THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Présentée par
Lukáš GAJDOŠ

Thèse dirigée par **Anne IMBERTY, Centre de recherches sur les macromolécules végétales**, et
codirigée par **Juliette DEVOS, Institut Laue-Langevin**
préparée au sein de **l'Institut Laue-Langevin**
dans **l'École Doctorale Chimie et Sciences du Vivant**

Neutron diffraction for deciphering lectin-glycan interactions involved in bacterial infection

Diffraction des neutrons pour l'étude des interactions protéine-sucre impliquées dans les infections bactériennes

Thèse soutenue publiquement le **20 Mai 2021**,
devant le jury composé de :
Monsieur, John HELLIWELL
Rapporteur, The University of Manchester, United Kingdom
Monsieur, Derek LOGAN
Rapporteur, Lund University, Sweden
Madame, Michaela WIMMEROVÁ
Examinateur, Masaryk University, Czech Republic
Madame, Andrea DESSEN
Présidente, Institut de Biologie Structurale, France

Venujem rodičom

Acknowledgements

First of all, I would like to thank all my PhD supervisors Dr. Anne Imberty, Dr. Juliette Devos, Prof. Trevor Forsyth, Dr. Matthew Blakeley and Dr. Michael Haertlein for their guidance and support throughout all these years. Thank you Anne for having me as an Erasmus student and trying to convince me to apply for the PhD. I am so glad I did and that you fought for that position till the end. You have been always supportive and carrying. Asking the right questions and giving the best advice. Thank you for being always there, reminding me deadlines and pushing me when I needed it. A huge thank you Juliette for everything you taught me in the lab, for being always cheerful, enthusiastic, and willing to help every time I needed. I could not have done it without you. You are doing an invaluable work and deserves the best. Thank you for all the fun in the lab, discovering new music, for the concerts, flying in the air, all the chats and laughs, drinks and travels to conferences.

Thank you Matthew for your help with neutrons, for your encouragement, optimism and trust. For being always supportive and being there when I needed. It has been a pleasure working with you.

Thank you Trevor for being always supportive, cheerful, making sure I have everything I need.

Thank you Michael for your thoughtful comments and suggestions during our meetings. I could not have asked for better supervisors who made everything so much easier.

I would like to thank Dr. Atul Kumar who was there in the very beginning when I first arrived to Grenoble. Thank you Atul for your curries, for your dinners and laughs with you and Deepti.

I also want to thank my former supervisor Prof. Michaela Wimmerová who introduced me to the world of lectins and found a unique opportunity for me to go to Grenoble that completely changed the course of my life.

Abstract

Lectins are ubiquitous proteins that can reversibly and specifically bind to carbohydrates. Carbohydrate-mediated recognition and adhesion are key events in the early steps of the interaction of pathogenic or symbiotic bacteria with a host. For example, *Pseudomonas aeruginosa* is an opportunistic bacterium that causes lethal lung infection in cystic fibrosis patients. *Photorhabdus laumondii* lives in a symbiotic relationship with nematodes and is involved in insecticidal activities of these worms. Both bacteria produce several lectins, i.e. glycan receptors, which are specific for the host glycoconjugates.

The aim of this PhD thesis was to use neutron protein crystallography to provide insights into lectin-carbohydrate interactions involved in bacterial infection. Fucose-specific lectins PLL and LecB from *Photorhabdus laumondii* and *Pseudomonas aeruginosa*, respectively, were produced in perdeuterated forms and crystallized with a ligand. A unique part of the thesis was the *in vivo* production of a perdeuterated monosaccharide, L-fucose, using genetically modified strains of *E. coli* bacteria designed by Dr. Eric Samain at CERMAV institute in Grenoble. Perdeuterated fucose has been successfully produced, purified and biophysically characterized. It was used as a ligand for crystallization trials of both PLL and LecB lectins.

Three neutron structures have been solved in this thesis. Two structures of PLL lectin from *P. laumondii* both in apo and ligand-bound form and one structure of LecB lectin from *P. aeruginosa* in the complex with perdeuterated fucose.

The results obtained in this thesis yielded one published publication, one submitted and one in preparation, which are all part of this manuscript.

Resumé

Les lectines sont des protéines omniprésentes qui peuvent lier les glucides de manière spécifique et réversible. La reconnaissance et l'adhésion à médiation par les hydrates de carbone sont des événements clés dans les premières étapes de l'interaction des bactéries pathogènes ou symbiotiques avec un hôte. Par exemple, *Pseudomonas aeruginosa* est une bactérie opportuniste qui provoque une infection pulmonaire mortelle chez les patients atteints de mucoviscidose. *Photorhabdus laumondii* vit en relation symbiotique avec les nématodes et participe aux activités insecticides de ces vers. Les deux bactéries produisent plusieurs lectines, c'est-à-dire des récepteurs de glycanes, qui sont spécifiques aux glycoconjugués de l'hôte.

Le but de cette thèse de doctorat était d'utiliser la cristallographie des protéines neutroniques pour fournir des informations sur les interactions entre lectines et glucides. Les lectines PLL et LecB spécifiques au fucus de *Photorhabdus laumondii* et *Pseudomonas aeruginosa*, respectivement, ont été choisies comme objet de notre étude. Une partie unique de la thèse a été la production in vivo d'un monosaccharide perdeutéré, le L-fucose, en utilisant des souches génétiquement modifiées de bactéries *E. coli* conçues par le Dr Eric Samain à l'institut CERMAV de Grenoble. Le fucose perdeutéré a été produit, purifié et caractérisé biophysiquement avec succès. Il a été utilisé comme ligand pour les essais de cristallisation des lectines PLL et LecB.

Trois structures neutroniques ont été résolues dans le cadre de cette thèse. Deux structures de la lectine PLL de *P. laumondii*, à la fois sous forme apo et sous forme de ligand, et une structure de la lectine LecB de *P. aeruginosa* dans le complexe avec le fucose perdeutéré.

Les résultats obtenus dans cette thèse ont donné lieu à une publication parue, une soumise et une en préparation qui font tous partie de ce manuscrit.

Table of content

List of figures

List of tables

List of abbreviations

ADP	Adenosine diphosphate
AIDS	Acquired immunodeficiency syndrome
AMP	Antimicrobial peptides
CoA	Coenzyme A
CBD	Carbohydrate binding domain
CEITEC	Central European Institute of Technology
CERMAV	Centre de Recherches sur les Macromolécules Végétales
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance receptor
ConA	Concanavalin A lectin
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
IPTG	Isopropyl β-D-1-thiogalactopyranoside
eDNA	Extracellular DNA
EPN	Entomopathogenic nematodes
EPS	Exopolysaccharides
ESI	Electrospray ionization
ESRF	European Synchrotron Radiation Facility
Fuc	Fucose
Fuc-d$_{12}$	Perdeuterated fucose
GDP	Guanosine diphosphate
HA	Influenza virus hemagglutinin
HCDC	High cell-density cultivation
HIV	Human immunodeficiency virus
HTX	High Throughput Crystallization
IBS	Institut de Biologie Structurale

IJ	Infective juvenile
ILL	Institut Laue-Langevin
IPS	3,5-dihydroxy-4-isopropylstilbene
ITC	Isothermal titration calorimetry
LB	Lysogeny Broth
LPS	Lipopolysaccharides
MBL	Mannose-binding lectin
Mbp	Megabase pair
MDR	Multidrug resistance
MS	Mass spectrometry
NADP	Nicotineamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
NMX	Neutron macromolecular crystallography
NRA	Nuclear reaction analysis
OD_{600}	Optical density at 600 nm
PAMP	Pathogen associated molecular pattern
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PO	Phenolooxidase
PPO	Prophenoloxidase
QS	Quorum sensing
RMSD	Root-mean-square-deviation
Rpm	Revolution per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TLC	Thin layer chromatography
UTI	Urinary tract infection

X M X   X-ray macromolecular crystallography

1
Introduction

1.1 Lectins

Lectins are carbohydrate-binding proteins that were discovered in 1888 by Peter Herman Stillmark who observed the agglutination of human red blood cells by extracts of the castor bean seeds (Stillmark, 1888). Formerly known as hemagglutinins or phytohemagglutinins, the term lectin (derived from Latin *legere*, to select) was introduced in 1954 to point out their ability to selectively bind and agglutinate different types of human ABO red blood cells (Boyd and Shapleigh, 1954). The agglutination activity of lectins comes from the fact that lectins are usually multivalent proteins with several carbohydrate-binding sites per subunit which allows crosslinking between cells by interacting with glycans on their surface (Lis and Sharon, 1998).

Lectins are thus described as carbohydrate-binding proteins that are of non-immune origin, that lack enzymatic activity and that can agglutinate cells (Goldstein et al., 1980). The non-immune origin indicates that they are not a product of an immune response such as carbohydrate-specific antibodies. Opposite to carbohydrate-processing enzymes that can also selectively bind carbohydrates, such as glycosidases and glycosyltransferases, lectins cannot modify the carbohydrate molecule upon binding.

