromatic beam, and this is produced using a perfect crystal, a monochromator, and the desired wavelength is selected by changing the angle of the incident beam, in accordance with Bragg's law [3]. The monochromatic beam is then focused by using, for example, a system of X-ray mirrors. Consequently, an X-ray beam of the desired size and shape is delivered to the sample position. Continuous development of focusing systems has led to the use of beam sizes as small as a few nanometers. This has allowed the study of smaller samples with an enhanced signal to noise and higher spatial resolution [4-7].

The tunability of the wavelength to reach the values that are optimum for a given experiment provides the most powerful way to determine the three-dimensional features of macromolecular structures. Diffraction performed at an energy close to a heavy element absorption edge produces a resonant effect for which scattered waves are reemitted with a phase delay, inducing small variations in the diffraction intensities. The differences in the intensities can be used to determine the position of the heavier atoms and ultimately the electron density map of the macromolecule's structure. This effect is known as multiwavelength anomalous diffraction (MAD) and today, with its single-wavelength variant (SAD), it is the most successful and widely used techniques to determine the 3D structure of complex systems such as biological molecules, which can be composed of thousands of atoms [8,9].

## Molecular structures

As early as 1930, the first crystal structures of organic compounds to be investigated were carbohydrates of low molecular weight. Over the following years, only eight additional crystal structures were reported. The determination of the three dimensional structure of the dehydrated form of sucrose, in 1947, was considered a significant contribution to the field. A major breakthrough occurred in 1951, when Bijvoet confirmed, without ambiguity, the D-configuration of glucose, which had been assigned from indirect reasoning by Emil Fischer in 1891 [10]. At the present time, the Cambridge Structural Database contains a few thousand entries for carbohydrate crystal structures, among which a limited number of molecules are relevant to glycobiology.

With the exception of sucrose and cyclic compounds, such as cyclodextrins or cyclo-amyloses, carbohydrates are reluctant to crystallize in form and size suitable for X-ray crystallographic analysis. This is even more pronounced for compounds having molecular weights ranging from 1000 to 5000 Da. Among the reasons, there is the difficulty to produce sufficient amount of material or the intrinsic occurrence of molecular disorder in solution, where several forms coexists (linear, five and six membered rings, anomeric mixture, etc.). It is also true that much less effort has been devoted to the production of organic crystals of medium sized biomolecules compared to biological macromolecules. Nevertheless, in many instances, ordered samples may be obtained, either in the form of molecular crystals of micrometric dimensions or in the form of polycrystalline materials.

## Small molecule crystals

In the quest to solve the crystal structures of cello-oligosaccharides, as model compounds of cellulose, several attempts to grow crystals of  $\beta$ -D-cellotetraose of a size suitable for X-ray

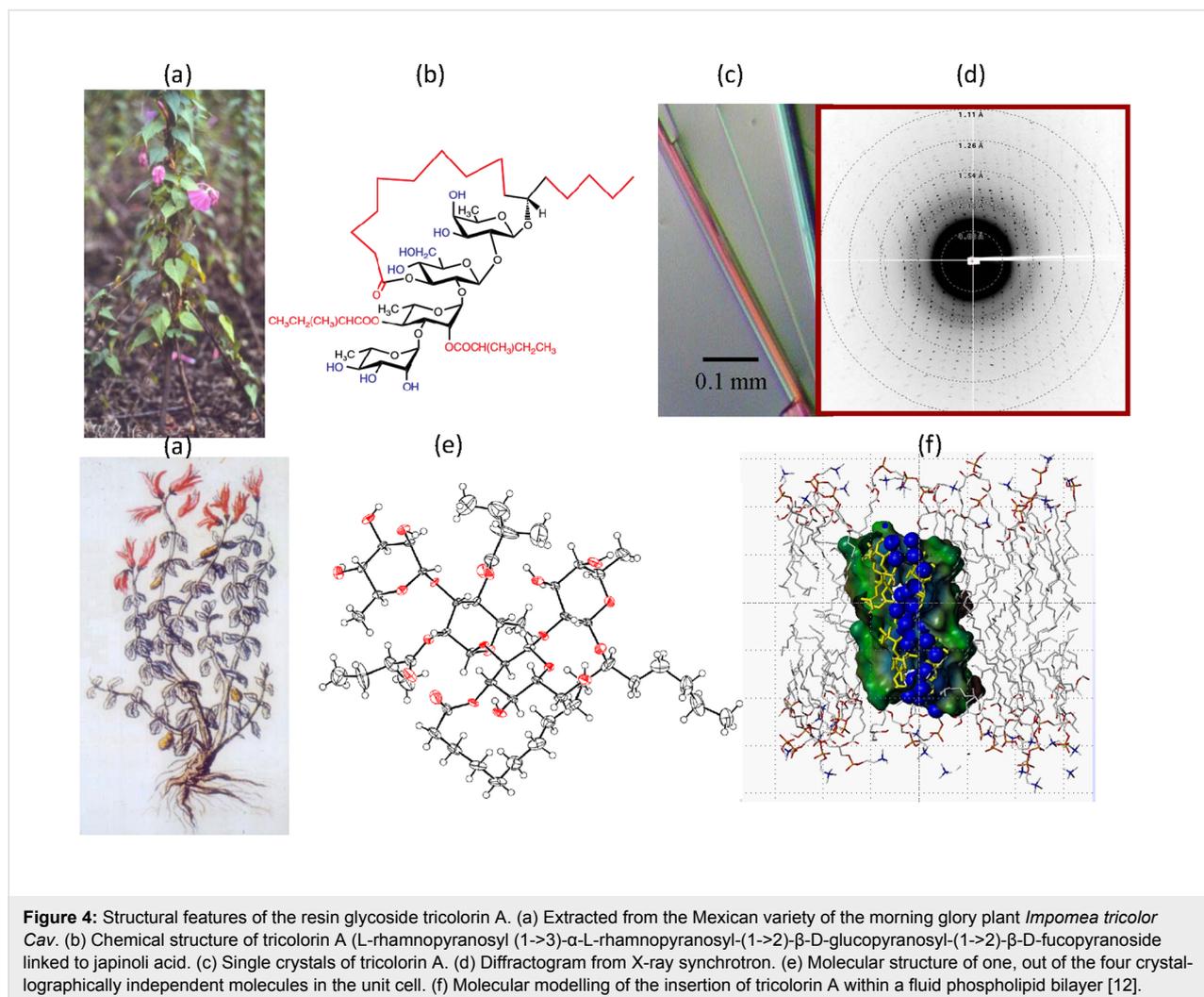
diffraction had failed. Despite many attempts, the best crystals ever obtained for celotriose and celotetraose were very thin laths having dimensions of only about 10  $\mu\text{m}$  in thickness. In the case of celotetraose, single crystals as small as  $0.40 \times 0.15 \times 0.015$  mm could be processed by an X-ray synchrotron beam and 3800 independent reflections were collected. The molecular and crystal structure was solved using molecular replacement methods, and refined to an R factor of 0.048 [11]. Those results were useful in the elucidation of the crystalline structure of cellulose.

The family of resin glycosides offers another example of difficulty in terms of single crystal growth. Glycolipids (or lipopoligosaccharides) comprise a carbohydrate moiety covalently linked to a lipid that confers on them an amphiphilic character, which makes them reluctant to crystallize. One member of the family is tricolorin A (L-rhamnopyranosyl (1- $\rightarrow$ 3)- $\alpha$ -L-rhamnopyranosyl (1- $\rightarrow$ 2)- $\beta$ -D-glucopyranosyl (1- $\rightarrow$ 2)- $\beta$ -D-fucopyranoside linked to japiniol acid forming a 19-membered ring

macrocyclic ester, extracted from *Convolvulaceae* species which have been used in traditional medicine throughout the world since ancient times. Small crystals, with dimensions of  $0.5 \times 0.01 \times 0.01$  mm, could be grown using protein crystallization methods. Data were collected using synchrotron radiation, and the structure was solved using direct methods. Four independent molecules were found in each asymmetric unit (which contains 284 non-hydrogen atoms) in a highly hydrated unit cell (Figure 4) [12].

### Polycrystalline material

Powder diffraction is a standard technique in material science that is used to investigate polycrystalline materials as many micrometer-sized crystals instead of a large single crystal. A powder diffraction pattern captures all possible crystal orientations simultaneously. The development of synchrotron radiation instrumentation dedicated to powder diffraction [13] allows to perform the experiments that were considered to be impractical before.



When using high quality data in conjunction with advanced computational methods, it is possible to solve and refine crystals structures of small organic molecules with limited torsional freedom. This approach is less powerful than single crystal diffraction because of a loss of information by reducing the 3D space on a 1D spectra. Nevertheless, the resolution of the crystalline structure of a synthetic pentasaccharide from heparin, illustrates the potential of this technique. From the experimentally recorded X-ray powder diffractogram (Figure 5), the unit cell dimensions and the space group were determined. The process was continued with a computational building of the pentasaccharide and a simulated annealing procedure in direct space to locate the molecule in the unit cell. Once the carbohydrate backbone was positioned, the refinement continued by an adjustment of the rotations of the glycosidic linkages and side chains. The final construction and model completion provided the crystal and molecular structure with a high confidence factor [14].

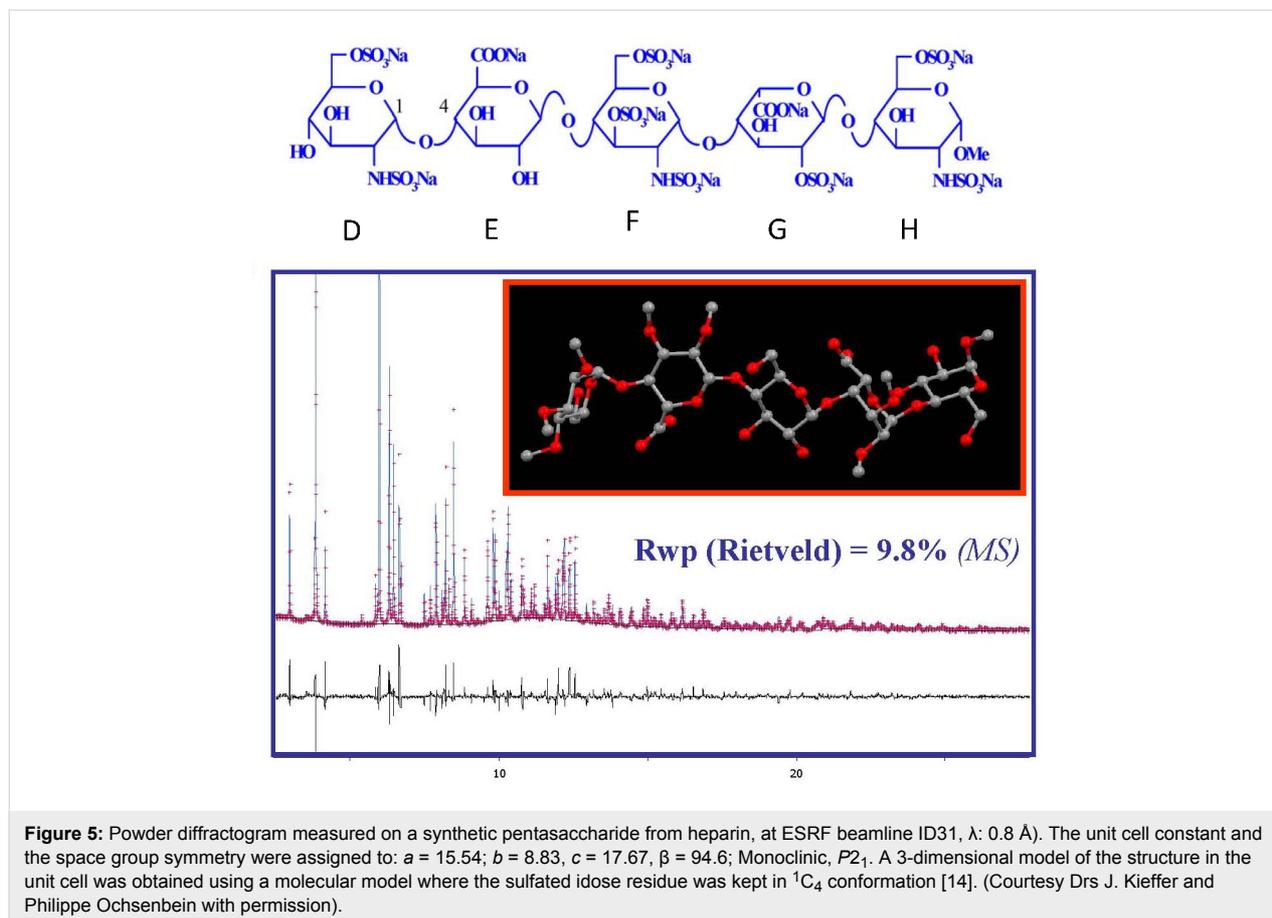
Recently remarkable examples of protein structures determined by using this technique were also reported [15–17]. The clear advantage over single crystal diffraction is the easier preparation of the crystalline samples. As a result of improvement in

the technique's high resolution, new possibilities exist, such as the investigations of the occurrence of phase transitions in large macromolecules as a function of temperature.

### Macromolecular structures

X-ray diffraction with synchrotron radiation is the most powerful method for revealing the three-dimensional structure of biological macromolecules. Among the 128,000 structures deposited in the Protein Data Bank (January 2017) more than 80% have been measured and solved at synchrotron radiation facilities [18].