Lectins are ubiquitous proteins and can be found in all living systems including viruses, bacteria, fungi, plants, and animals. They represent a heterogeneous group of proteins that differ in size, oligomeric arrangement, structure, carbohydrate-specificity as well as the composition of their binding sites. Lectins bind mono- and oligosaccharides reversibly and with high specificity. The binding selectivity comes from a balance between several types of interactions: i) specific hydrogen bonds and/or metal coordination are established between the sugar hydroxyl groups that can act

both as donors and acceptors of hydrogen bonds and the amino acids of the protein, ii) apolar face of sugars establishes π-stacking contact with side groups of aromatic amino acids. The affinity for a monosaccharide is usually low with dissociation constants in the millimolar range (Lis and Sharon, 1998). However, lectins from pathogens often exhibit much higher affinities towards di-, tri-, and tetrasaccharides that are most likely their natural ligands and that are present on the surface of eukaryotic cells. Based on their specificity to monosaccharides, lectins could first be divided into five groups with the highest affinity to monosaccharide: D-mannose, D-galactose/*N*-acetyl-D-galactosamine, *N*-acetyl-D-glucosamine, L-fucose, and *N*-acetylneuraminic acid (Lis and Sharon, 1998). These monosaccharides are typical constituents of glycans of eukaryotic cells. Nowadays, this classification based on sugar specificity is becoming obsolete because of the difference in fine specificities of lectins as well as with increasing number of lectins that show low affinity towards simple monosaccharides.

Lectins are usually multivalent proteins with the presence of several binding sites originating from a repetition of tandem sequences from oligomerization. The apparent low affinity of single binding interactions is enhanced by the overall binding avidity effect of multivalent interactions (Lis and Sharon, 1998).

Lectins are widely used in biotechnology thanks to their stability, availability and high specificity. Immobilized lectins are used for the detection and purification of glycoproteins and glycolipids using affinity chromatography (Freeze, 1995). They are used in histo- and cytochemistry with their ability to detect sugars in tissues and cells. Lectins are also investigated for their therapeutical prospects in cancer treatment since they are able to induce apoptosis and autophagy of cancer cells (Yau et al., 2015). Recently, lectins have shown potential in cancer diagnostics as biomarkers that can detect glycosylation modifications on malignant cells. For example, *Lens culinaris* agglutinin from lentil is specific to α-1,6-fucosylated glycoconjugates and could be used as a diagnostic marker in thyroid cancer, prostate cancer and hepatocellular carcinoma (Kanai et al., 2009; Zhou et al., 2017). A GlcNAc-binding fungal lectin from *Psathyrella*

velutina (PVL) can specifically bind to cancer cells in lung, breast and colon carcinomas and can be used as a labelling tool (Audfray et al., 2015). Another fungal lectin PhoSL from *Pholiota squarrosa* has been identified to exclusively recognize α-1,6-fucosylated *N*-glycans that are associated with cancer such as liver and melanoma cancers and could be used in glycoprofiling and detecting of core fucosylation in proteins (Cabanettes et al., 2018).

1.1.1   Lectins to decipher the "glycocode"

Lectins are able to "read" so-called "glycocode". The code originates from the complexity of glycans that are the most diverse and abundant group of biomolecules on earth due to the many regio- and stereochemistries how simple monosaccharides combine into branched oligo- and polysaccharides. While three different amino acids can generate only 6 different tripeptides, three different hexoses (monosaccharides with six carbon atoms) could theoretically produce up to 27 648 unique trisaccharides (Laine, 1994). Of course, not all of these possibilities are present in nature but it is noteworthy to say that the Carbohydrate Structure Database currently contains almost 25 000 different natural glycan compounds from prokaryotic, plant and fungal origin (http://csdb.glycoscience.ru/database/) (Egorova and Toukach, 2018). The term glycan refers to the carbohydrate part of glycoproteins or glycolipids that are typically located at the cell surface, on secreted mucin proteins and in the extracellular matrix (Varki and Kornfeld, 2017). Each eukaryotic cell is covered with glycans referred to as glycocalyx. Glycocalyx represents the first information about the cell to the outside environment. It has many functions, it acts as a protective barrier against microorganisms, regulates membrane protein diffusion and is involved in cell-signaling and cell differentiation, among others (Möckl, 2020). Bacteria also produce various polysaccharides, for example, lipopolysaccharides (LPS) are a part of the outer membrane of Gram-negative bacteria that provide structural integrity, membrane protection, serve in surface adhesion and induce a strong immune response in animals (Raetz and Whitfield, 2002). The complexity of polysaccharides can also be seen in

plants with a large heterogeneity of plant cell wall polysaccharides that contain more than 15 different monosaccharides linked by various glycosidic bonds (Burton et al., 2010).

Lectins have the unique ability to specifically recognize and bind different glycans which promotes their biological activity. Since glycans lack the diversity of functional groups that are present on other biomolecules, the specific binding is achieved by stereospecific complementarity between the glycan hydroxyl groups and the amino acids in the lectin binding sites. Lectins adopt a limited number of folds, predominantly β-prism, β-trefoil, β-sandwich, or β-propeller (Figure 1) (Bonnardel et al., 2019; Imberty and H. Prestegard, 2015).
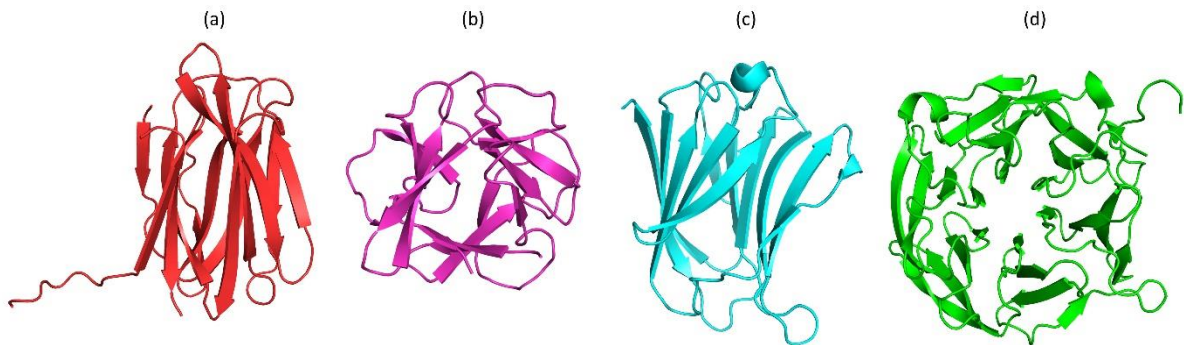
Figure 1: Examples of lectin folds. (a) β-prism fold of Jacalin lectin (PDB: 1KU8). (b) β-trefoil lectin from *Entamoeba histolytica* (PDB: 6IFA). (c) β-sandwich fold of Galectin-3C (PDB: 3AYE). (d) β-propeller fold of *Aleuria aurantia* lectin (PDB: 1OFZ).

The common structural features may be used in classification of lectins. For example, simple lectins consist of a small number of subunits with each containing a carbohydrate-binding site. To this class of lectins belong practically all plant lectins and galectins from animals. Mosaic lectins consist of several types of binding domains, out of which only one possesses a carbohydrate-binding site. This class includes C-, P- and I-type animal lectins and viral hemagglutinins. Finally, macromolecular assemblies can be found in bacteria as filamentous organelles arranged in defined order such as

pili or fimbriae. Each class can be subdivided into structural families with similar sequences and properties (Ambrosi et al., 2005).

A new classification of lectins has been proposed recently (Bonnardel et al., 2021). It is based on three levels: i) the fold level accounting for similarities in lectin domains (β-propellers, β-sandwich *etc.*); ii) the class level characterized by sequence similarities with 20 % cut-off between different classes and iii) the family level that groups lectins with more than 70 % sequence identity.

1.1.1.1   Human glycome

Human glycome is composed of a large repertoire of glycoconjugates which are carbohydrates moieties covalently linked to proteins and lipids and has a huge biological importance (Reily et al., 2019). Unlike proteins, the glycome is produced in a non-template manner. The glycosylation of protein and lipid molecules is catalysed by glycosytransferases and glycosidases and occurs in the endoplasmic reticulum and Golgi apparatus. *N*-glycans are covalently attached to asparagine side chains of proteins, while *O*-glycans are linked to side chains of serine and threonine (Figure 2). Glycosphingolipids, one of the most abundant glycolipids in humans, are decorated by a glycan attached at the C1 hydroxyl position of a ceramide. Proteoglycans are heavily glycosylated proteins with long linear sugar repeats, called glycosaminoglycans, attached via *O*-glycosylation and are present mostly in the extracellular matrix.
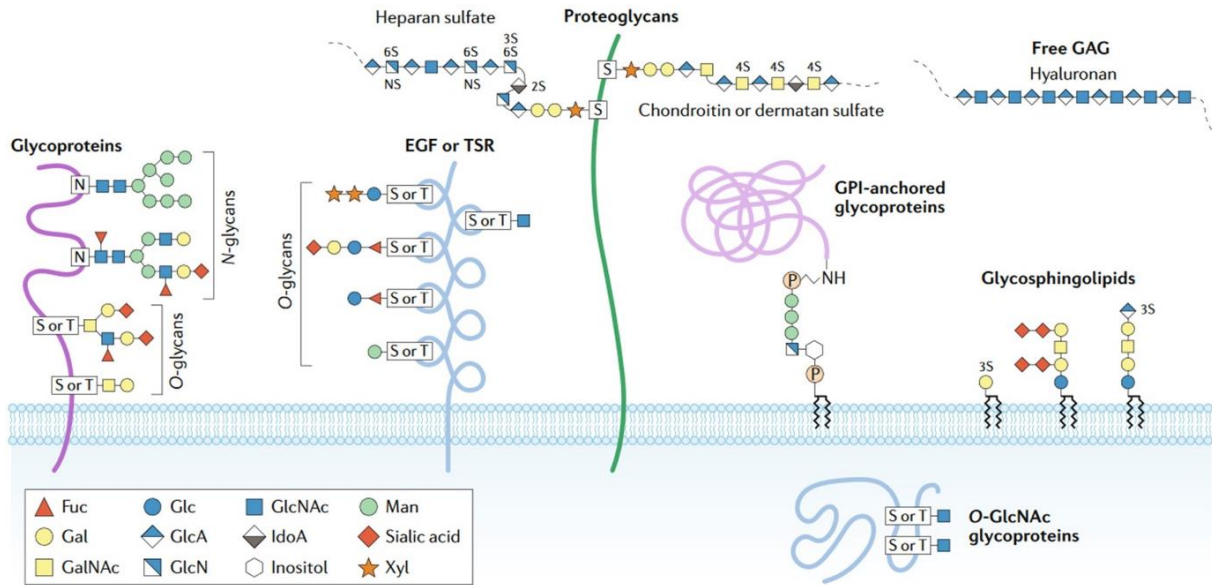
Figure 2: Major types of glycosylations in human proteins and lipids. Glycans are covalently attached to proteins and lipids to form various glycoconjugates. Figure is taken from (Reily et al., 2019).