Macromolecular crystallography beamlines underwent a constant evolution over last decade that had a dramatic impact on the throughput and on the complexity of the structures determined. However, despite the development in nano-volume liquid handling for high-throughput screens, the crystallization of biological macromolecules still represents an important bottleneck in structure determination. Nanoliter handling devices allow the screening of hundreds of crystallization conditions even with a limited amount of sample of a few tens of microliters [19]. Furthermore, a successful example of automation in crystal harvesting were recently reported [20],



while robots are now used to handle cryo-cooled samples at most synchrotron sources. Automation allows for reliable sample exchange and the evaluation of hundreds of samples per day. The development in pixel array detector technology has reduced the time for data collection to minutes or less and significantly improved the quality of the data thanks to no read-out noise and point spread across the pixels. Furthermore the advent of microfocus [21] and microbeam beamlines [22,23] completely dedicated to macromolecular crystallography permits diffraction data collection from smaller samples, of the order of a few micrometers. By matching the X-ray beam to the crystal size, it maximizes the diffraction signal-to-noise and reduces background scattering from crystal holder and mother liquor. New beamline graphical control software [24-26] facilitates beamline operation without exposing the complexity of the hardware. This allows the implementation of elaborate experiments even to users that are less familiar with computational tools. Beamline control software is interfaced with a laboratory information management system (LIMS), a metadata management system [27]. It is used to track samples, record experiment details and report experimental protocols and results from automatic post-experiment data processing protocols [28]. The synergy among these components has recently given rise to completely automated data collection experiments [29].

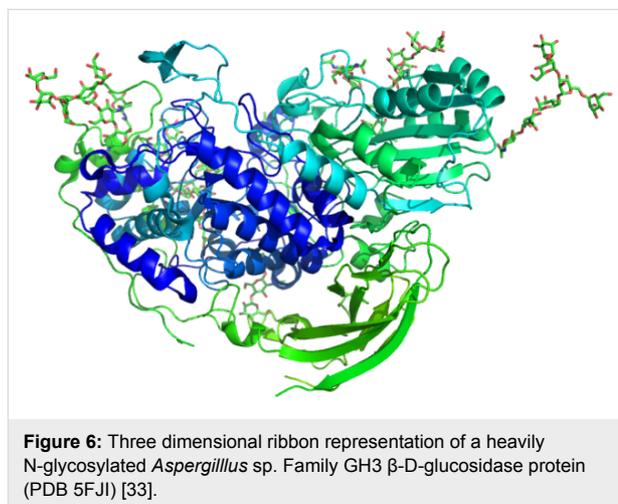
## Glycoproteins

In recent years, the expression and production of recombinant proteins was of great benefit to the whole structural biology community, with more than 85% of the protein structures deposited in the Protein Data Bank being expressed in *Escherichia coli*. However, many proteins require post-translational modifications for correct biological activity and it is estimated that more than 50% of all human proteins are glycosylated, whereas proteins expressed in *E. coli* do not contain any glycan chains. For proteins that require post-translational modification, eukaryotic expression systems are usually preferred [30].

The crystallization of glycoproteins faces several obstacles, including the micro-heterogeneity of glycans at the surface of the protein. For a given glycoprotein, there may exist considerable variations of N-linked glycan chains from protein to protein. Such a heterogeneous macromolecular mix is not suitable for crystal formation. Large post-translational modifications also have the effect of increasing surface entropy and hinder crystal packing. For this reason, it is sometime necessary to manipulate the glycoform to facilitate the crystallization. In the case of the human IgE-FcεRIα [31], Man5-GlcNAc-GlcNAc-Asp-linked glycoforms produced better crystals than in the case where only the Man-GlcNAc-GlcNAc-Asp form was present. There are other cases where crystallization may be facilitated by

the presence of glycans that form stabilizing intermolecular contacts within the crystal. Platforms for the expression and crystallization of glycoproteins are available and can typically be successful in a few weeks [32].

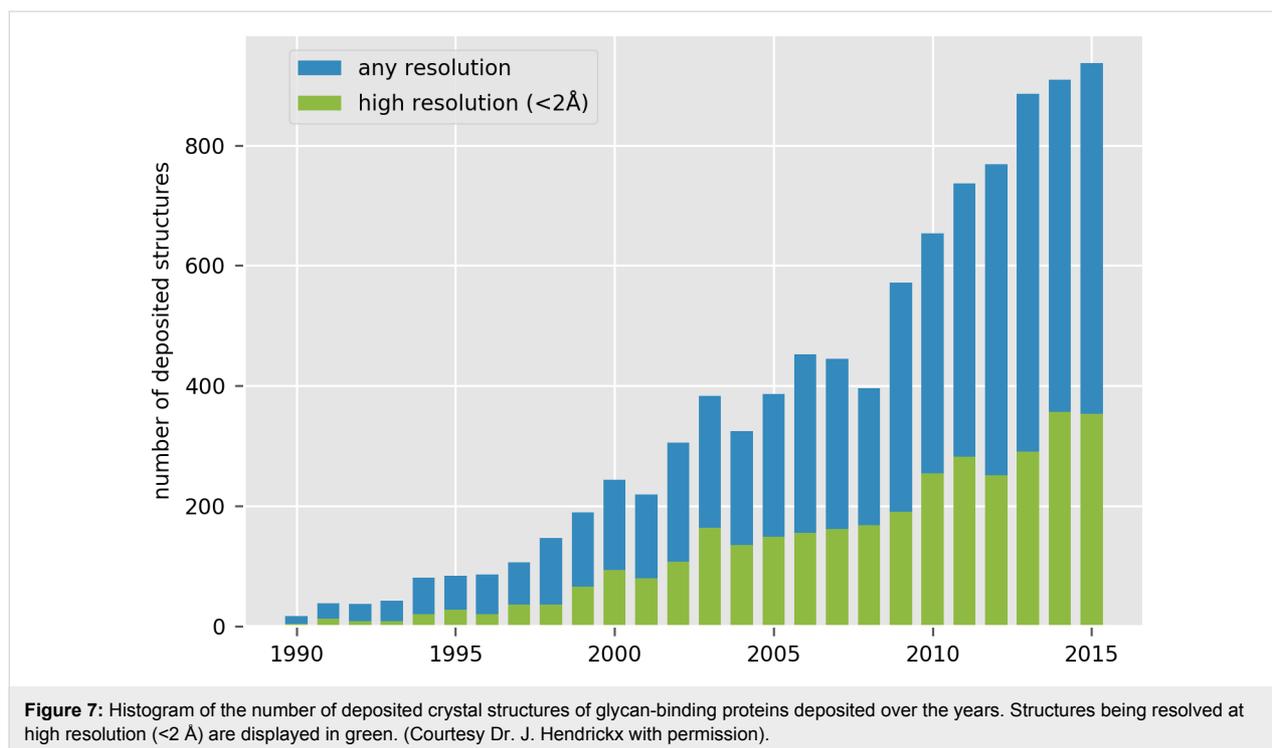
Nevertheless, in the large majority of glycosylated structures, only the electron density map of the initial N-linked GlcNAc is present and can be modelled. In most of the cases, the glycan chains are exposed to the solvent and highly flexible. In such instances, the glycan can be modelled only up to the last visible residue (Figure 6).



The rise in the deposition of glycosylated protein structures reinforced the need for appropriate model restraints for model building and refinement crystallographic software. Model refinement without correct restraints will nearly always result in distortion and particular caution should be given to crystallographic reports where there is a wrong linkage distance specification or a mistaken anomer and handedness. Automated detection, building and validation of sugar models starting from X-ray diffraction data are being implemented [34].

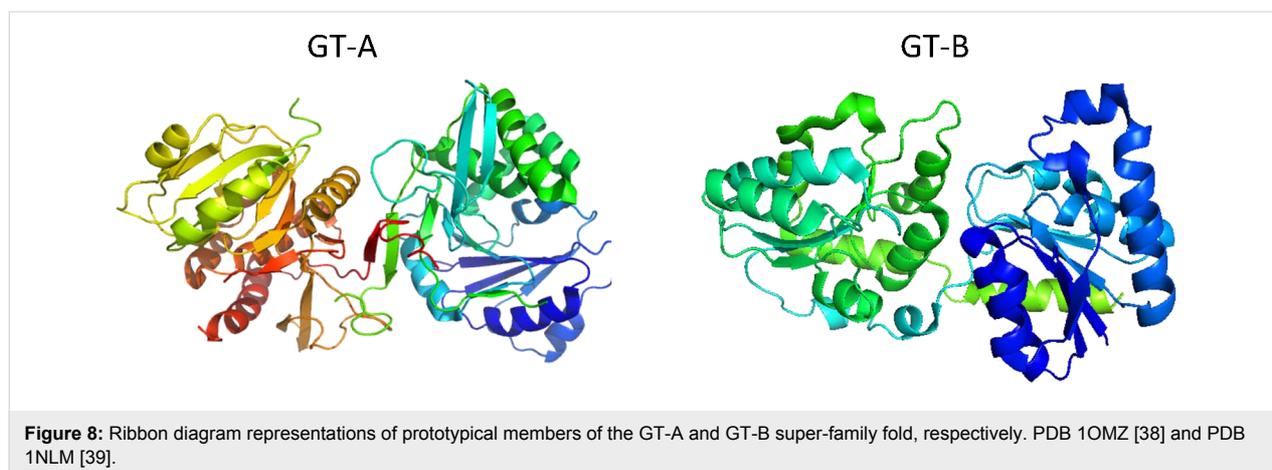
## Carbohydrate interacting proteins

The carbohydrate-mediated recognition events that have a high biological relevance give a pivotal role to the study of protein-carbohydrate interactions. Those interactions drive several distinct biological events, going from the enzymes involved in the biosynthesis, to the hydrolysis and modifications. Transporters and proteins purely involved in recognition (lectin, antibodies, carbohydrate binding modules, glycosaminoglycan binding proteins) are the other important classes of carbohydrate-binding proteins. Figure 7 shows the evolution of the number of carbohydrate interacting proteins that have been solved over the last 25 years, with a particular emphasis on the number of structures determined at high resolution.



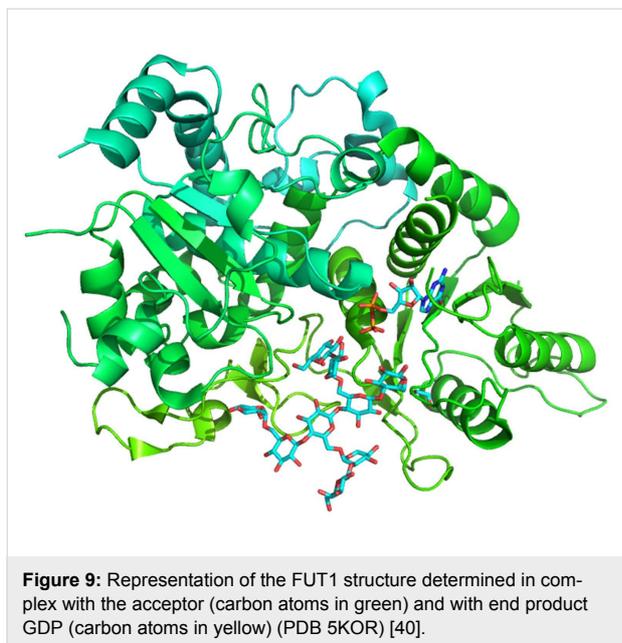
**Glycosyl transferases:** The biosynthesis of oligosaccharides is performed by a ubiquitous class of enzymes: the glycosyl transferases (GTs). The catalytic mechanism underlying the biosynthesis of glycosidic linkage requires the transfer of a sugar residue from a donor to an acceptor [35]. Acceptor substrates are carbohydrates, proteins, lipids, DNA, flavonol, antibiotics and steroids. In contrast, glycosyl donor substrates are mostly sugar nucleotides, such as UDP-GlcNAc, UDP-Gal, GDP-Man, and the GTs that process them are often referred to as Leloir enzymes. In certain cases, lipid-linked sugars, e.g., dolichol phosphate saccharides and unsubstituted phosphates are also utilized. The transfer of saccharides by GTs is regio-specific and stereo-specific: depending on the anomeric configuration of

the transferred saccharide, two possible stereo-chemical outcomes occur, either inversion or retention. Based on the CAZy classification, the number of GT families amounts to 90, in a context where sequence homology is low <http://www.cazy.org> [36]. The increased number of sequenced genomes is paralleled by an increasing number of accession entries for the GTs crystal structures in the PDB, which amounts to 900. Unlike glycoside hydrolases which display a large variety of different folds, the structures of GTs solved today can be clustered in two types of folds (and variants of these folds), namely GT-A and GT-B (Figure 8). Different folds are nevertheless observed for GTs that use lipid phosphate donor substrates. The achievement of the enzyme-transition



state complex requires a particular arrangement of the active site that is the result of concomitant protein dynamics, plasticity of GTs and conformation changes that allow for substrate recognition and catalysis [37].

In plants, GTs are also involved in the biosynthesis of hemicellulose. Xyloglucan is one of the main hemicellulose components in the cell walls of dicots. Its biosynthesis involves different GTs, including a fucosyltransferase, FUT1 that belongs to the glycosyltransferase family 37. The determination of the crystal structure revealed yet another variant of a GT-B fold and could explain FUT1 substrate specificity (Figure 9). Furthermore, the determination in complex with a minimal xyloglucan oligosaccharide acceptor and GDP lead to the understanding of the FUT1 mechanism [40].



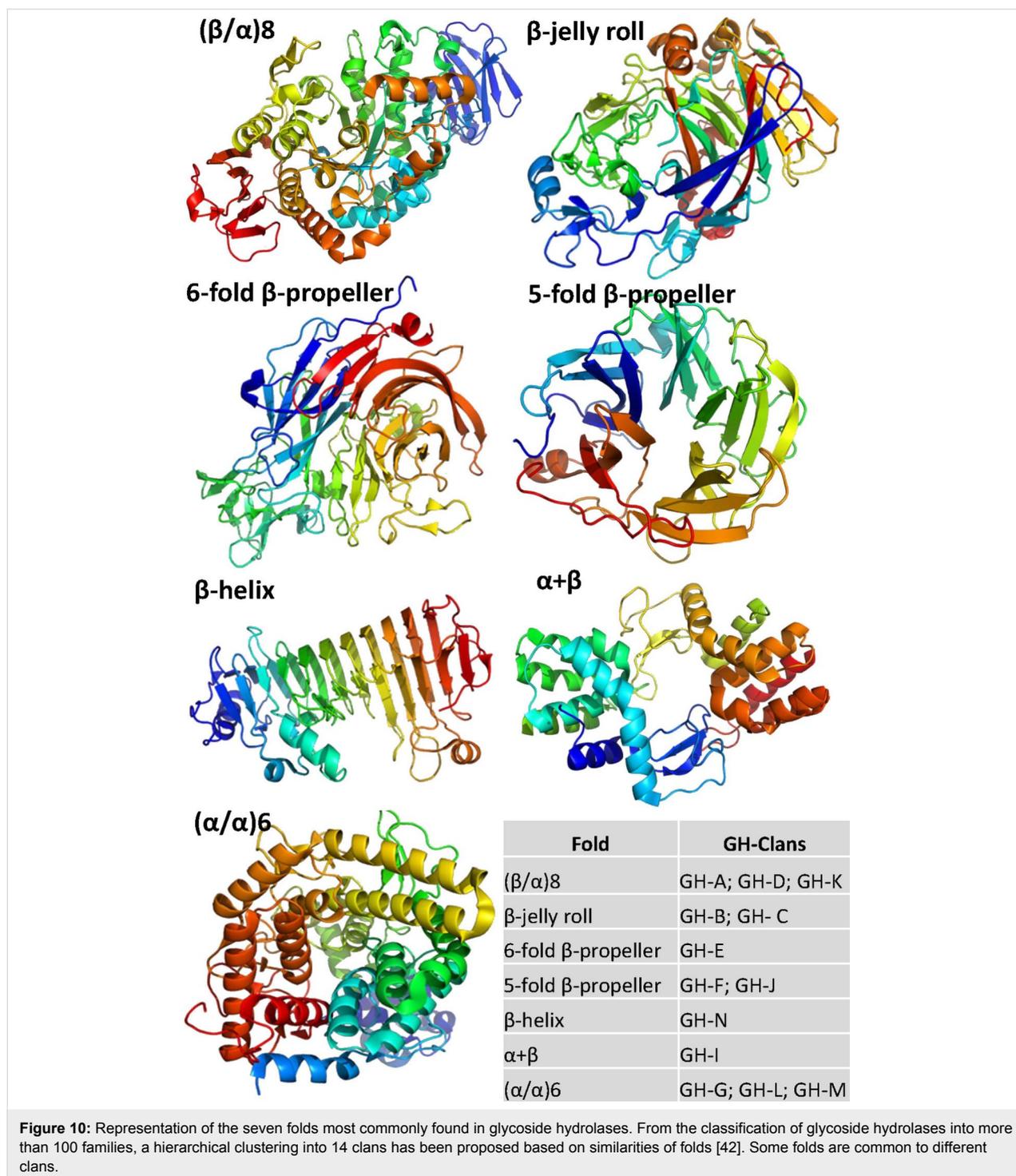
**Carbohydrate esterases:** Carbohydrate esterases perform the de-O or de-N-acylation of carbohydrates. From a mechanistic point of view, this family of enzymes is divided into two classes, according to the dual role played by the carbohydrate. One class is exemplified by the pectin methyl esterase in which case the carbohydrate plays the role of the “acid”. In another class, the carbohydrate acts as an alcohol, as in acetylated xylan. A classification based on amino acid sequence similarities has been proposed yielding 16 families [41]. Among the 100 crystal structures which have been solved, 30 were obtained in complex with carbohydrates, mainly pectic oligosaccharides.

**Polysaccharide lyases:** Polysaccharide lyases (PLs) constitute a family of enzymes that cleave uronic acid-containing polysac-

charide chains. The underlying mechanism is a  $\beta$ -elimination mechanism which generates an unsaturated hexenuronic acid residue and a new reducing end of the polysaccharide. At the present time, the reported number of crystal structures amounts to 190, among which 64 are complexed with carbohydrate ligands. These enzymes show a large variety of folds. Based on amino acid sequence similarities, polysaccharide lyases have been classified in 24 families [41].

**Glycoside hydrolases:** The hydrolysis of carbohydrates is the result of the action of a wide spread group of enzymes: the glycosyl hydrolases (GHs). GHs cleave the glycosidic linkage between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. They can catalyse the hydrolysis of O-, N-, S-linked glycosides, as well. The catalytic event can occur either in the middle (-endo) or at the end (-exo) of the substrate. The hydrolysis of the glycosidic bond implies a general acid (proton donor) and nucleophile/base, and involves two amino acid residues of the enzyme. Depending upon the position of these catalytic residues with respect to the substrate cleavable bond, the outcome of the reaction is either an inversion (inverting mechanism) or a retention (retaining mechanism) of the anomeric configuration. At the present time, about 4500 crystal structures of GHs have been deposited in the PDB. Approximately 30% of them are complexed with carbohydrate ligands. A classification of GH (more than 100 families) has been established, first based on amino acid sequences similarities and further consolidated by the availability of 3D dimensional structures [42]. The analysis of the GH structures present in the CAZy database helped not only to decipher the hydrolytic mechanism, but also reveal the evolutionary relationships between these enzymes. An extended classification based on the fold of the proteins, allowed the identification of 14 main clans (Figure 10).

**Carbohydrate binding modules:** Carbohydrate binding modules (CBM) are defined as a sequence of contiguous amino acids within a carbohydrate-active enzyme with a discrete fold having carbohydrate-binding activity. Initially CBMs were classified as a cellulose-binding domain, but their occurrence in other carbohydrate active enzymes required a dedicated classification, separate from other non-catalytic proteins, and similar to lectins, antibodies and sugar-transport molecules. Depositions in the PDB for CBMs amount to 900. In the CAZy database, CBMs are classified within 80 families based on amino acid sequence similarities, while a three-dimensional structural classification clusters CBM into seven fold families [43]. The most represented fold is the  $\beta$ -sandwich comprised of two  $\beta$ -sheets, each consisting of three to six anti-parallel  $\beta$ -strands. As a large proportion of crystal structures are complexed with carbohydrates (from monosaccharides to oligosaccharides),



three CBM types have been classified based on their sugar recognition modes: surface binders, "endo-type" binders and "exo-type" binders [44].

**Lectins:** Lectins constitute a unique and diverse family of proteins that reversibly bind monosaccharides and oligosaccharides, with utmost specificity, without displaying any catalytic

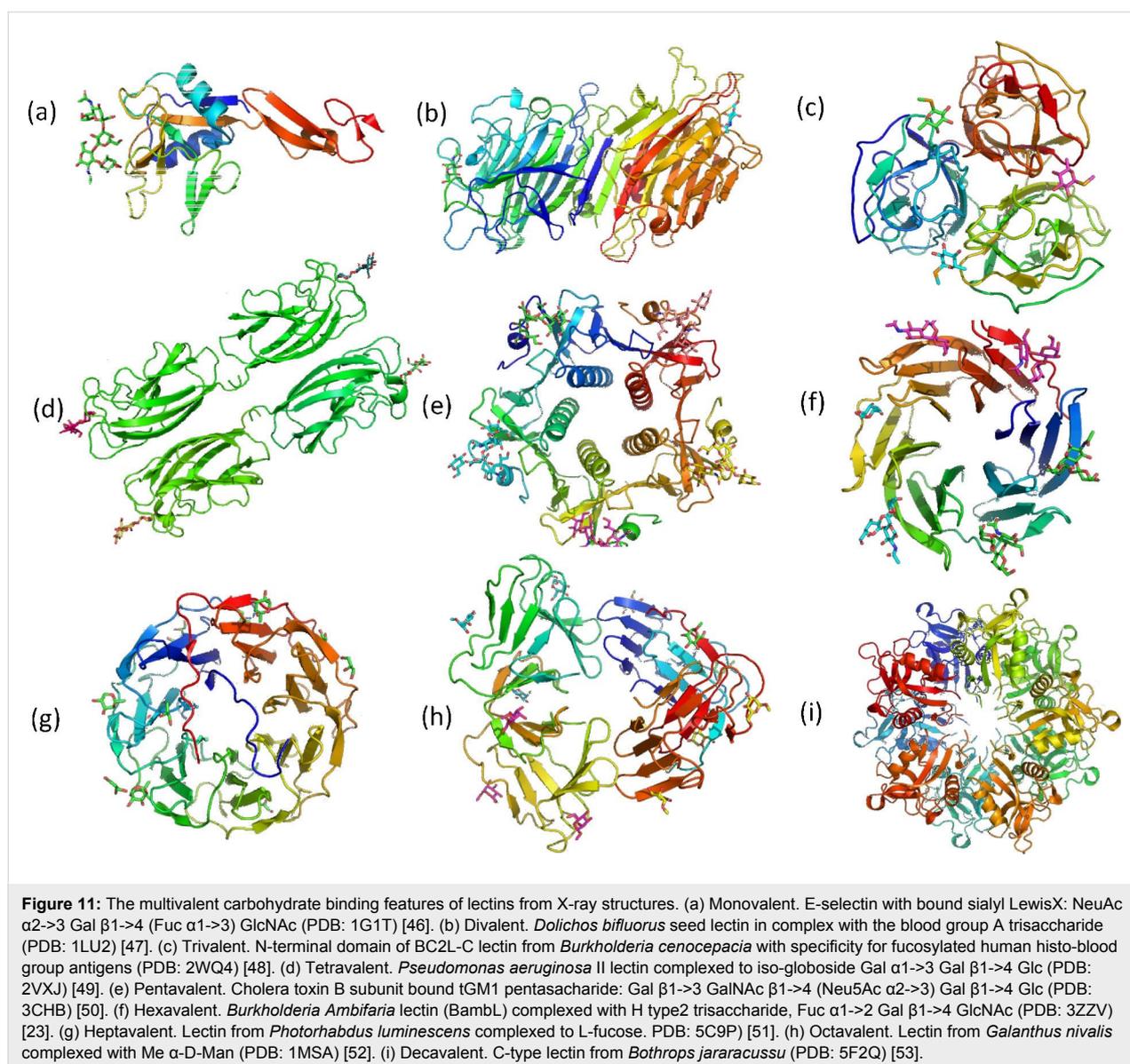
or immunological activity. At the present time, the number of crystal structures of lectins deposited in the PDB amounts to about 1,500. Interestingly, about 60% of them were obtained ligated to carbohydrates, which range from monosaccharides to 10 residue-long oligosaccharides. Lectins occur in plants, animals, algae, bacteria, fungi and yeasts, and viruses. Their involvement in key biologically-important recognition pro-

cesses is well documented, as in the case of embryogenesis, fertilization, inflammation and metastasis. Lectins play a key role in parasite-symbiotic recognition in microbes and invertebrates of plants and vertebrates. The present role assigned to lectin lies in their ability to decipher sugar-encoded information, i.e., they are a molecular reader of the glyco-code.

The plethora of three-dimensional structures of lectins, both in unbound form or complexed with oligosaccharides, lead to their organization in a dedicated database, available at <http://glyco3d.cermav.cnrs.fr> [45]. Information contained in the database provided description of the main features of this important class of proteins. Lectins exhibit a variable oligomeric assembly that ranges from mono- to deca-valency (Figure 11).

Lectins present sugar-binding sites that are in most cases relatively shallow, and are located near the surface and therefore accessible to solvent. One or two calcium ions, identified in several lectin families of different origins, are involved in the carbohydrate binding by direct coordination to the sugar hydroxy groups. The comparison of detailed conformational features of oligosaccharides and their modes of interaction with the protein led to the development of different molecular modelling methods.

A somehow indirect application of the fine specificity of the binding of oligosaccharides to lectins has been elegantly developed to solve the phase problem in protein crystallography. Selenium-labelled carbohydrates can bind to the combining site of lectins at relatively low concentration, and provide sufficient



anomalous signals for MAD or SAD methods of phasing to work, as was exemplified by the structure solution of the F17-G fibrial adhesion [54]. This elegant approach was used to elucidate the crystal structure of *Ralstonia solanacearum* lectin [55], *Parkia platycephala* lectin [56] and *Psathyrella velutina* lectin [57].

**Anti-carbohydrate antibodies:** Carbohydrate determinants are expressed on the cell surface through glycoproteins and glycolipids where they are exposed to a wide range of contexts, surroundings and surface densities. It is within such a landscape that antibodies recognize carbohydrate determinants. Data from many systems have shown that the minimum epitopes are often found at the extremity of the determinant. As a result, the presentation of the carbohydrate on the target cell may be such that antibodies with similar specificity exhibit different selective cell-profiling. Up to now, crystallographic studies of carbohydrate-antibodies mainly concentrated on systems where carbohydrates are complexed with antibody (Fab) or variable fragments (Fv). The organization of a small database of high-resolution three-dimensional structures of carbohydrate-antibody complexes [45] provides a way to classify the different types of bindings. Antibodies that recognize a terminal carbohydrate motif present a cavity-like binding feature, while a groove-like binding site is found in antibodies that bind to internal carbohydrate motifs. This is a mechanism typically found in bacterial polysaccharides. There is an occurrence of very large cavities which are open at both ends. The “side-on” entry of the antigen is often at the origin of the occurrence of conformational antigens.