More than 130 diseases are caused by congenital disorders of glycosylation characterized by defects in glycan modification pathways often resulting in severe muscular, developmental and neurological malfunctions (Chang et al., 2018). Glycans play important roles in immunity and inflammation. Human lectins including galectins, Siglecs and selectins recognize different glycans and mediate cell responses, including apoptosis, cell adhesion or cell trafficking. Another example of a specific role of glycans is the ABO blood group system, which is based on a presence or absence of A, B and H antigens that are fucosylated glycans present on membrane glycoproteins and glycolipids on the surface of erythrocytes, epithelial and endothelial cells.

L-Fucose is an unusual 6-deoxy monosaccharide commonly found in *N-* and *O-*linked glycans in humans (Becker and Lowe, 2003). It often exists as a terminal modification of glycoconjugates that have various biological functions. Another type of human blood group system with fucosylated antigens is Lewis antigen system with Lewis and Sialyl Lewis antigens (Figure 3).
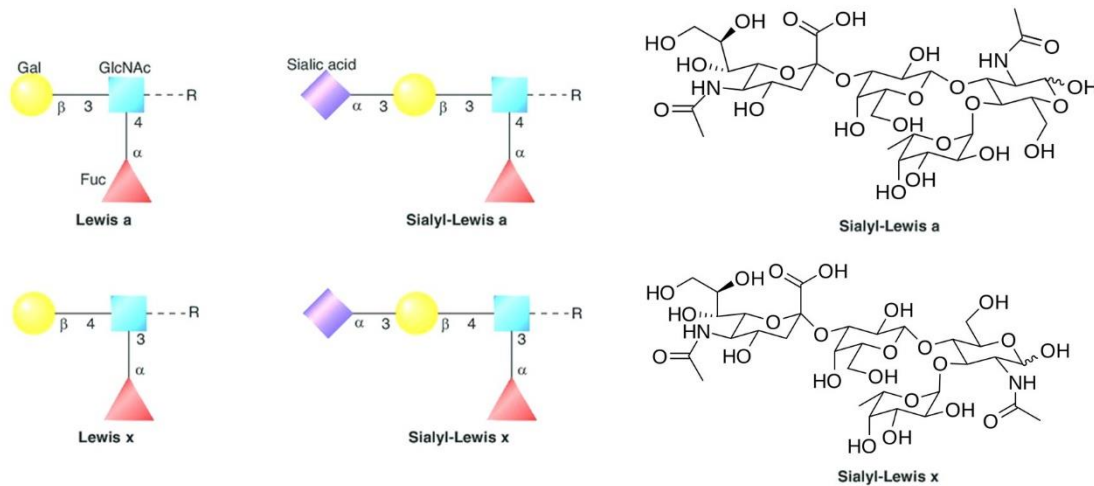
Figure 3: Schematic and chemical structures of Lewis a and Lewis x antigens. Figure is adapted from (Magalhes et al., 2010).

Their overexpression is often connected to breast, colorectal or hepatocellular cancer and can be associated with decreased survival, metastasis or cancer progression (Blanas et al., 2018). Fucosylated antigens have important roles in blood transfusion, contribute to selectin-dependent leukocyte adhesion (Vestweber and Blanks, 1999) and promote cell adhesion during embryogenesis (Bird and Kimber, 1984). Fucosylated blood group antigens are also important for host-microbe interactions, for example, they serve as binding epitopes for gastric pathogen *Helicobacter pylori* that binds Lewis b antigen containing α(1,2)- and α(1,4)-linked fucose (Hooper and Gordon, 2001).

1.1.1.2   Plant lectins

Plant lectins are a large group of lectins sharing common structural similarities with different carbohydrate specificities. They have been studied for a long time and yet their functions are still under debate. Some are involved in the symbiosis between nitrogen-fixing bacteria mainly from the genus *Rhizobium* and leguminous plants (Bohlool and Schmidt, 1974). A more general theory suggests that plant lectins act as defense agents against various phytopathogenic and predatory organisms such as fungi, insects as well as other vertebrates and invertebrates (De Hoff et al., 2009).

Legume lectins are the largest family of plant lectins also known as L-type lectins, more than 100 members of this family have already been characterized (Loris et al., 1998). They all require Ca^{2+} ion and a transitional metal ion (usually Mn^{2+}) for the carbohydrate-binding activity. They are generally oligomeric proteins composed of two to four subunits of about 250 amino acids each, with one binding site per subunit. All legume lectins exhibit high sequence homology with about 20 % of invariant amino acids and close to 20 % of similar ones (Ambrosi et al., 2005). The conserved amino acids are involved in the carbohydrate-binding sites and in metal-coordination. Amongst legume lectins, concanavalin A (ConA) from the seeds of *Canavalia ensiformis* (jack bean) occupies a special place as it was the first lectin whose primary sequence and three-dimensional structure were determined (Figure 4a). Concanavalin A is specific for mannose and glucose and the interesting feature is its so-called "circular homology". For ConA, this homology is obtained by aligning the residue 119 (Ser) with the N-terminus of other legume lectins proceeding to the C-terminus of ConA and continuing to the N-terminus along the sequence of other legume lectins.

Another important family of plant lectins are the ribosome-inactivating protein type II lectins (RIP-II) with prototypical lectin ricin (also called RCA-II) from the poisonous plant *Ricinus communis* (castor bean) that was the first lectin to be discovered (Stillmark, 1888). RIP-II lectins belong to the R-type lectin superfamily composed of members from plants, animals and bacteria with a conserved ricin-like carbohydrate-binding domain (CBD). Ricin is one of the most poisonous toxins with a lethal dose (LD_{50}) of 3-5 μg per kilogram of bodyweight. Ricin is composed of two chains A and B linked by a disulphide bridge (Figure 4b). Chain A is an *N*-glycoside hydrolase that inactivates the 60S ribosomal subunit and chain B binds carbohydrates. Chain B contains two CBDs with a β-trefoil structure that can bind divalently β-galactosides or *N*-acetylgalactosamines present on the cell surface. Ricin is then transported by endosomes to the endoplasmic reticulum where the two chains are separated by reduction of the disulphide bridge. Chain A is released into the cytosol

where it cleaves one adenine residue in 60S RNA resulting in the inability of ribosomes to promote protein biosynthesis (Endo and Tsurugi, 1987).
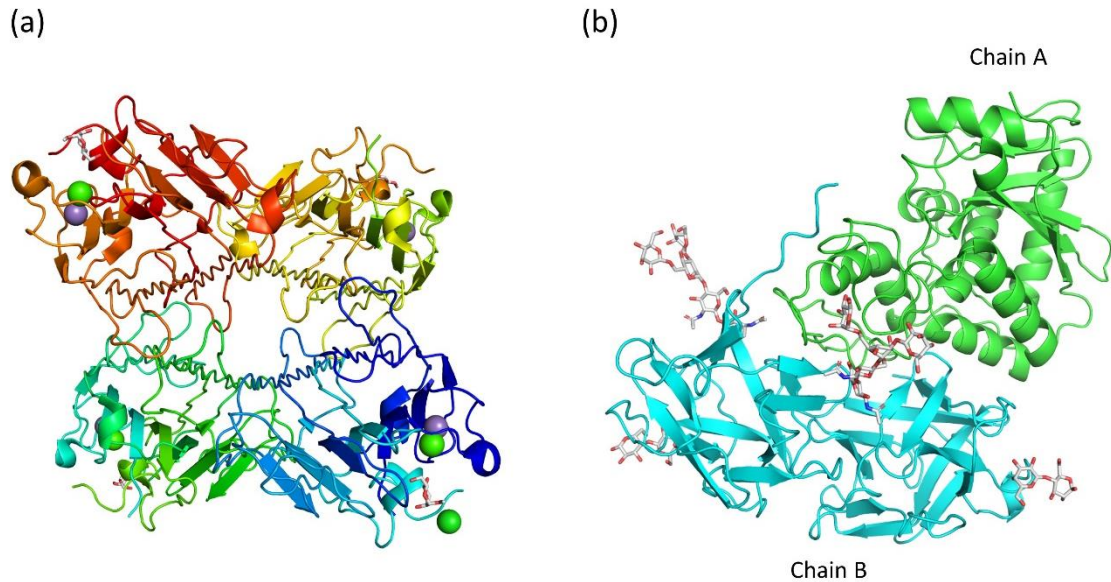
Figure 4: Cartoon representation of plant lectins. Carbohydrates are depicted as grey sticks, Ca^{2+} and Mn^{2+} ions as light green and violet spheres, respectively. (a) Concanavalin A in complex with α-methyl-D-mannopyranoside (PDB: 5CNA). (b) Ricin in complex with lactose (PDB: 2AAI).