**Glycosaminoglycan–protein complexes:** As members of the proteoglycan family, glycosaminoglycans (GAGs) are linear polysaccharides, constituted by 40 to 100 repeating disaccharide units, which are usually found to be linked to core proteins. These polysaccharides are components of the peri- and extracellular matrix and are present on surfaces or close to surfaces of animal cells. Based on their core repeat disaccharide units, glycosaminoglycans are classified in four groups: the heparin/heparin group; the chondroitin/dermatan sulfate group, the keratin sulfate group, and the hyaluronic group. With the exception of the later, many sources of structural micro-heterogeneities occur as the epimerization at the C-5 position of uronic acids, and N- and O-sulfation. Of paramount interest is the elucidation of the role of GAGs in their interactions with such important proteins as extracellular matrix proteins, chemokines, growth factors, complement proteins, enzymes, and viruses [58].

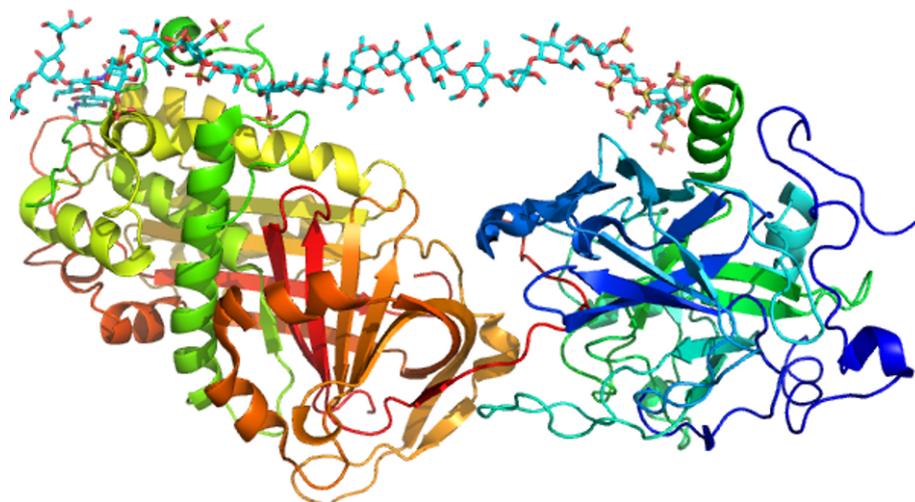
The three-dimensional structures of proteins co-crystallized in interaction with GAG fragments have been organized in a data-

base (<http://glyco3d.cermav.cnrs.fr/>). Because of their relevance for pharmaceutical application, most of these fragments are heparin oligosaccharides. Crystallization of proteins in complex with GAGs is very difficult because of the high degree of heterogeneity and intrinsic flexibility of GAGs. The crystal structure of a fragment as long as a hexadecasaccharide could be co-crystallized as complexed with thrombin and antithrombin at 2.5 Å resolution. For the time being, this is one of the largest oligosaccharide structures ever established throughout macromolecular X-ray crystallography (Figure 12) [59].

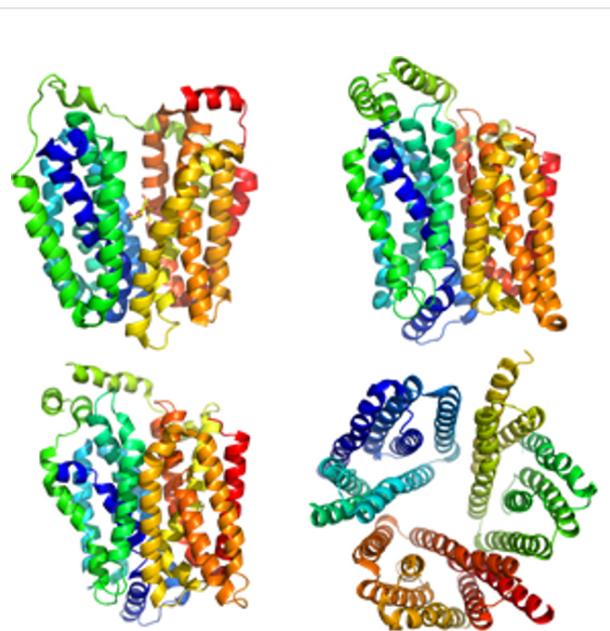
**Transporters:** Soluble sugars serve many purposes in complex organisms. Their cellular exchange relies on transport proteins that are responsible for uptake or release. To date, three main families of eukaryotic transporters have been identified GLUTs, SGLTs, and SWEETs – the most recently discovered sugar transport family, which is responsible for cellular export. In mammals, 14 monosaccharide transport proteins GLUTs are responsible for the diffusion of glucose, galactose, fructose, urate, myoinositol, and dehydroascorbic acid. SGLTs are sodium-glucose symporters that couple the transport of glucose to sodium ions. SWEETs have been characterized the most recently. Major carbohydrate transporters mediate an active uptake and efflux of various mono- and disaccharides. The low affinity of these proteins for sugars seems to be a characteristic feature of transporters involved in high turnover rates, rather than a highly specific transport at low levels of substrates. The structure of the first transporter to be determined was the one of lactose permease LacY [60]. Later on, the structures of different bacterial homologues were also solved. It is only recently that the structure of human GLUT1 was reported [61]. Nevertheless, the joint difficulty to solubilize and crystallize membrane proteins, explains the paucity of crystal structures deposited in the database [62-72]. This is even worse for those proteins involved in the transport of sugars (Figure 13)

### Kinetic crystallography

Since the biological activity of many proteins is preserved in the crystalline state, the possibility to investigate the dynamic process of their mechanisms is absolutely intriguing. In kinetic crystallography, a biological reaction is initiated in the crystal and the fates of the transient species formed are followed by determining the structural changes. Depending upon the time scale of the reaction and the scheme used for its initiation, time-resolved crystallography requires either the use of fast diffraction techniques such as Laue diffraction (polychromatic beam), or the capture of intermediates by trapping methods. These trapping strategies require the complementary use of UV–visible single-crystal spectroscopy. Providing extreme care is taken to avoid artefacts, these methods are in principle available to a



**Figure 12:** Three-dimensional depiction of the ternary complex formed by a heparin mimetic in interaction with antithrombin. The structure has been solved at 2.6 Å resolution (PDB 1TB6) [59]. The basis of the antithrombotic properties of therapeutic heparin could partly be deciphered by the availability of such a three-dimensional structure.



**Figure 13:** 3D representation of different sugar transporter structures: (left to right, top to down) lactose permease structure (PDB 1PV7, [60]), of the human glucose transporter GLUT1 (PDB 4PYP, [61]), of the bovine fructose transporter (GLUT 5) (PDB 4YB9, [73]) and of a SWEET transporter of *Orzyva sativa* (PDB 5CTG, [72]).

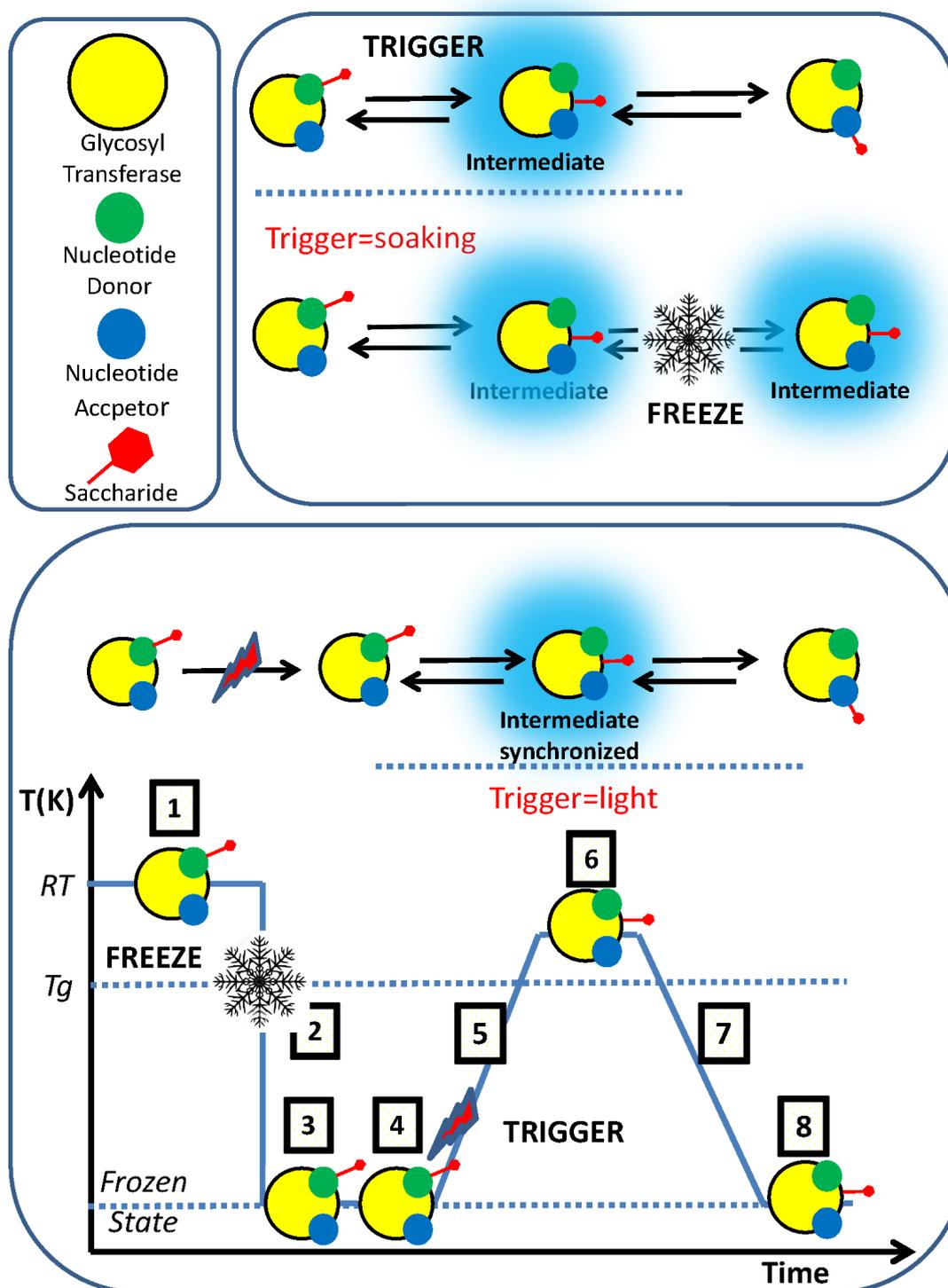
wide range of biological systems. Two types of intermediate trapping schemes are available.

In the so-called “trigger-freeze” experiment, a large fraction of molecules is brought into the intermediate state of interest at room temperature, which is trapped by flash-cooling. While in a

“freeze-trigger” experiment, the sample is first flash-cooled, and then the reaction is triggered potentially after a transient and controlled temperature increase [74].

The ‘trigger-freeze’ approach consists in the use of various soaking times for the crystal with substrate sugars in presence of H<sub>2</sub>O or glycerol (“trigger step”) followed by the freezing step ranging from a few minutes to several hours. The use of a “freeze-trigger” approach solves the synchronization issue but introduces experimental complications as a photo-activable analogue called a caged-compound is required. For this, the substrates have to be modified by adding a photolabile group that prevents the reaction from occurring. The principles of the ideal cage-compound based kinetic crystallography experiment are presented in Figure 14.

The studies of the mechanism of blood group glycosyl transferase have been investigated by kinetic crystallography approaches with the aim of characterizing the double-displacement mechanism which involves the formation of a covalently bound glycosyl-enzyme intermediate, by trapping and solving the X-ray structure of this intermediate [75]. The A and B antigenic determinants are synthesised by the blood group A (GTA) and the blood group B (GTB) glycosyltransferases which transfer GalNAc from UDP-GalNAc for the A type and a Gal residue from UDP-Gal for the B-type. A mutant of the galactosyl transferase was created with the capacity to act as GTA and GTB [76,77]. The first attempt was conducted using the “trigger-freeze” method. The experiments were inconclusive presumably due to the lack of synchronisation of the reaction within the crystals and because the reaction time scale is shorter



**Figure 14:** Kinetic crystallography. Protein crystals are soaked with the cage compound (Step 1) followed by flash-cooling (Step 2). The structure determination of the complex is solved (Step 3) to see if and how the cage-compound binds to the protein active site. Step 4 is the cleaving of the cage compound using a laser of appropriate wavelength. In this state, the substrate is available for hydrolysis but the frozen state prevents this from happening. Step 5 consists of a slow increase of the temperature to cross the glass transition and reach a temperature region where the protein has a greater degree of conformational flexibility but with a reaction rate slower than at room temperature. In Step 6, with the enzyme at room temperature, the reaction can proceed with a reorganization of the active site to transfer the sugar to a nucleophilic amino acid. An intermediate is formed and trapped during Step 7, which consists of decreasing the temperature to go back to the frozen state. Step 8 is the structural determination of the protein and elucidation of the intermediate. (Adapted from [75] with permission from Dr. G. Batot).

than the time scale of substrate diffusion in the crystals. For example, when UDP-GalNAc was soaked for 24 minutes, experiments resulted in structures with UDP-GalNAc in several conformations that are difficult to interpret. The “freeze-trigger” route was started using a series of cage compounds that had been synthesized. They all included a substrate donor, either UDP-Gal or UDP-GalNAc with an additional group on the sugar or on the uracil. Photolysis at 100 K was monitored by UV–vis absorption, both in solution and in crystals in order to assess the efficiency of the laser ablation in the crystalline glycosyl transferase. The four steps of the “freeze-trigger” process could be validated throughout by elucidation of the crystal structure of the glycosyl transferase, which has the active site occupied in a semi-closed conformation of the substrate with various levels of ordering of the internal flexible loop.