1.1.1.3   Animal lectins

Animal lectins have been isolated from both invertebrates and vertebrates. In vertebrates lectins are either present as an integral part of cell membranes or in soluble forms. The functions of animal lectins are extremely diverse from cell-adhesion, clearance of glycoproteins, apoptosis processes, innate immune response, fertilization, cell proliferation, leukocyte trafficking, cancer cell metastasis, to pathogen clearance, *etc* (Gabius, 1997; Kilpatrick, 2002). They are divided into several families based on common structural similarities. Currently (March 2021) there are 149 unique animal lectins structurally characterized with 687 X-ray crystal structures (Bonnardel et al., 2019). Animal lectins can be classified into five families based on the common structural and binding features (Gabius, 1997).

C-type lectins represent a superfamily of calcium-dependent lectins that use one calcium atom to coordinate two hydroxyl groups of the carbohydrate ligand. They are characterized by an extracellular carbohydrate-binding domain consisting of 115-130 amino acids, of which 14 are invariant and 18 highly conserved (Drickamer, 1999). C-type lectins are grouped into three major subgroups based on the common overall architecture: endocytic lectins, collectins and selectins (Drickamer and Taylor, 2015). Endocytic lectins are transmembrane glycoproteins that function as receptors inducing endocytosis. Collectins (collagen-containing C-type lectins) play roles in innate immunity. For example, the binding of serum mannose-binding lectin to microorganisms activates the lectin pathway of the complement system (Russell et al., 2005). Selectins (E-, L- and P-type) is a family of transmembrane lectins with an extracellular CBD in the N-terminal region. E-selectin is a glycoprotein expressed by endothelial cells activated by inflammatory cytokines. All selectins bind sialylated carbohydrates such as Sialyl Lewis x and Sialyl Lewis a tetrasaccharides present on certain leukocytes. E-selectin (Figure 5a) functions in modulating inflammatory response as it mediates leukocyte-rolling on the endothelium and recruits neutrophils, monocytes, T-lymphocytes and natural killer cells to the site of inflammation (McEver, 2015). L-selectin is found on the surface of leukocytes and is described as a "homing receptor" involved in the attachment of lymphocytes to high endothelial venules of lymph nodes. P-selectin is present on the endothelial cells and platelets and plays a role in the recruitment of white blood cells to the site of injury during inflammation (McEver, 2015).
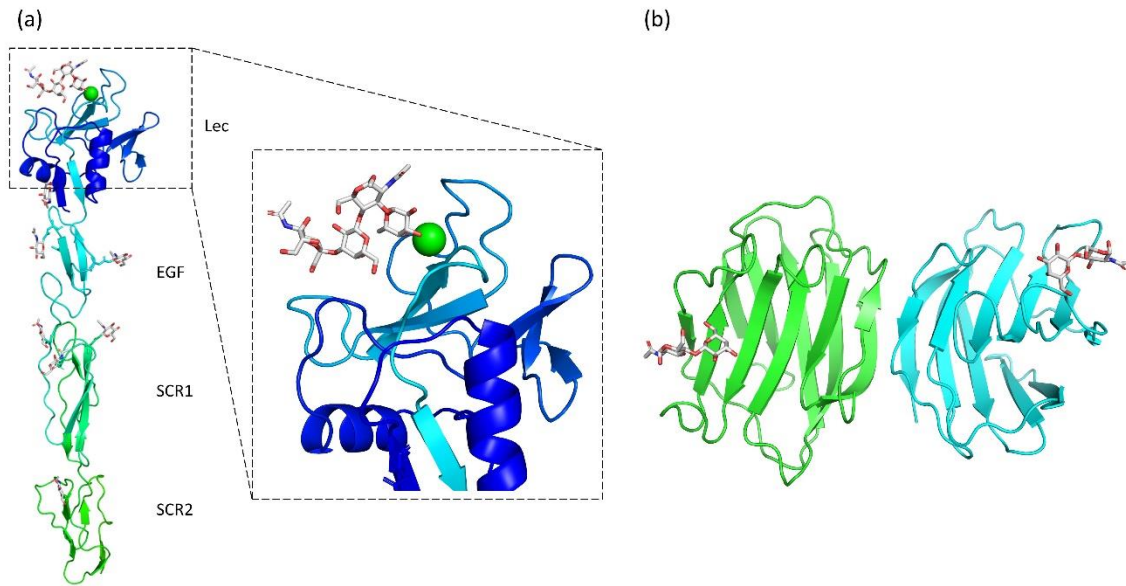
Figure 5: Cartoon representation of selected animal lectins. Carbohydrates are represented as grey sticks, calcium atom as s green sphere. (a) E-selectin with N-terminal Lec domain, the epidermal growth factor-like domain (EGF) and the first two short consensus repeats (SCR2 domains). The Lec domain binds sialylated Lewis x via calcium coordination (PDB: 4CSY). (b) Galectin-7 in complex with *N*-acetyl-D-lactosamine (PDB: 1GAN).

Galectins represent a widespread family of soluble lectins with a characteristic carbohydrate-binding domain composed of ~130 amino acids. Galectins usually bind β-galactoside-containing glycoconjugates (Johannes et al., 2018). There are 15 galectins found in mammals and based on the molecular form of the CBD, galectins can be divided into three classes: (a) the prototype galectins which contain a single CBD are usually non-covalently linked homodimers (Figure 5b), (b) the chimeric galectins, of which galectin-3 is the only known galectin in vertebrates, are characterized by a single CBD and a long non-CBD N-terminal domain rich in proline, glycine and tyrosine residues, and (c) the tandem-repeat galectins containing two CBDs within a single polypeptide (Cummings and Liu, 2009). Galectins can be found in the cytosol, nucleus, and in extracellular matrix or in circulation. The various localization is reflected in their diverse functions. They were first thought to interact only with endogenous "self" glycans, mediate developmental processes and regulate immune homeostasis. Recent studies suggest new roles of host galectins as pathogen-recognition receptors that target

glycans present on the surfaces of viruses and bacteria and which mediate the immune response (Vasta, 2009).

Siglecs (sialic acid-binding immunoglobulin-like lectins) are transmembrane proteins that can bind glycoconjugates containing terminal sialic acid. They are mostly present on the surface of hematopoietic cells and function in cell-cell interactions in adaptive and innate immune response as well as in pathogen recognition leading to altered immune responses (Crocker et al., 2007).

Pentraxins were named based on the cyclic arrangement of their subunits into discs with pentameric radial symmetry. They are serum proteins that bind glycans in a calcium-dependant manner. The best-known members of this family are serum amyloid P component protein (SAP) and C-reactive protein which is an acute-phase protein produced by hepatocytes in response to inflammation (Du Clos, 2013).

The P-type lectins can recognize *N*-glycans containing phosphorylated-mannose residues. These lectins play essential roles in trafficking newly synthesized lysosomal enzymes bearing mannose-6-phosphate moieties into lysosomes (Dahms and Hancock, 2002).

1.1.1.4   Fungal lectins

Fungi represent a rich source of lectins with unique carbohydrate specificities. It has been stated that 80 % of reported fungal lectins arise from mushrooms, 15 % from microfungi and 3 % from yeasts (Singh et al., 2011). Interestingly, only 10 % of mushroom species are known which makes them an enormous unexplored source of new potential lectins (Hassan et al., 2015). Their functions are still not well known but several have been proposed. They might serve as storage proteins and have a role in development and morphogenesis, mediate host recognition and association with algae or cyanobacteria in lichens (Díaz et al., 2011). Similar to plant lectins, they also participate in the defense of fungi as they possess insecticidal, vermicidal and antiviral activities. Lectins from pathogenic fungi are also involved in infectious processes (Varrot et al., 2013).

Some fungal lectins act as pathogens towards animals and plants. Several lectins from microfungi and yeasts have been involved in human infections. For instance, lectins from opportunistic pathogens *Aspergillus fumigatus* and *Candida glabrata* are responsible for nosocomial infections (Figure 6). AFL1 lectin is a fucose-binding lectin from *A. fumigatus* that adopts a 6-bladed β-propeller fold with six binding sites located in between the adjacent blades. AFL1 is present on the surface of conidia and binds to human fucosylated blood group oligosaccharides. It has a strong pro-inflammatory effect on human bronchial cells (Houser et al., 2013). Pathogenic yeast *C. glabrata* produces several Epa adhesins with N-terminal calcium dependant lectin domain that bind terminal galactose and promote adherence to mammalian cells (Zupancic et al., 2008).
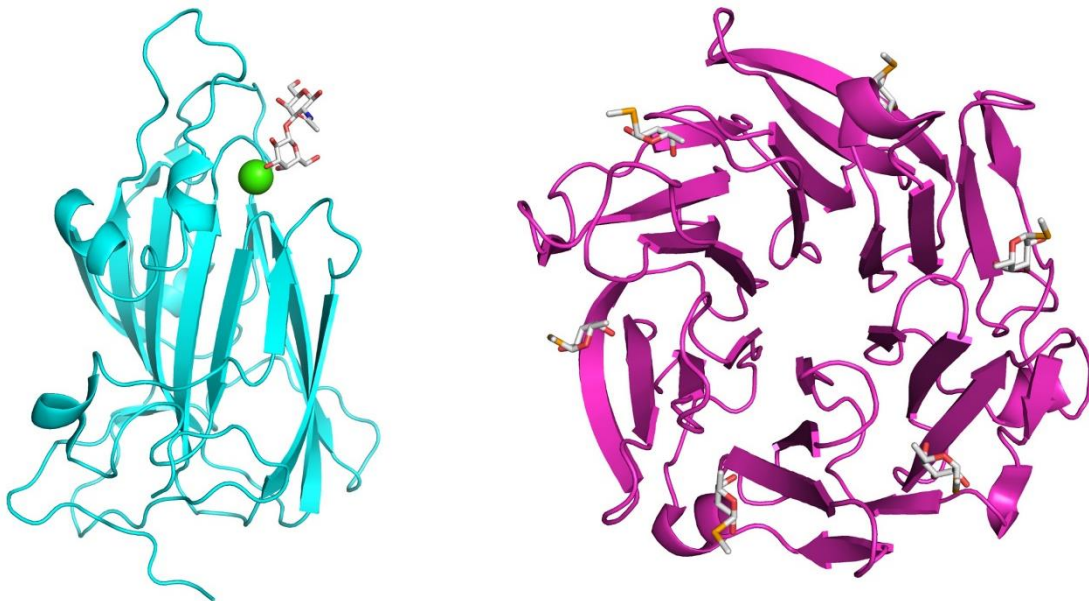
Figure 6: (Left) Epa1 adhesin from yeast *Candida glabrata* in complex with human Galβ1- 3GalNAc mucin epitope (PDB: 4ASL). (Right) AFL1 lectin from *Aspergillus fumigatus* in complex with seleno fucoside (PDB: 4AGI). Carbohydrates are represented as grey sticks and calcium atom as a green sphere.

1.1.1.5   Microbial lectins

Many microorganisms including viruses, bacteria and parasites recognize glycans present on host cells as binding epitopes (Figure 7). They produce sugar-binding

proteins such as lectins that serve as adhesins and toxins and mediate interactions with the host glycoconjugates. These interactions are often a prerequisite for infection. To date (March 2021), there have been 111, 113 and 6 different bacterial, viral and parasitic lectins respectively, structurally characterized resulting in 918 3D X-ray structures (Bonnardel et al., 2019).
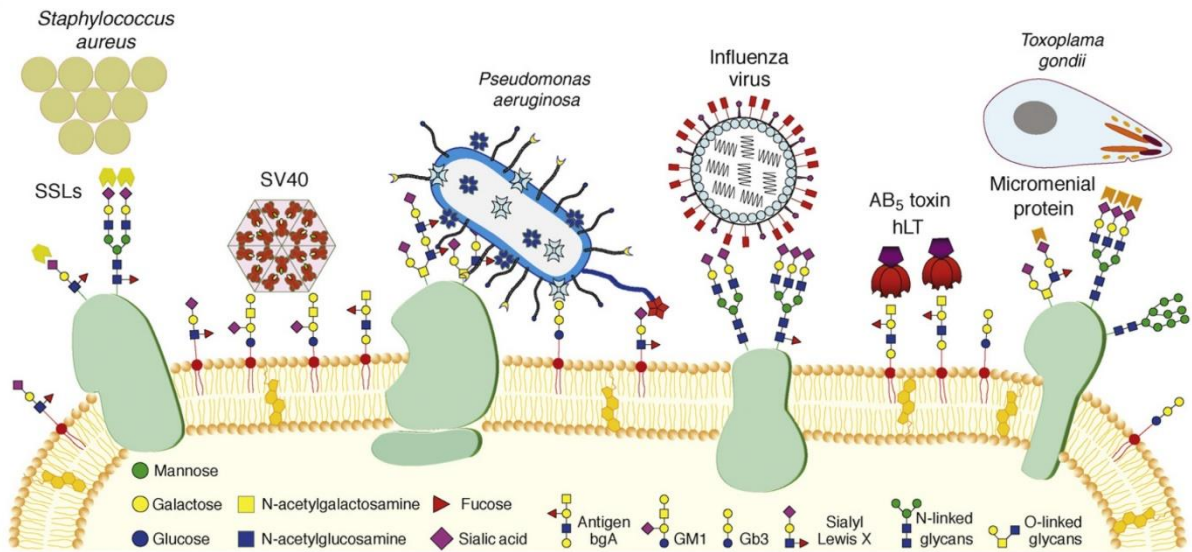
Figure 7: Schematic illustration of pathogenic organisms and their lectin-based strategies employed in cell-adhesion to host surface glycoconjugates. Figure is taken and modified from (Imberty and Varrot, 2008).

One of the most studied viral lectins is the influenza virus hemagglutinin (HA) that binds terminal sialic acid-containing glycoconjugates present on the host airway cell surface (Figure 7) inducing endocytosis and infection (Nizet et al., 2017). Interestingly, the host tropism of the viral infection may be influenced by different glycosylation patterns of the host cell glycoconjugates. Human HA recognizes α-2,6-linked sialosides while the avian influenza virus binds α-2,3-linked ones. The former are abundantly present in the upper respiratory tract while the latter are found in the intestinal tract and lower part of the respiratory tract (Stevens et al., 2006). The change in receptor specificity is an important step in crossing species-barrier and adaptation to human hosts, and this change can be induced by a point mutation in

HA (Yamada et al., 2006). Human-adapted avian influenza viruses such as H5N1 are therefore considered a global threat.

Bacterial lectins act as adhesins and toxins often in a form of complex protein appendages known as fimbriae and pili growing from the surface of bacteria that specifically recognize glycoproteins and glycolipids on the host cell (Figure 8). The most studied bacterial adhesin is FimH (Figure 9a) lectin located at the tip of the type-1 fimbriae from uropathogenic *Escherichia coli* that can bind to terminal α(1-2)-, α(1-3)-, and α(1-6)-linked mannosides of glycoproteins (Sauer et al., 2019). FimH lectin from uropathogenic *E. coli* binds glycoprotein Uroplakin Ia present on the epithelial surface of a bladder causing urinary tract infection (Sauer et al., 2019).
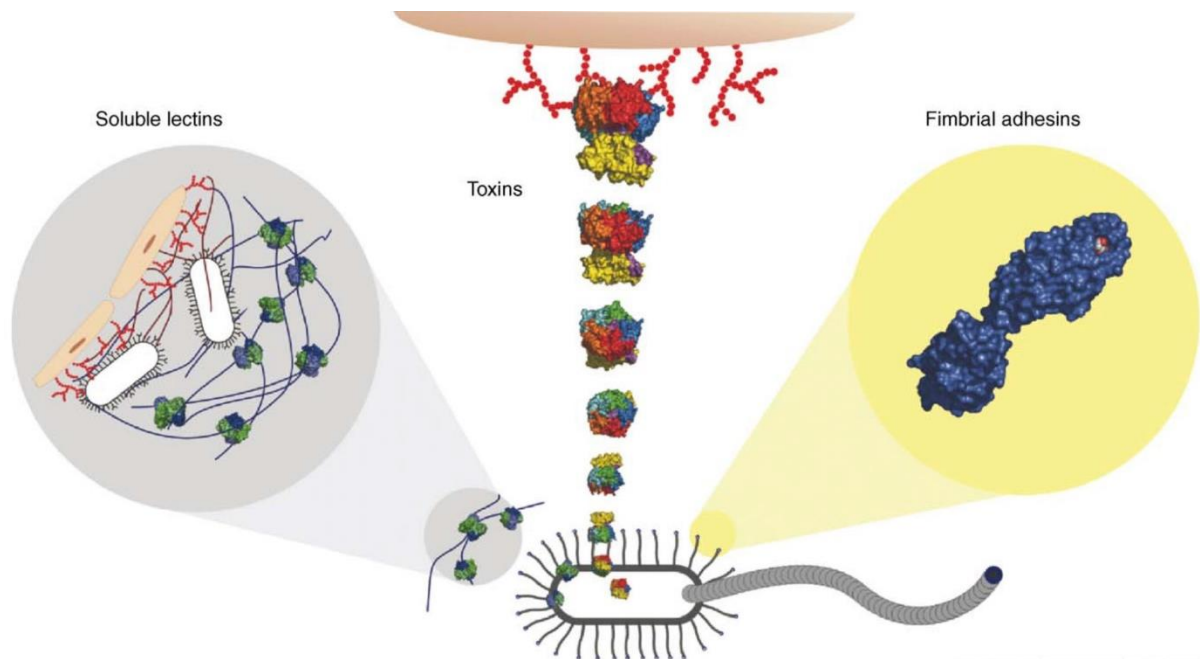
Figure 8: Schematic illustration of different types of lectins produced in bacteria. Figure taken and modified from (Imberty et al., 2005).

Bacteria can also secrete soluble lectins that are involved in adhesion, biofilm formation and cytotoxicity. Opportunistic bacteria *Pseudomonas aeruginosa* and *Burkholderia cenocepacia* that cause pulmonary infections in immunocompromised patients express soluble lectins present on the bacterial surface that bind human fucosylated oligosaccharides and are involved in bacterial adhesion in the early stages of infections (Imberty et al., 2005).

Numerous secreted toxins also bind glycoconjugates. For instance, hexameric AB_5 toxin from bacteria *Cholera vibrio* (cholera toxin) binds to Galβ1–3GalNAc moiety of GM1 ganglioside receptors via its pentameric B subunit (Figure 9b). Recently it has been shown that cholera toxin can also bind to fucosylated structures such as Lewis x trisaccharide histo-blood group antigen via secondary binding sites on the B subunit (Heim et al., 2019). Subunit A has an ADP-ribosylation activity that activates regulatory G-protein which in turn activates adenylate cyclase. The ultimate result is a severe alteration of ion homeostasis with efflux of chloride ions that leads to the secretion of water to the intestinal lumen causing severe diarrhoea (Nizet et al., 2017).