### Small angle X-ray scattering

Small angle X-ray scattering is a universal technique whereby X-rays are recorded that have been elastically scattered at a low angle from samples in solution. Analysis of the scattered X-rays allows low-resolution structural information to be obtained, such as average particle size, distribution and shape. Different kinds of samples beside soluble proteins can be studied by this technique including nucleic acids, protein-based complexes, lipids, membrane proteins and surfactants, glycoproteins, virus, polymers and colloids [78,79].

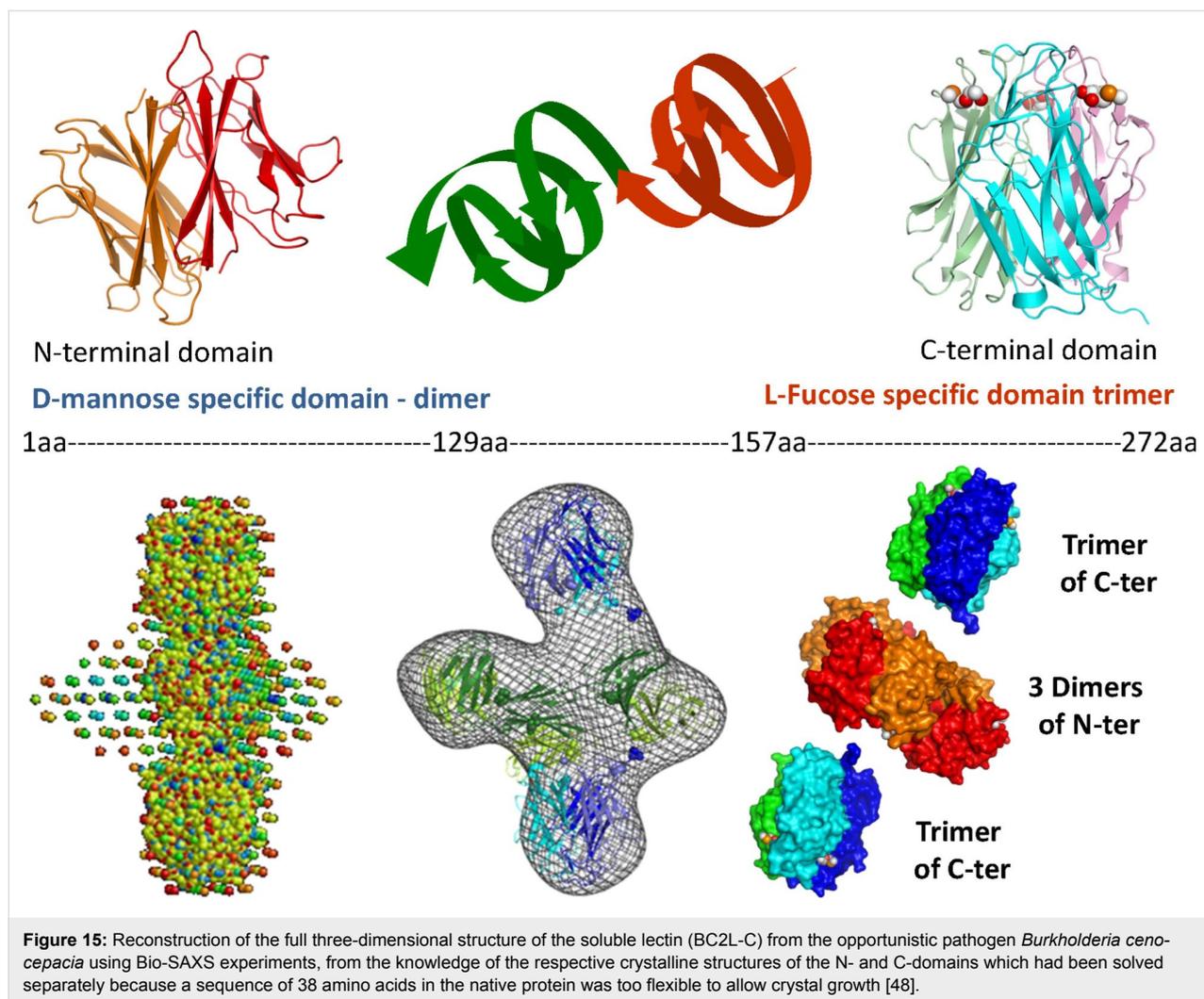
**Proteins:** SAXS applied to biological materials (BioSAXS) is a complementary tool to protein crystallography and has become an invaluable resource for structural biologists [80]. Although at a much lower resolution than protein crystallography, BioSAXS permits the structural analysis of macromolecules at more physiological conditions, besides being suitable for the study of heterogenous systems that are unlikely to crystallise. Furthermore, the experiments in solution allow the effect of other factors, such as pH, ion concentration, or temperature, on the overall protein structure to be studied. Samples for structure analysis should be highly monodisperse. Besides sample quality control by using complementary analysis, such as dynamic light scattering, native gel, ultracentrifugation, many BioSAXS beamlines at synchrotrons are nowadays equipped with size exclusion chromatography devices directly connected with the sample cell and the data acquisition systems [81].

An illustration of how BioSAXS experiments can help to complete data obtained by protein crystallography is given by the characterization of the full structural assembly of the lectin of *Burkholderia cenocepacia*, an opportunistic bacterial pathogen. Throughout biochemical characterization, the lectin, BC2L-C was shown to be composed of two distinct domains, each

displaying unique specificities and biological activities. The protein is a super lectin that binds independently to fucosylated human histo-blood group epitopes and to mannose/heptose glycoconjugates. The N-terminal domain is a fucose-binding lectin having similarity with tumour necrosis factor. The structure of the other domain (C-terminal part) which belongs to the superfamily of calcium-dependent lectins displays specificity for mannose and L-glycero-D-manno-heptose monosaccharides. The two domains are linked by a conformationally flexible sequence of 38 amino acids which was detrimental for crystallization. The respective crystal structures of the N- and C-domains could be solved separately and eventually used to establish the overall structure of the assembly by small-angle X-ray scattering (and further confirmed by electron microscopy). Figure 15 displays the reconstruction of the full macromolecular complex as a flexible arrangement of three mannose/heptose-specific dimers flanked by two fucose-specific TNF- $\alpha$ -like trimers [48]. This study (along with many other examples) highlights the potential of SAXS to decipher the global state of glycoproteins and carbohydrate binding proteins in solution so as to greatly amplify the high resolution 3D structural information derived from macromolecular crystallography of domains of small proteins.

**Colloids:** Scattering methods light, neutron and X-ray have long been the methods of choice to investigate the states of soft-condensed materials, which include solutions and gels. Differences in wavelengths and scatterers can be used for combined measurements yielding supplemental information. Following the instrumental developments of light sources, SAXS has become a common tool for the investigations of the state of materials in solution at the nanoscale. Many studies have been devoted to polysaccharides, for which structural change have been observed in real time. There exists extensive literature on this subject and the role played by polysaccharide association structures in food and in biomedical applications, as hydrogels, triggers the development of novel experiments and tools, such as optical tweezers, while making use of synchrotron radiation. Description of the details of molecular interactions occurring between complex materials such as polysaccharides and mucin is among the many new achievements that yields to the rational design of muco-adhesive polysaccharide-based nano-formulations [77].

The availability of new instrumentation that combines wide and small angle X-ray scattering and high resolution ultra-small angle X-ray scattering in a time-resolved manner is creating an opportunity to investigate the microstructure and non-equilibrium dynamics of soft matter on a length scale from a few angstroms to micrometers and on a timescale descending to the millisecond.



## Grazing incidence X-ray reflectometry

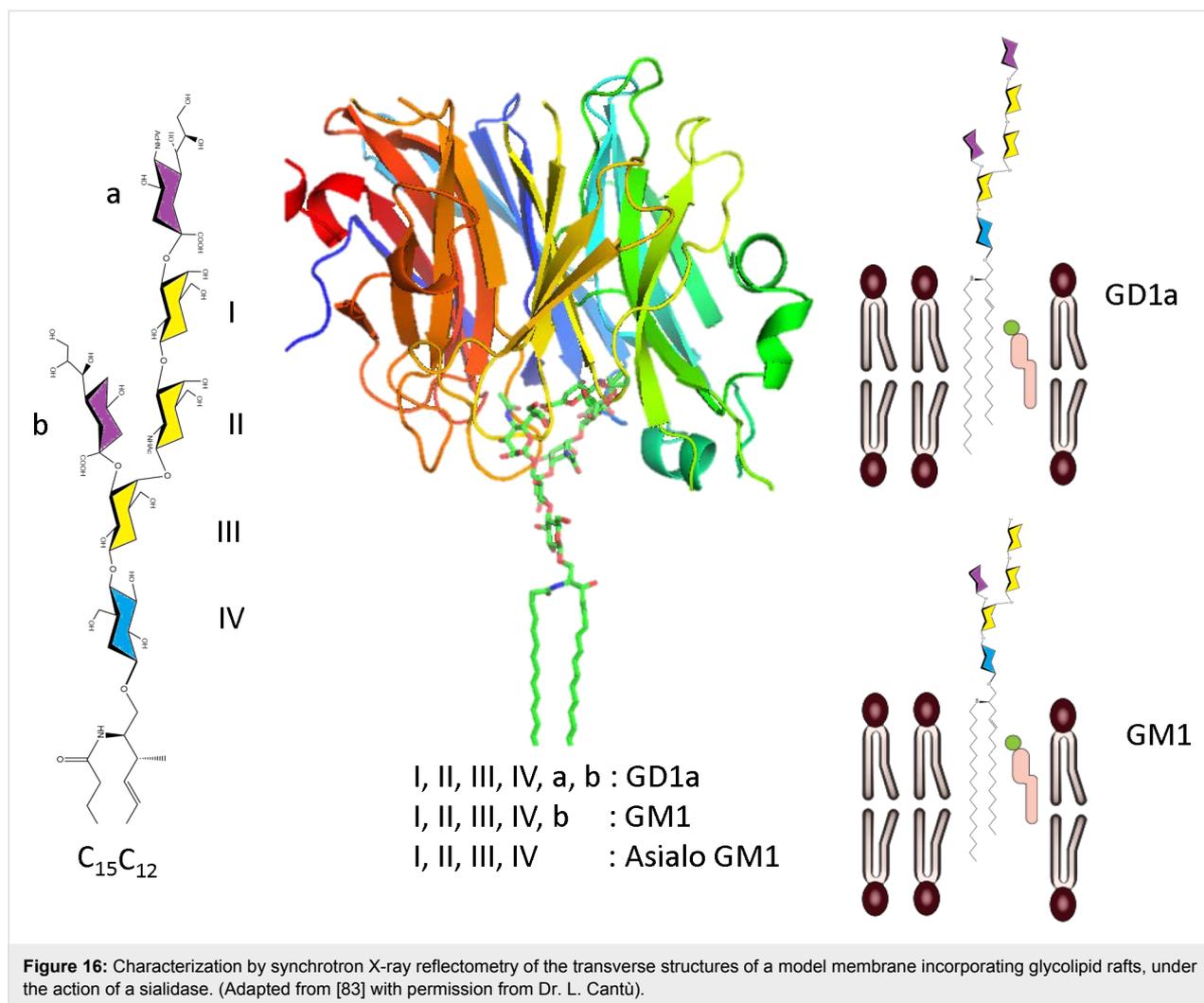
**Glycolipids:** Despite their importance in the constitution and dynamics of plasma membranes, the structural and physico-chemical features of gangliosides have been somehow neglected presumably because of the lack of appropriate experimental techniques. X-ray reflectometry is a surface-sensitive analytical technique based on the measure of the intensity of X-ray reflected by a flat surface. Any deviation from surface flatness will result in deviation of the reflected beam which can be analyzed to obtain the density profile of the interface normal to the surface [82].

Synchrotron X-ray reflectometry has been used to access the transverse structure of a biomimetic plasma membrane incorporating glycolipid rafts. The in situ chemical conversion of GD1a ganglioside into its metabolic product under the action of sialidase was investigated. The outcome of the sialidase action is not limited to the creation of GM1 and AsialoGM1 gangliosides as it is accompanied by a reshaping of the membrane

which involves a rearrangement of the headgroups on the surface (Figure 16) [83].

## Polysaccharide structures

In contrast to other macromolecules and because of the lack of regular crystalline order, X-ray diffraction of polysaccharides usually leads to an insufficient number of reflections to permit structural determination based on the data alone. Such a lack of experimental data must be complemented by modelling techniques. As such, the process of structural elucidation combines the calculation of diffraction intensities from various low energy models with those intensities collected on X-diffractograms. In this context, it is even most appropriate to use the term ‘model’ in place of ‘structure’. These experimental limitations explain why so few models of polysaccharides have been reported: there are just over 100, counting all polymorphs, variants, derivatives, and complexes. Because of their ubiquitous occurrence, polymorphs of celluloses and starches have attracted most of the attention.