(a)   (b)
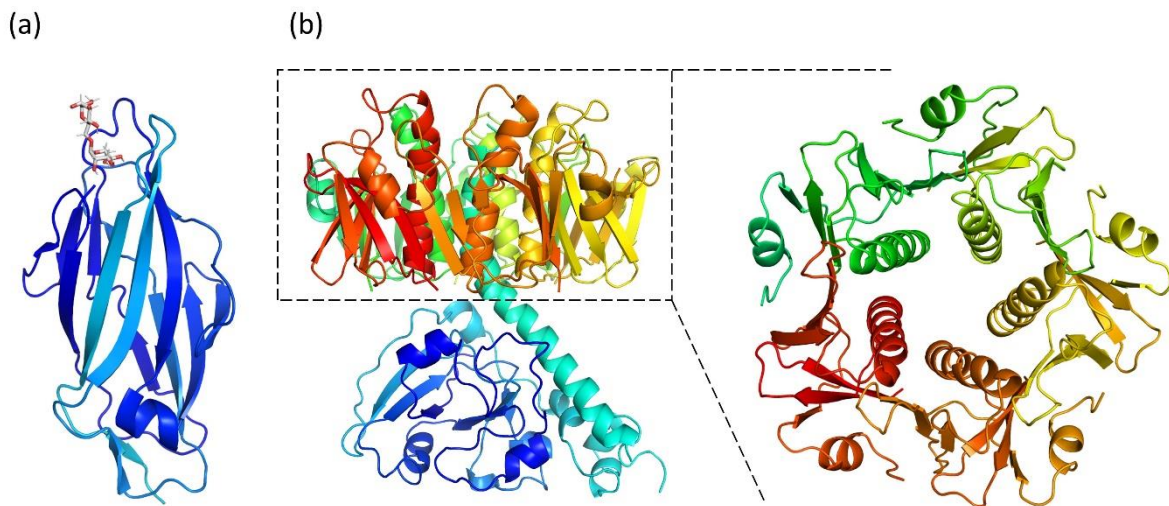
Figure 9: Cartoon representation of bacterial lectins. (a) FimH adhesin from *E. coli* in complex with Man(α1,2-)Man represented as grey sticks (PDB: 6GTX). (b) Cholera toxin with a detail on the pentameric subunit B (dashed box) that bind glycans (PDB: 1XTC).

In addition to viruses and bacteria, parasites produce lectins that bind glycan epitopes in a process of adhesion. The malaria parasite *Plasmodium falciparum* expresses an adhesin called erythrocyte-binding-antigen-175 (EBA-175) that binds the major membrane glycoprotein on human red blood cells called glycophorin A. This protein is highly sialylated and the terminal Neu5Ac(α2,3)-Gal-determinant is a preferred target for EBA-175 (Orlandi et al., 1992). The adhesin-glycoprotein binding triggers *Plasmodium* merozoites invasion of red blood cells where the parasite

undergoes reproduction (Koch et al., 2017). Another example is the human parasite *Entamoeba histolytica* causing amoebiasis. This parasite expresses a 260-kDa surface-anchored lectin on trophozoites that binds to terminal Gal/GalNAc glycans of glycoproteins and glycolipids present on the intestinal epithelium. The interaction mediates the adhesion and cytolysis of colonic epithelial cells and cells of the immune system such as macrophages, neutrophils and T-lymphocytes (Vines et al., 1998).

1.1.2   Structural basis of lectin-glycan recognition

A high number of X-ray crystal structures of lectin-carbohydrate complexes provided detailed information on sugar-binding modes at the atomic level. The Unilectin3D database (https://www.unilectin.eu/unilectin3D/) currently gathers 1456 three-dimensional X-ray structures of lectin-glycan complexes. The binding sites are usually found on the surface of the protein in the form of a shallow depression or a cleft. They are generally formed by non-contiguous amino acids within a single polypeptide, but occasionally several polypeptides can be involved in forming a binding site (Varrot et al., 2011). Large conformation changes are rarely observed upon carbohydrate binding. The binding of a carbohydrate may induce mild changes in the conformation of the amino acids to improve the stereocomplementarity of the binding. Oligosaccharides often exhibit conformational freedom and higher flexibility due to the free rotation around the glycosidic bonds between different monosaccharide constituents.

Lectins bind carbohydrates predominantly via a network of hydrogen bonds between sugar hydroxyl groups and main chain and side chain atoms of amino acids in the binding sites. These hydrogen bonds are either direct or can be bridged by water molecules. An important driving force behind the carbohydrate recognition is the CH/π dispersion interaction also known as hydrophobic interaction between the apolar face of a sugar molecule and aromatic amino acid residues (Figure 10) (Asensio et al., 2013). Recent studies indicate that the CH/π stacking interaction occurs in roughly 40 % of protein-carbohydrate complexes and can be the main driving force of the binding in some cases (Houser et al., 2020).

α-D-glucose   β-D-mannose   α-D-galactose
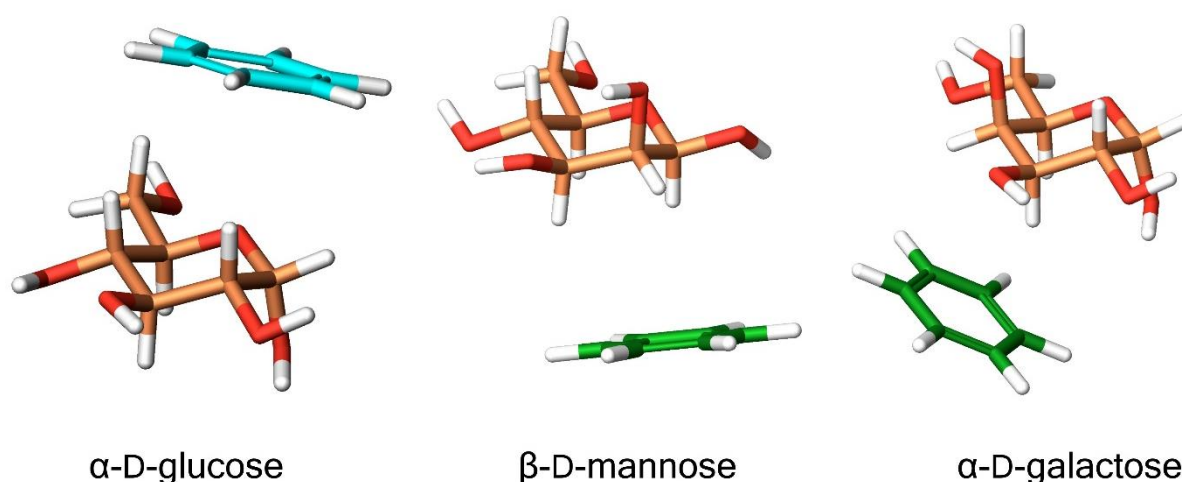
Figure 10: Examples of carbohydrate CH/π stacking interactions with aromatic amino acid residues. Carbohydrates are represented as sticks with carbon atoms coloured in orange. Figure is adapted from (Asensio et al., 2013).

In some cases, metal ions are involved in the binding sites where they bridge the interaction between a carbohydrate and a protein (Figure 11). The typical example is the C-type family of animal lectins that use calcium atoms in sugar recognition. In this case, upon the carbohydrate-binding, two sugar hydroxyl groups coordinate one calcium atom. Calcium-dependent sugar recognition has been also observed in soluble lectins from opportunistic bacteria such as *Pseudomonas aeruginosa*, *Ralstonia solanacearum* and *Chromobacterium violaceum*. *P. aeruginosa* produces two soluble lectins LecA and LecB specific to D-galactose and L-fucose respectively. In LecA, two hydroxyl groups OH3 and OH4 of galactose coordinate one calcium atom which is similar to the C-type lectins (Cioci et al., 2003). A unique binding mode has been observed in LecB lectin where three hydroxyl groups of fucose directly coordinate two calcium atoms resulting in un unusually high (micromolar) affinity (Mitchell et al., 2002).

Finally, electrostatic interactions can be involved in the sugar-binding even though they are rarely observed and are limited to special monosaccharides such as sialic acid.
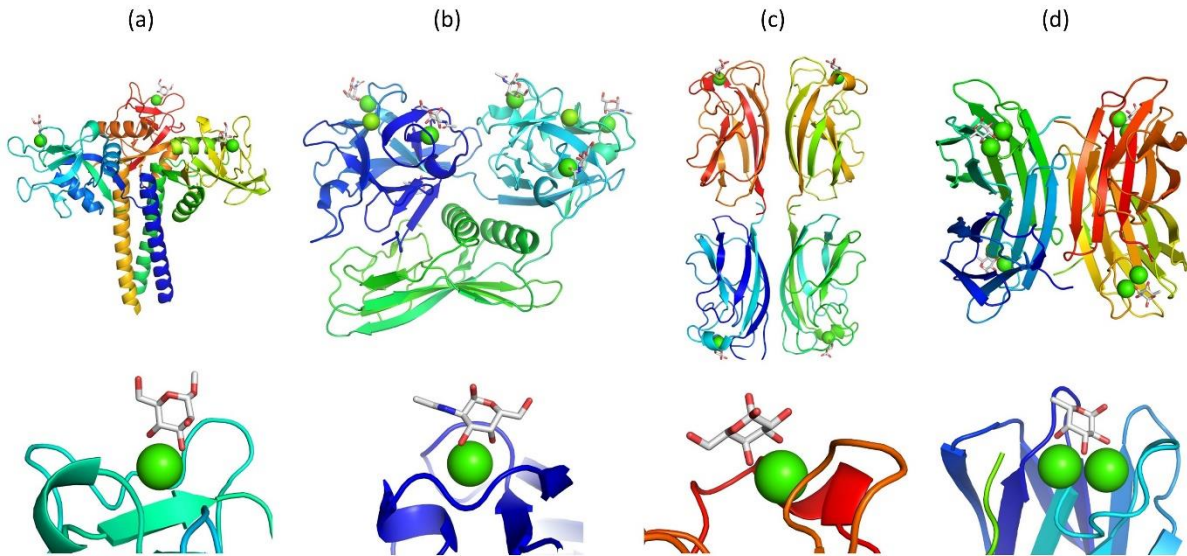
Figure 11: Examples of calcium-binding lectins. Proteins are depicted as cartoons, carbohydrates as grey sticks and calcium atoms as green spheres. (a) Human mannose-binding protein A in complex with D-mannose (PDB: 1KWU). (b) Sea cucumber CEL-III in complex with GalNAc (PDB: 2Z48). (c) LecA lectin in complex with D-galactose (PDB: 1OKO), (d) LecB lectin in complex with L-fucose (PDB: 1GZT).

1.2 Pseudomonas aeruginosa

Pseudomonas aeruginosa (*P. aeruginosa*) is a gram-negative, facultatively anaerobic, rod-shaped bacterium that can cause severe infections in immunocompromised patients. *P. aeruginosa* is a human opportunistic pathogen that is able to live in various environmental niches including soil, water and many human-associated environments such as moist surfaces of medical devices like catheters causing cross-infections in hospitals. This makes *Pseudomonas* a leading pathogen in nosocomial (hospital-acquired) infections. It typically infects urinary and respiratory tracts, wounds, burns and blood (Moradali et al., 2017).

The genome of *P. aeruginosa* is especially large for a prokaryote organism (5.5-7 Mbp) and consists of a single circular chromosome with an accessory genome

consisting of extrachromosomal elements such as plasmids, transposons, genomic islands or insertion sequences that account for high intra- and interclonal genome diversity (Kung et al., 2010; Stover et al., 2000). The genome contains one of the highest number of predicted regulatory genes (8.4 %) of all bacterial genomes, encoding transcriptional regulators and two-component regulatory systems. Another important group of proteins based on genome analysis are outer membrane proteins involved in adhesion, motility, transport of antibiotics, export of extracellular virulence factors and import of nutrients (Stover et al., 2000). Bacteria's preferred carbon source is not sugar but rather amino acids, organic acids and aromatic compounds. The metabolic diversity of *P. aeruginosa* is reflected in its genome with an unusually high number of genes coding for enzymes involved in β-oxidation metabolism such as acyl-CoA dehydrogenase (25 genes) and enoyl-CoA hydratase/isomerase (16 genes) (Stover et al., 2000). This correlates with its vast variety of cell membrane transport systems for mono-, di- and tri-carboxylates, and the absence of sugar transporters as well as enzymes involved in glycolysis (Lessie and Phibbs, 1984). *P. aeruginosa* utilizes several types of protein secretion systems to secrete virulence factors, toxins, lipases and proteases. Six different types of secretion systems have been identified and are called type I to type VI secretion systems (T1SS-T6SS) and *P. aeruginosa* possesses all of them but T4SS (Bleves et al., 2010). They represent complex assemblies that transport proteins through the inner and outer cell membrane to the extracellular medium. The role of secreted proteins is to capture essential elements such as iron, hydrolyze carbon compounds for metabolism, colonize the host tissue causing infection and escape the host immune system (Filloux, 2011). The genomic size together with its diversity and dynamics explains the evolutionary adaptability of *P. aeruginosa* to thrive in different environmental conditions.

Similar to other gram-negative bacteria, *P. aeruginosa* has developed a regulatory system known as quorum sensing (QS). Quorum sensing is a communication mechanism between cells that allows for a cell-density-dependent gene regulation. Bacteria use QS to monitor the cell-population density by detecting small signal

molecules produced by bacteria. The concentration of the signal molecule increases with the increasing bacterial population and when it reaches a threshold value, bacteria can respond by altering expression of target genes. These genes are involved in bacterial virulence, biofilm formation, production of secondary metabolites and stress-adaptation mechanisms. The small signal molecules in *P. aeruginos*a are acyl-homoserine lactones sensed by *las* and *rhl* QS systems, alkyl-quinolones used by *pqs* system and 2-(2-hydroxy-phenyl)thiazole-4-carbaldehyde used by *iqs* system (Lee and Zhang, 2014).

The ability of *P. aeruginosa* to form biofilms is an important prerequisite for colonization and for establishing chronic infection in human tissues. Biofilms are complex multicellular aggregates of bacteria composed of exopolysaccharides (EPS), exogenous DNA (eDNA) and proteins that hold microbial cells together attached to a surface and provide them with a protective environment. *P. aeruginosa* produces at least three extracellular polysaccharides, alginate, Psl and Pel that comprise the major part of the biofilm matrix (Jennings et al., 2015). Alginate is composed of linear copolymers of α-D-mannuronic acid and β-L-guluronic acid and is the major exopolysaccharide in mucoid strains of *P. aeruginosa* colonizing lungs of cystic fibrosis patients. Psl is a D-mannose-rich polysaccharide containing also D-glucose, D-galactose and L-rhamnose (Ma et al., 2007). Pel exopolysaccharide is composed of *N*-acetyl glucosamine (GlcNAc)- and *N*-acetyl galactosamine (GalNAc)-rich polysaccharides that crosslink eDNA in the biofilm (Jennings et al., 2015). Both Psl and Pel function as structural scaffolds maintaining the structural stability of the biofilm. The extracellular DNA interacts with EPS and has structural roles in the 3D matrix structure (Wang et al., 2015). It has also been observed that eDNA acidifies biofilms and thus delays penetration of positively charged aminoglycoside antibiotics across the biofilm (Wilton et al., 2016). Finally, the protein content of the biofilm matrix is not well characterized. One of the best-characterized matrix proteins is CdrA adhesin that binds to Psl exopolysaccharide and provides enhanced aggregate integrity by preventing CdrA proteolysis (Reichhardt et al., 2018).

P. aeruginosa has an enormous capacity to resist antibiotic treatment due to the utilization of many intrinsic and acquired drug-resistance mechanisms. The intrinsic mechanisms include reduced permeability of the outer membrane, expression of efflux pumps that expel the antibiotics out of the cell and production of antibiotic-modifying enzymes that inactivate them (Lambert, 2002). Most of the commonly applied antibiotics target inner cell components and must be able to penetrate the bacterial cell membrane. For instance, aminoglycoside antibiotics such as gentamycin and amikacin inhibit the protein synthesis by binding to the ribosomal 30S subunits (Mingeot-Leclercq et al., 1999). β-lactam ring-containing antibiotics such as penicillin and ampicillin block bacterial cell wall synthesis by targeting enzymes involved in peptidoglycan biosynthesis (Poole, 2004). Quinolone antibiotics such as levofloxacin target enzymes involved in DNA replication such as gyrase and topoisomerase IV (Aldred et al., 2014).

The outer membrane of *P. aeruginosa* is about 12 to 100-fold less permeable than that of *E. coli* and antibiotics penetrate the membrane either through porin channels or via interactions with LPS present on the outer membrane (Pang et al., 2019; Yoshimura and Nikaido, 1982). Bacteria utilize numerous efflux pumps which are big protein transmembrane complexes pumping the antibiotics out of the cell. Finally, another type of intrinsic antibiotic-resistance mechanism is expression of antibiotic-inactivating enzymes such as β-lactamase that hydrolyzes β-lactam ring and aminoglycoside-modifying enzymes such as aminoglycoside phosphotransferase, aminoglycoside nucleotidyltransferase and aminoglycoside acetyltransferase (Poole, 2005). Bacteria can acquire antibiotic-resistance through mutational changes or horizontal transfer of genes from other bacteria carrying plasmids with antibiotic-resistance genes that contribute to the development of multi-drug resistant (MDR) virulent strains.

Together with the adaptive antibiotic resistance mechanisms including biofilm formation and evolution of bacterial persister cells that are highly tolerant to

antibiotics, this makes *Pseudomonas aeruginosa* a dangerous pathogen in patients with defective immune system.

1.2.1 *P. aeruginosa* infections immunocompromised patients

Cystic fibrosis (CF) is one of the most common (the incidence is estimated to be between 1/3000 and 1/6000) autosomal recessive disease (Scotet et al., 2020) caused by a mutation in a gene encoding the cystic fibrosis transmembrane conductance regulator (CFTR) (Kerem et al., 1989). More than 2000 mutations in the *CFTR* gene have been identified thus far, although with a single deletion of phenylalanine 508 occurring in 90 % of cystic fibrosis cases (Lukacs and Verkman, 2012). CFTR is an ion channel transferring chloride ions across the epithelial cell membrane and its dysregulation results in multiple-organ dysfunction, one of which is a build-up of thick mucus in the lungs.

CF patients suffer from airway obstruction with hyperviscous mucus, reduced antimicrobial functions and impaired mucociliary clearance which make them more prone to severe and persistent bacterial lung infections. The median age of death of CF patients was 32.4 years in 2019 based on the 2019 Patient Registry Annual Data Report of Cystic Fibrosis Foundation (https://www.cff.org/About-Us/Reports-and-Financials/2019-Annual-Report/). The primary cause of death was respiratory/cardiorespiratory-related disease accounting for 62 % of all deaths. *P. aeruginosa* is known as the main cause of morbidity and mortality in CF (Figure 12).

Interestingly, the mutant version of CFTR has also effects on glycosylation phenotype with increased fucosylation and decreased sialylation of cell surface glycoconjugates of CF airway epithelial cells (Rhim et al., 2004; Scanlin et al., 1985). In addition, the human airway mucins from CF patients also show glycosylation modifications with a higher amount of sialylated and sulphated Lewis x antigens (Lamblin et al., 2001). *P. aeruginosa* uses flagellar and fimbrial adhesins and non-pili

lectins for specific recognition of these glycoconjugates allowing adherence to host tissue in early stages of infection (Bucior et al., 2012; Scharfman et al., 2001).