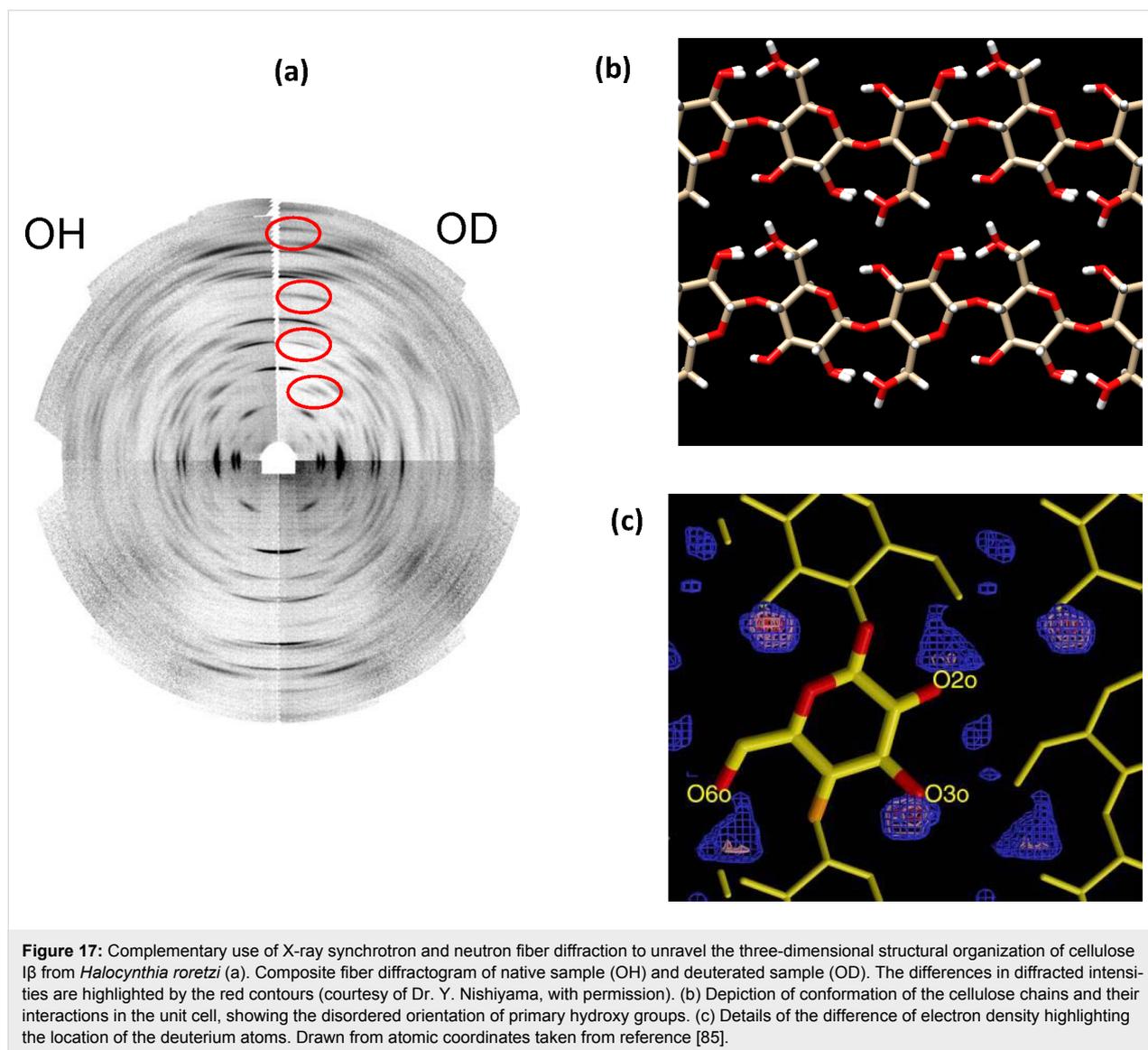
### Fibrillar structures: cellulose and starch

**Cellulose:** The first X-ray fiber diffractograms of native cellulose were reported more than one century ago. The results of the investigations that have been undertaken left many of the structural details unclear as conflicting structural models were reported. One particular obstacle to be overcome in the study of cellulose microfibrils is the co-existence of a mixture of two crystal forms I $\alpha$  [84] and I $\beta$  [85]. In light of this allomorphy, the elucidation of the structure of cellulose I, awaited the mature developments of large scale facilities of synchrotron and neutron sources, and the mastering of deuteration methods of the intra-crystalline regions of the native cellulose samples without altering the overall structural integrity. Through an ingenious combination of synchrotron and neutron fiber diffraction, a highly accurate structural model could be established. The samples diffracted to better than 1 Å resolution, and provided the determination of C- and O-atoms positions from the set of X-ray diffracted intensities (Figure 17). In addition, the position of hydrogen atoms were determined from Fourier-differ-

ence analysis from the set of neutron diffracted intensities collected from hydrogenated and deuterated samples [85].

This resulted in a description of the three-dimensional features of both allomorphs of native cellulose. Nevertheless, a detailed elucidation of the biosynthetic mechanism is still required to understand the occurrence of two different structural arrangements within the same microfibrils. Some unexpected features still need to be elucidated and this would require the use of complementary methods.

**Starch:** The complexity of starch in terms of the nature and size of its macromolecular components (amylopectin, amylose) has always been an obstacle to the elucidation of the structural components and their arrangements, which are at the origin of the birefringence of a starch granule. The structure of the crystalline domains of the two allomorphs of starch granules found in cereal and tubers had been established from a series of experimental observations (X-ray and electron crystallography) and

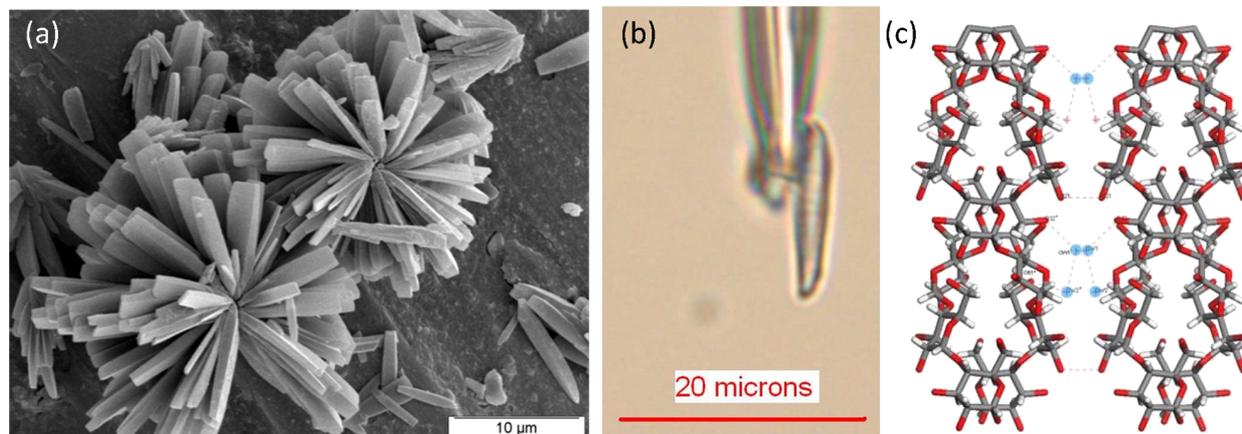


molecular modelling. While displaying differences in their mode of interactions, both allomorphs are characterized by a parallel arrangement of parallel-stranded left-handed double helices [86,87]. A second look at the crystal structure of the A-polymorph became possible when microcrystals were grown from short chains of synthetic starch and diffraction data collected using a micron-sized beam at a synchrotron source. While this new investigation corroborated the essential features of the original model, some additional fine details were revealed (Figure 18) [88].

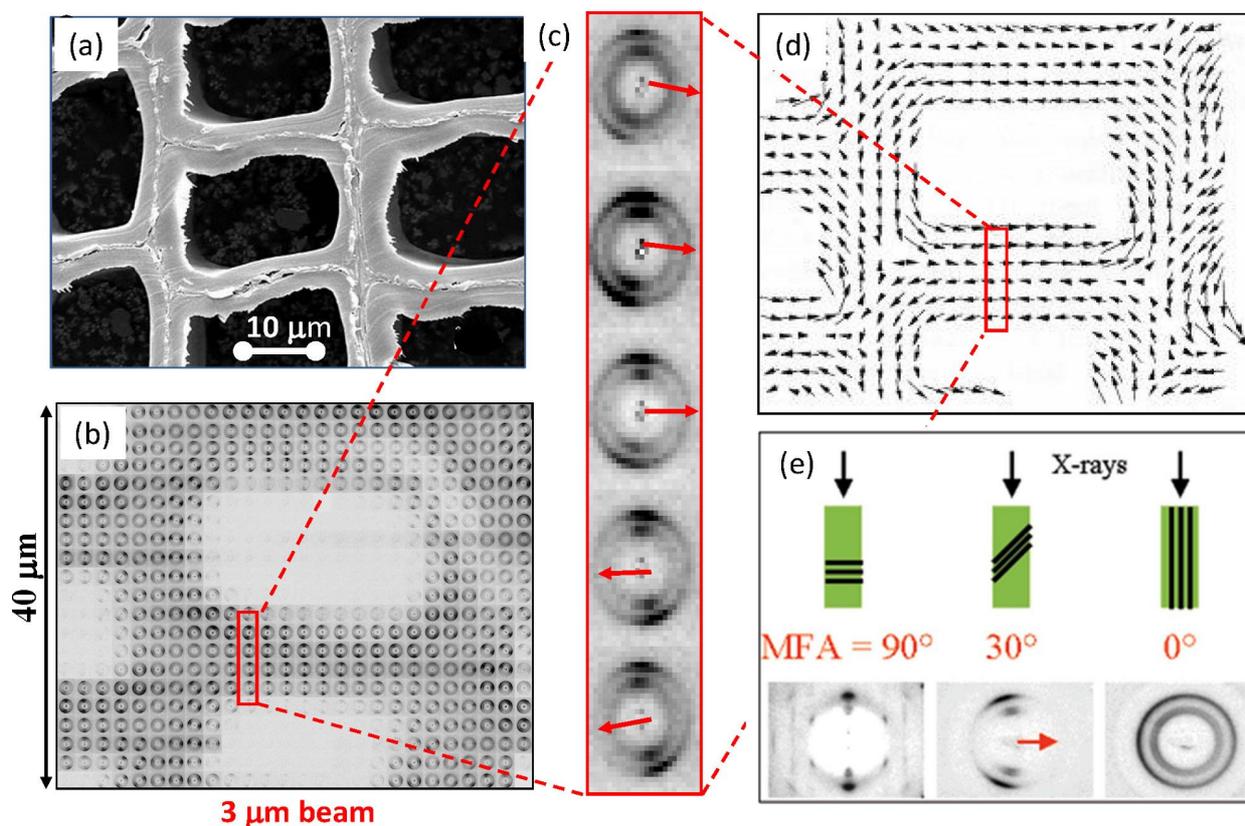
### Multiscale organization

**Cellulose:** Knowledge of the structure of a material is necessary to understand its properties. In the case of cellulose, it is also the key to ascertain the processes of biosynthesis. Cellulosic materials in nature often have many levels of structural

complexity, whose organization depends on the source organism. In wood, a cohesive interlaced network of crystalline microfibrils of cellulose composes a first level of interacting components of the cell walls. The typical dimensions of the cellulosic fibers, which are composed of 30–40 cellulose chains, have lengths in the region of 1–2 nm and width of about 35 Å. The elucidation of the structural organization of these microfibrils came from the use of a micro-focused X-ray beam of 3  $\mu\text{m}$  on a wood section of 10  $\mu\text{m}$  thick and oriented perpendicular to the incident beam [89]. As depicted in Figure 19, a complete distribution map of the orientation of the axes of the cellulose microfibril, of a specimen of 52  $\times$  42  $\mu\text{m}$ , was established through a series of 546 diffraction patterns. These results can be translated into three-dimensions and establish the existence of an ultra-structural organisation in which the orientations of the cellulose fibrils follow a super-helicoidal fashion.



**Figure 18:** Scanning electron micrograph of high-quality micrometer-sized A-amylose microcrystals grown from short chains of synthetic starch (a) and of a single crystal glued to a borosilicate glass capillary tip (b). (c) Three dimensional representation of the crystal and molecular structure derived from X-ray synchrotron diffraction. (a) and (c) taken from reference [91] with permission from Actualité Chimique (<http://www.lactualitechimique.org>). (b) (Courtesy, J. L. Putaux; a very similar image of the subject was published in reference [88]).



**Figure 19:** Cartography of distribution and orientation of cellulose in wood using a 3 μm X-ray beam. The scanning of a 10 μm thick wood section, by increments of 2 μm, (a) a collection of "fiber" like diffractograms was collected (b). From each fiber diffractogram, the attribution of the diffraction spots indicates the local orientation of the microfibril axis (c). This is depicted by arrows which indicate a marked local asymmetry in the microfibril (d). The integration of the degrees of disorientation over the full map gives the orientation of the microfibril angle (MFA) along the direction of propagation (e) (adapted from reference [89], and with permission of the International Union of Crystallography, <http://journals.iucr.org>).

TEMPO-mediated oxidation is one of several methods that can be used to extract nascent crystals of cellulose, or cellulose microfibrils from biomass. For this process to be optimal, some fundamental aspects of the structural and ultrastructural characterization of the cellulosic material have to be ascertained. Indeed, the extraction process should be adapted to the specificity of the various sources (e.g., wood, cotton, jute, bamboo, etc.). For dispersions in aqueous suspension, the structure of cellulose nanofibers (and their aggregates) can be characterized by SAXS. This technique has permitted quite significant insight to be gained about the structure of cellulose from a variety of botanical origins. In the case of wood pulp, cellulose nanofibers displayed a ribbon shape of about one micrometer in length. The cross-sections sizes were found to cluster in two groups with dimensions of  $3\text{ nm} \times 8\text{ nm}$  and  $9\text{ nm} \times 20\text{ nm}$ , respectively. Quite different results were obtained for the structure of microfibril fractions extracted from never-dried delignified spruce wood. In this case, the observed morphology was of the type “nanostrips” that had a characteristic thickness and width of about  $0.5\text{ nm}$  and  $4\text{ nm}$ , respectively. The thickness is an indication that the nanostrips are made up of only one monolayer of cellular material, which indicated the occurrence of “two-dimensional” crystals that could be further investigated by wide-angle X-ray diffraction [90].

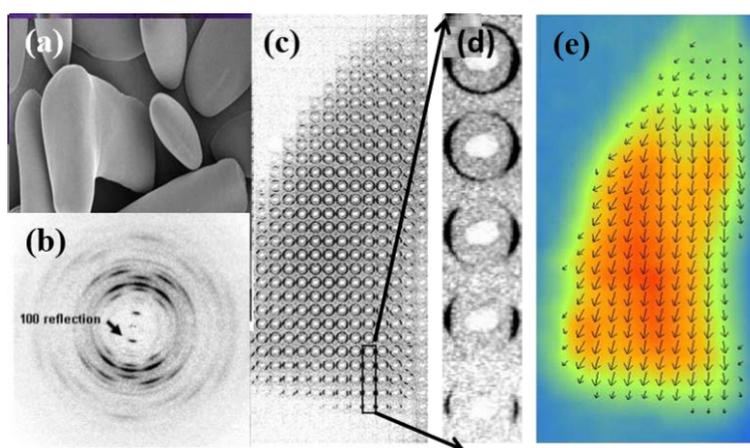
**Starch:** Depending upon their botanical origin, starch granules display an elliptical shape with dimensions ranging from  $0.1$  to  $100\text{ }\mu\text{m}$ . The advent of micro-focus X-ray diffraction from synchrotron radiation offered the possibility to explore the arrangements of the crystalline domains which are at the origin

of the birefringence of the starch granule. Using a  $2\text{ }\mu\text{m}$  wide X-ray beam, a complete cartography of the relative orientation on a single granule could be drawn. Two-dimensional fiber diffraction patterns were collected for each domain on a grid of  $4 \times 4\text{ }\mu\text{m}$ . Information about the nature of the crystalline structure (for allomorphs A, B and C) was obtained confirming the orientation of amylopectin double helices in the crystalline lamellae as well as the location of these domains and their relative orientation with respect to the granule. The most detailed investigation performed on potato starch (B allomorph) indicates that the double helices do not seem to point towards a single focus but rather towards the surface of an inner ellipsoid. Thus, the double helices have a radial orientation, and are perpendicular to the surface of the granule.

At the resolution of these ultra-structural features ( $10\text{ }\mu\text{m}$ ) there is no discontinuity of orientation, i.e., no disclination of orientation. Between  $10\text{ }\mu\text{m}$  steps, the change of the direction of the double-helices is gradual, which is consistent with a radial orientation (Figure 20) [92].

## Conclusion

The aim of the present article was to describe how synchrotron radiation has benefited the field of structural glycoscience in the studies of complex carbohydrates. The atomic structures of numerous (macro)-molecules have been revealed, from molecular single crystals all the way to the complexity of polysaccharide architectures, throughout the field of protein–carbohydrate interactions. Seemingly, the study of the less well explored area of colloids and glycolipids in their membrane environments can



**Figure 20:** Structural micro-diffraction scanning of a starch granule from *Phajus grandifolius* with dimensions  $50 \times 200\text{ }\mu\text{m}$ . (a) Scanning electron micrograph. (b) X-ray diffractogram collected from a peripheral region of the starch grain. The width of the azimuthal (100) reflection indicates the level of crystallinity. (c,d) The cartography of the crystalline domains collected on a single grain, on a grid having  $4 \times 4\text{ }\mu\text{m}$  dimensions, using a  $1\text{ }\mu\text{m}$  X-ray beam. The total surface explored was greater than  $5,000\text{ }\mu\text{m}^2$ . (e) The width and orientation of the (100) reflection on each diffractogram reflects the level of crystallinity of the explored section along with their relative orientation with respect to the fiber axis. The experiment was performed at  $100\text{ K}$  to limit the degradation of the grain in the X-ray beam. (Taken from reference [91] with permission from Actualité Chimique <http://www.lactualitechimique.org>).

be tackled. The increasing speed of data collection times and photon flux are opening the way to time-resolved studies. The application of kinetic crystallography to elucidate glyco-enzymatic mechanisms is still at its infancy. Complementary to instrumental developments, the contribution of organic synthesis will be essential for the development of cage compounds, tailored to initiate light-activated reactions.

The results that have been presented were obtained on third-generation synchrotron sources. More sophisticated fourth-generation X-ray linear sources (X-ray Free Electron Lasers – XFEL) are operating at Stanford (USA), Hamburg (Germany) and in Harima (Japan). The brightness of the X-ray beams are then orders of magnitude greater and with short pulses, down to a few femtoseconds. A world full of novel experiments can be envisaged involving diffraction as well as the possibility to image non-periodic materials. Furthermore, different third-generation sources are planning major upgrades of their machine lattice to produce diffraction limited storage rings (DLSR) that will open new avenues in the science performed at these sources.

Synchrotron radiation offers much more than diffraction experiments and many other experiments can be performed making use of either the spectroscopy or imaging techniques. Spectroscopy techniques allow identification and characterization of molecular substances and their dynamics. Imaging techniques use the light-source beam to obtain pictures with spatial resolution of the sample under study. Integration of the results gathered from such experiments is a requirement to get deeper insight into the structures and mechanisms of vital biological processes in plants, animals and human.

## Acknowledgments

The authors are indebted to the many colleagues who provided material to illustrate the wealth of applications and results that have been achieved using synchrotron light over the recent years. Gary Admans (ESRF) is gratefully acknowledged for critical reading of the manuscript. Part of this work was supported by the Cross Disciplinary Program Glyco@Alps, within the framework "Investissements d'avenir" program (ANR-15-IDEX-02). The three-dimensional representations of the proteins were drawn using the PyMol Molecular Graphics Systems, Version 1.8 Schrödinger, LLC.

## References

- Eriksson, M.; van der Veen, J. F.; Quitmann, C. *J. Synchrotron Radiat.* **2014**, *21*, 837–842. doi:10.1107/S1600577514019286
- Bianconi, A.; Dell'Arccia, M.; Durham, P. J.; Pendry, B. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1982**, *26*, 6502–6508. doi:10.1103/PhysRevB.26.6502
- Bragg, W. H.; Bragg, W. L. *Proc. R. Soc. London, Ser. A* **1913**, *88*, 428–438. doi:10.1098/rspa.1913.0040
- Martinez-Criado, G.; Borfecchia, E.; Mino, L.; Lamberti, C. Micro and Nano X-ray beams. In *Characterization of Semiconductor Heterostructures and Nanostructures*; Lamberti, C.; Agostini, G., Eds.; 2013; pp 361–412. doi:10.1016/b978-0-444-59551-5.00009-1
- Riekel, C. *Rep. Prog. Phys.* **2000**, *63*, 233. doi:10.1088/0034-4885/63/3/201
- Kirkpatrick, P.; Baez, A. V. *J. Opt. Soc. Am.* **1948**, *38*, 766–774. doi:10.1364/JOSA.38.000766
- Snigirev, A.; Kohn, V.; Snigireva, I.; Lengeler, B. *Nature* **1996**, *384*, 49–51. doi:10.1038/384049a0
- Hendrickson, W. A.; Teeter, M. M. *Nature* **1981**, *290*, 107–113. doi:10.1038/290107a0
- Karle, J. *Int. J. Quantum Chem.* **1980**, *18*, 357–367. doi:10.1002/qua.560180734
- Fischer, E. *Ber. Dtsch. Chem. Ges.* **1891**, *24*, 1836–1845. doi:10.1002/cber.189102401311
- Raymond, S.; Heyraud, A.; Tran Qui, D.; Kvik, A.; Chanzy, H. *Macromolecules* **1995**, *28*, 2096–2100. doi:10.1021/ma00110a051
- Rencurosi, A.; Mitchell, E. P.; Cioci, G.; Pérez, S.; Pereda-Miranda, R.; Imbert, A. *Angew. Chem., Int. Ed.* **2004**, *43*, 5918–5922. doi:10.1002/anie.200460327
- Brunelli, M.; Wright, J. P.; Vaughan, G. B. M.; Mora, A. J.; Fitch, A. N. *Angew. Chem., Int. Ed.* **2003**, *42*, 2029–2032. doi:10.1002/anie.200250607
- Ochsenbein, P.; Kieffer, J.; El Hajji, M. In *European Powder Diffraction Conference, Darmstadt, Germany*, Darmstadt, Germany; 2010.
- Collings, I.; Watier, Y.; Giffard, M.; Dagogo, S.; Kahn, R.; Bonneté, F.; Wright, J. P.; Fitch, A. N.; Margiolaki, I. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 539–548. doi:10.1107/S0907444910005354
- Karavassili, F.; Giannopoulou, A. E.; Kotsiliti, E.; Knight, L.; Norrman, M.; Schluckebier, G.; Drube, L.; Fitch, A. N.; Wright, J. P.; Margiolaki, I. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *D68*, 1632–1641. doi:10.1107/S0907444912039339
- Margiolaki, I.; Giannopoulou, A. E.; Wright, J. P.; Knight, L.; Norrman, M.; Schluckebier, G.; Fitch, A. N.; Von Dreele, R. B. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *D69*, 978–990. doi:10.1107/S0907444913003867
- Rose, P. W.; Prlić, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Kramer Green, R.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. *Nucleic Acids Res.* **2017**, *45*, D271–D281. doi:10.1093/nar/gkw1000
- Stevens, R. C. *Curr. Opin. Struct. Biol.* **2000**, *10*, 558–563. doi:10.1016/S0959-440X(00)00131-7
- Cipriani, F.; Röwer, M.; Landret, C.; Zander, U.; Felisaz, F.; Márquez, J. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *D68*, 1393–1399. doi:10.1107/S0907444912031459
- Flot, D.; Mairs, T.; Giraud, T.; Guizarro, M.; Lesourd, M.; Rey, V.; van Brussel, D.; Morawe, C.; Borel, C.; Hignette, O.; Chavanne, J.; Nurizzo, D.; McSweeney, S.; Mitchell, E. *J. Synchrotron Radiat.* **2010**, *17*, 107–118. doi:10.1107/S0909049509041168

22. de Sanctis, D.; Beteva, A.; Caserotto, H.; Dobias, F.; Gabadinho, J.; Giraud, T.; Gobbo, A.; Guijarro, M.; Lentini, M.; Lavault, B.; Mairs, T.; McSweeney, S.; Petitdemange, S.; Rey-Bakaikoa, V.; Surr, J.; Theveneau, P.; Leonard, G. A.; Mueller-Dieckmann, C. *J. Synchrotron Radiat.* **2012**, *19*, 455–461. doi:10.1107/S0909049512009715
23. Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoën-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerová, M.; Le Pendu, J.; Römer, W.; Varrot, A.; Imberty, A. *J. Biol. Chem.* **2012**, *287*, 4335–4347. doi:10.1074/jbc.M111.314831
24. Gabadinho, J.; Beteva, A.; Guijarro, M.; Rey-Bakaikoa, V.; Spruce, D.; Bowler, M. W.; Brockhauser, S.; Flot, D.; Gordon, E. J.; Hall, D. R.; Lavault, B.; McCarthy, A. A.; McCarthy, J.; Mitchell, E.; Monaco, S.; MuellerDieckmann, C.; Nurizzo, D.; Ravelli, R. B. G.; Thibault, X.; Walsh, M. A.; Leonard, G. A.; McSweeney, S. M. *J. Synchrotron Radiat.* **2010**, *17*, 700–707. doi:10.1107/S0909049510020005
25. de Sanctis, D.; Oscarsson, M.; Popov, A.; Svensson, O.; Leonard, G. *Acta Crystallogr., Sect. D* **2016**, *D72*, 413–420. doi:10.1107/S2059798316001042
26. Stepanov, S.; Makarov, O.; Hilgart, M.; Pothineni, S. B.; Urakhchin, A.; Devarapalli, S.; Yoder, D.; Becker, M.; Ogata, C.; Sanishvili, R.; Venugopalan, N.; Smith, J. L.; Fischetti, R. F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *D67*, 176–188. doi:10.1107/S0907444910053916
27. Delagenière, S.; Brenchereau, P.; Launer, L.; Ashton, A. W.; Leal, R.; Veyrier, S.; Gabadinho, J.; Gordon, E. J.; Jones, S. D.; Levik, K. E.; McSweeney, S. M.; Monaco, S.; Nanao, M.; Spruce, D.; Svensson, O.; Walsh, M. A.; Leonard, G. A. *Bioinformatics* **2011**, *27*, 3186–3192. doi:10.1093/bioinformatics/btr535
28. Monaco, S.; Gordon, E.; Bowler, M. W.; Delagenière, S.; Guijarro, M.; Spruce, D.; Svensson, O.; McSweeney, S. M.; McCarthy, A. A.; Leonard, G.; Nanao, M. H. *J. Appl. Crystallogr.* **2013**, *46*, 804–810. doi:10.1107/S0021889813006195
29. Bowler, M. W.; Nurizzo, D.; Barrett, R.; Beteva, A.; Bodin, M.; Caserotto, H.; Delagenière, S.; Dobias, F.; Flot, D.; Giraud, T.; Guichard, N.; Guijarro, M.; Lentini, M.; Leonard, G. A.; McSweeney, S.; Oskarsson, M.; Schmidt, W.; Snigirev, A.; von Stetten, D.; Surr, J.; Svensson, O.; Theveneau, P.; Mueller-Dieckmann, C. *J. Synchrotron Radiat.* **2015**, *22*, 1540–1547. doi:10.1107/S1600577515016604
30. Higgins, S. J. *Protein Expression: A Practical Approach*; Oxford University Press: Oxford, U.K., 1999; p 304.
31. Garman, S. C.; Wurzburg, B. A.; Tarchevskaya, S. S.; Kinet, J.-P.; Jardetzky, T. S. *Nature* **2000**, *406*, 259–266. doi:10.1038/35018500
32. Lee, J. E.; Fusco, M. L.; Saphire, E. O. *Nat. Protoc.* **2009**, *4*, 592–604. doi:10.1038/nprot.2009.29
33. Agirre, J.; Ariza, A.; Offen, W. A.; Turkenburg, J. P.; Roberts, S. M.; McNicholas, S.; Harris, P. V.; McBrayer, B.; Dohnalek, J.; Cowtan, K. D.; Davies, G. J.; Wilson, K. S. *Acta Crystallogr., Sect. D* **2016**, *D72*, 254–265. doi:10.1107/s2059798315024237
34. Agirre, J.; Iglesias-Fernández, J.; Rovira, C.; Davies, G. J.; Wilson, K. S.; Cowtan, K. D. *Nat. Struct. Biol.* **2015**, *22*, 833–834. doi:10.1038/nsmb.3115
35. Breton, C.; Fournel-Gigleux, S.; Palcic, M. M. *Curr. Opin. Struct. Biol.* **2012**, *22*, 540–549. doi:10.1016/j.sbi.2012.06.007
36. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. *J. Mol. Biol.* **2003**, *328*, 307–317. doi:10.1016/S0022-2836(03)00307-3
37. Albesa-Jové, D.; Guerin, M. E. *Curr. Opin. Struct. Biol.* **2016**, *40*, 23–32. doi:10.1016/j.sbi.2016.07.007
38. Pedersen, L. C.; Dong, J.; Taniguchi, F.; Kitagawa, H.; Krahn, J. M.; Pedersen, L. G.; Sugahara, K.; Negishi, M. *J. Biol. Chem.* **2003**, *278*, 14420–14428. doi:10.1074/jbc.M210532200
39. Hu, Y.; Chen, L.; Ha, S.; Gross, B.; Falcone, B.; Walker, D.; Mokhtarzadeh, M.; Walker, S. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 845–849. doi:10.1073/pnas.0235749100
40. Rocha, J.; Cicéron, F.; de Sanctis, D.; Lelimosin, M.; Chazalet, V.; Lerouxel, O.; Breton, C. *Plant Cell* **2016**, *28*, 2352–2364. doi:10.1105/tpc.16.00519
41. Lombard, V.; Bernard, T.; Rancurel, C.; Brumer, H.; Coutinho, P. M.; Henrissat, B. *Biochem. J.* **2010**, *432*, 437–444. doi:10.1042/BJ20101185
42. Henrissat, B.; Davies, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637–644. doi:10.1016/S0959-440X(97)80072-3
43. Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. *J. Biochem. J.* **2004**, *382*, 769–781. doi:10.1042/BJ20040892
44. Gilbert, H. J.; Knox, J. P.; Boraston, A. B. *Curr. Opin. Struct. Biol.* **2013**, *23*, 669–677. doi:10.1016/j.sbi.2013.05.005
45. Pérez, S.; Sarkar, A.; Rivet, A.; Breton, A.; Imberty, A. Glyco3D: A Portal for Structural Glycosciences. In *Glycoinformatics*; Lütke, T.; Frank, M., Eds.; Methods in Molecular Biology, Vol. 1273; Springer Science + Business Media: New York, 2015; pp 241–258. <http://glyco3d.cermav.cnrs.fr>
46. Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467–479. doi:10.1016/S0092-8674(00)00138-0
47. Hamelryck, T. W.; Loris, R.; Bouckaert, J.; Dao-Thi, M.-H.; Strecker, G.; Imberty, A.; Fernandez, E.; Wyns, L.; Etzler, M. E. *J. Mol. Biol.* **1999**, *286*, 1161–1177. doi:10.1006/jmbi.1998.2534
48. Šulák, O.; Cioci, G.; Delia, M.; Lahmann, M.; Varrot, A.; Imberty, A.; Wimmerová, M. *Structure* **2010**, *18*, 59–72. doi:10.1016/j.str.2009.10.021
49. Blanchard, B.; Nurisso, A.; Hollville, E.; Tétaud, C.; Wiels, J.; Pokorná, M.; Wimmerová, M.; Varrot, A.; Imberty, A. *J. Mol. Biol.* **2008**, *383*, 837–853. doi:10.1016/j.jmb.2008.08.028
50. Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. *J. Mol. Biol.* **1998**, *282*, 1043–1059. doi:10.1006/jmbi.1998.2076
51. Kumar, A.; Sýkorová, P.; Demo, G.; Dobeš, P.; Hryšl, P.; Wimmerová, M. *J. Biol. Chem.* **2016**, *291*, 25032–25049. doi:10.1074/jbc.M115.693473
52. Hester, G.; Kaku, H.; Goldstein, I. J.; Schubert Wright, C. *Nat. Struct. Biol.* **1995**, *2*, 472–479. doi:10.1038/nsb0695-472
53. Sartim, M. A.; Pinheiro, M. P.; de Pádua, R. A.; Sampaio, S. V.; Nonato, M. C. *Toxicol.* **2017**, *126*, 59–69. doi:10.1016/j.toxicol.2016.12.007
54. Buts, L.; Loris, R.; De Genst, E.; Oscarson, S.; Lahmann, M.; Messens, J.; Brosens, E.; Wyns, L.; De Greve, H.; Bouckaert, J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *D59*, 1012–1015. doi:10.1107/S0907444903007170
55. Kostlánová, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerová, M.; Imberty, A. *J. Biol. Chem.* **2005**, *280*, 27839–27849. doi:10.1074/jbc.M505184200
56. Gallego del Sol, F.; Gómez, J.; Hoos, S.; Nagano, C. S.; Cavada, B. S.; England, P.; Calvete, J. J. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2005**, *F61*, 326–331. doi:10.1107/S1744309105004835
57. Cioci, G.; Mitchell, E. P.; Chazalet, V.; Debray, H.; Oscarson, S.; Lahmann, M.; Gautier, C.; Breton, C.; Pérez, S.; Imberty, A. *J. Mol. Biol.* **2006**, *357*, 1575–1591. doi:10.1016/j.jmb.2006.01.066

58. Imberty, A.; Lortat-Jacob, H.; Pérez, S. *Carbohydr. Res.* **2007**, *342*, 430–439. doi:10.1016/j.carres.2006.12.019
59. Li, W.; Johnson, D. J.; Esmon, C. T.; Huntington, J. A. *Nat. Struct. Mol. Biol.* **2004**, *11*, 857–862. doi:10.1038/nsmb811
60. Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S. *Science* **2003**, *301*, 610–615. doi:10.1126/science.1088196
61. Deng, D.; Xu, C.; Sun, P.; Wu, J.; Yan, C.; Hu, M.; Yan, N. *Nature* **2014**, *510*, 121–125. doi:10.1038/nature13306
62. Bianchi, L.; Diez-Sampedro, A. *PLoS One* **2010**, *5*, e10241. doi:10.1371/journal.pone.0010241
63. Cao, Y.; Jin, X.; Levin, E. J.; Huang, H.; Zong, Y.; Quick, M.; Weng, J.; Pan, Y.; Love, J.; Punta, M.; Rost, B.; Hendrickson, W. A.; Javitch, J. A.; Rajashankar, K. R.; Zhou, M. *Nature* **2011**, *473*, 50–54. doi:10.1038/nature09939
64. Dang, S.; Sun, L.; Huang, Y.; Lu, F.; Liu, Y.; Gong, H.; Wang, J.; Yan, N. *Nature* **2010**, *467*, 734–738. doi:10.1038/nature09406
65. Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J. *Science* **2008**, *321*, 810–814. doi:10.1126/science.1160406
66. Kumar, H.; Kasho, V.; Smirnova, I.; Finer-Moore, J. S.; Kaback, H. R.; Stroud, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 1784–1788. doi:10.1073/pnas.1324141111
67. McCoy, J. G.; Ren, Z.; Stanevich, V.; Lee, J.; Mitra, S.; Levin, E. J.; Poget, S.; Quick, M.; Im, W.; Zhou, M. *Structure* **2016**, *24*, 956–964. doi:10.1016/j.str.2016.04.003
68. Oldham, M. L.; Chen, J. *Science* **2011**, *332*, 1202–1205. doi:10.1126/science.1200767
69. Sun, L.; Zeng, X.; Yan, C.; Sun, X.; Gong, X.; Rao, Y.; Yan, N. *Nature* **2012**, *490*, 361–366. doi:10.1038/nature11524
70. Wang, J.; Yan, C.; Li, Y.; Hirata, K.; Yamamoto, M.; Yan, N.; Hu, Q. *Cell Res.* **2014**, *24*, 1486–1489. doi:10.1038/cr.2014.144
71. Watanabe, A.; Choe, S.; Chaptal, V.; Rosenberg, J. M.; Wright, E. M.; Grabe, M.; Abramson, J. *Nature* **2010**, *468*, 988–991. doi:10.1038/nature09580
72. Xu, Y.; Tao, Y.; Cheung, L. S.; Fan, C.; Chen, L. Q.; Xu, S.; Perry, K.; Frommer, W. B.; Feng, L. *Nature* **2014**, *515*, 448–452. doi:10.1038/nature13670
73. Nomura, N.; Verdon, G.; Kang, H. J.; Shimamura, T.; Nomura, Y.; Sonoda, Y.; Hussien, S. A.; Qureshi, A. A.; Coincon, M.; Sato, Y.; Abe, H.; Nakada-Nakura, Y.; Hino, T.; Arakawa, T.; Kusano-Arai, O.; Iwanari, H.; Murata, T.; Kobayashi, T.; Hamakubo, T.; Kasahara, M.; Iwata, S.; Drew, D. *Nature* **2015**, *526*, 397–401. doi:10.1038/nature14909
74. Bourgeois, D.; Weik, M. *Crystallogr. Rev.* **2009**, *15*, 87–118. doi:10.1080/08893110802604868
75. Batot, G. Une nouvelle approche pour étudier le mécanisme des glycosyltransférases. Ph.D. Thesis, Université de Grenoble, Grenoble, France, 2013.
76. Jørgensen, R.; Batot, G.; Mannerstedt, K.; Imberty, A.; Breton, C.; Hindsgaul, O.; Royant, A.; Palčić, M. M. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2014**, *F70*, 1015–1021. doi:10.1107/S2053230X1401259X
77. Menchicchi, B.; Fuenzalida, J. P.; Hensel, A.; Swamy, M. J.; David, L.; Rochas, C.; Goycoolea, F. M. *Biomacromolecules* **2015**, *16*, 924–935. doi:10.1021/bm501832y
78. Pernot, P.; Round, A.; Barrett, R.; De Maria Antolinos, A.; Gobbo, A.; Gordon, E.; Huet, J.; Kieffer, J.; Lentini, M.; Mattenet, M.; Morawe, C.; Mueller-Dieckmann, C.; Ohlsson, S.; Schmid, W.; Surr, J.; Theveneau, P.; Zerrad, L.; McSweeney, S. *J. Synchrotron Radiat.* **2013**, *20*, 660–664. doi:10.1107/S0909049513010431
79. Round, A.; Felisaz, F.; Fodinger, L.; Gobbo, A.; Huet, J.; Villard, C.; Blanchet, C. E.; Pernot, P.; McSweeney, S.; Roessle, M.; Svergun, D. I.; Cipriani, F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *D71*, 67–75. doi:10.1107/S1399004714026959
80. Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. *Protein Pept. Lett.* **2016**, *23*, 217–231. doi:10.2174/0929866523666160106153655
81. De Maria Antolinos, A.; Pernot, P.; Brennich, M. E.; Kieffer, J.; Bowler, M. W.; Delageniere, S.; Ohlsson, S.; Malbet Monaco, S.; Ashton, A.; Franke, D.; Svergun, D.; McSweeney, S.; Gordon, E.; Round, A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *D71*, 76–85. doi:10.1107/S1399004714019609
82. Daillant, J.; Gibaud, A. *X-ray and Neutron Reflectivity: Principles and Applications*; Springer, 1999.
83. Rondelli, V.; Brocca, P.; Fragneto, G.; Daillant, J.; Tringali, C.; Cantu, C.; Del Favero, E. *Biochim. Biophys. Acta, Biomembr.* **2017**, *1859*, 845–851. doi:10.1016/j.bbamem.2017.01.012
84. Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306. doi:10.1021/ja037055w
85. Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082. doi:10.1021/ja0257319
86. Imberty, A.; Chanzy, H.; Pérez, S.; Buléon, A.; Tran, V. *J. Mol. Biol.* **1988**, *201*, 365–378. doi:10.1016/0022-2836(88)90144-1
87. Imberty, A.; Pérez, S. *Biopolymers* **1988**, *27*, 1205–1221. doi:10.1002/bip.360270803
88. Popov, D.; Buléon, A.; Burghammer, M.; Chanzy, H.; Montesanti, N.; Putaux, J.-L.; Potocki-Véronèse, G.; Riekkel, C. *Macromolecules* **2009**, *42*, 1167–1174. doi:10.1021/ma801789j
89. Lichtenegger, H.; Müller, M.; Paris, O.; Riekkel, C.; Fratzl, P. *J. Appl. Crystallogr.* **1999**, *32*, 1127–1133. doi:10.1107/S0021889899010961
90. Hsiao, B. S. Structure characterization of cellulose nanofibers-microfibrils. *ALBA Synchrotron - Marie Curie Meeting Room*; Barcelona, 2015.
91. Pérez, S.; Burghammer, M. *Actual. Chim.* **2011**, *356–357*, 68–72.
92. Chanzy, H.; Putaux, J.-L.; Dupeyre, D.; Davies, R.; Burghammer, M.; Montanari, S.; Riekkel, C. *J. Struct. Biol.* **2006**, *154*, 100–110. doi:10.1016/j.jsb.2005.11.007

## License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (<http://www.beilstein-journals.org/bjoc>)

The definitive version of this article is the electronic one which can be found at: [doi:10.3762/bjoc.13.114](https://doi.org/10.3762/bjoc.13.114)