The prevalence of *P. aeruginosa* infection in CF patients has been continually decreasing over recent years which might be an outcome of implementation of pulmonary guidelines and pharmacological approaches in the prevention and eradication of initial infection (Mogayzel et al., 2014). Nevertheless, the prevalence of MDR *P. aeruginosa* infection has stayed constant over time (Figure 13).
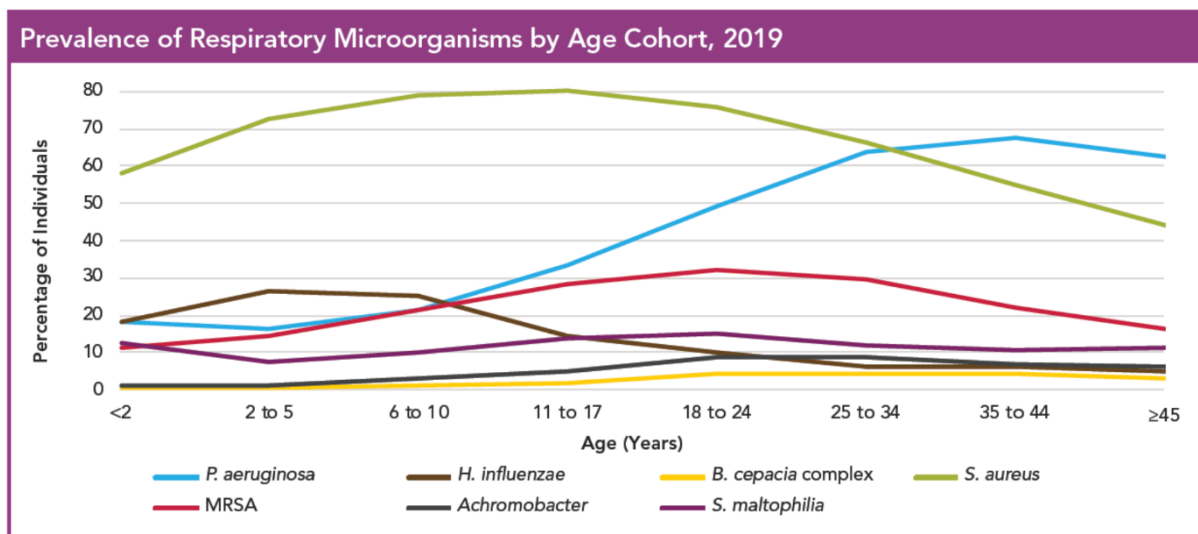
Figure 12: Prevalence of respiratory microorganisms in cystic fibrosis patients in the United States (Cystic Fibrosis Foundation, 2020).
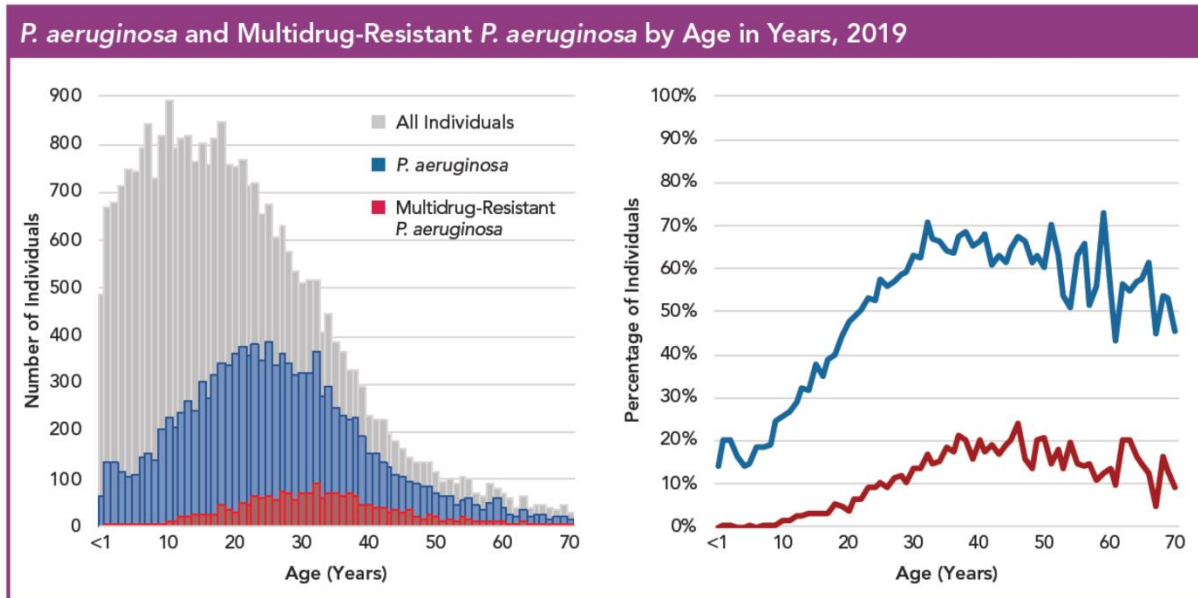
Figure 13: Prevalence of *Pseudomonas aeruginosa* and multi-drug resistant *P. aeruginosa* infection in cystic fibrosis patients by age in the United States (Cystic Fibrosis Foundation, 2020).

1.2.2 Role of lectins in *Pseudomonas* infection

P. aeruginosa produces several lectins as a part of surface pili that function in bacterial adherence, motility, virulence and biofilm formation. The type IV pili are several micrometers long proteinaceous appendages present on the surface of bacteria mostly composed of pilin (PilA protein) that acts both as a structural component and an adhesin. PilA was observed to bind glycosphingolipids asialo-GM1 and asialo-GM2 (Gupta et al., 1994) by its C-terminal terminal region (Lee et al., 1994) that recognizes βGalNAc(1-4)βGal disaccharide motif (Hazes et al., 2000; Sheth et al., 1994). Another type of fimbriae produced by bacteria is a flagellum which is a unique organelle responsible for a "swimming" movement of bacteria. Flagella consist of polymerized flagellin proteins (FliC) that binds to glycolipids GM1 and asialo-GM1 (Feldman et al., 1998) and to mucin MUC1 (Lillehoj et al., 2002).

    P. aeruginosa also produces two soluble non-pili lectins LecA (PA-IL) and LecB (PA-IIL) specific to D-galactose- and L-fucose-containing glycans respectively, and both have been identified as virulence factors in *P. aeruginosa* infection. Both lectins are

expressed under the control of quorum sensing and RpoS sigma factor that acts as a regulator of the stress response and activates expression of several virulence factors (Schuster et al., 2004; Winzer et al., 2000). Both lectins are present mostly in the bacterial cytoplasm with abundant fractions in the outer membrane (Glick and Garber, 1983). LecB binds to the outer membrane porin OprF that also acts as an adhesin (Funken et al., 2012). Furthermore, LecB also binds to the branched Psl exopolysaccharide of *P. aeruginosa* biofilm which leads to improved retention of cells in the biofilm matrix and contributes to the biofilm structural integrity (Passos da Silva et al., 2019). This is in agreement with a LecB-deficient mutant strain of *P. aeruginosa* that was impaired in biofilm formation (Tielker et al., 2005). LecA lectin has also been shown to contribute to biofilm development with decreased biofilm formation and dispersal of biofilms after applying competitive inhibitors of LecA (Diggle et al., 2006).

Both LecA and LecB lectins are implicated in the lung injury observed on *in vitro* and *in vivo* models suggesting direct effects of both lectins on pathogenicity of *P. aeruginosa* infection. The LecA/LecB-mutant strain of *P. aeruginosa* showed increased bacterial clearance and reduced adhesion and bacterial dissemination. The same outcome was observed in the wild-type strain with co-administration of specific lectin-competitive carbohydrates such as Me-α-Gal for LecA and Me-α-Fuc for LecB that proved to be effective inhibitors (Chemani et al., 2009). LecA has been shown to have a dose-dependent cytotoxic effect on the respiratory epithelial cells contributing to the epithelial damage (Bajolet-Laudinat et al., 1994). It has also been demonstrated that LecA interaction with host cell glycosphingolipid Gb3 promotes membrane engulfment of *P. aeruginosa* by human lung epithelial cells (Eierhoff et al., 2014). Additionally, LecA lectin induced an increase in permeability of intestinal epithelium to exotoxin A (Laughlin et al., 2000). LecB has been shown to stop the ciliary beating of human airways which is an important defense mechanism against microbes (Adam et al., 1997). Taken together, it has been shown that both LecA and LecB lectins contribute

significantly to the pathogenicity of *P. aeruginosa* infection and that these lectins can be viewed as potential targets for a drug-development.

Both LecA and LecB lectins require calcium atoms for carbohydrate-binding. Metal ions directly interact with the ligand which is a striking feature for both lectins. A detailed description of the overall structure and ligand-binding mode will be described below.

1.2.2.1   LecA (PA-IL) lectin

LecA displays medium-range affinity for D-galactose with an association constant (K_a) of 3.4 x 10^4 M^{-1}. Among monosaccharides, the specificity is strictly towards galactose with an exception of *N*-acetyl-D-galactosamine that binds also to LecA with lower affinity. The presence of a hydrophobic moiety on the anomeric position, in either α or β configuration, increases the affinity with the strongest binding obtained for phenyl-β-thiogalactoside (Kadam et al., 2013). LecA also binds to adenine and acyl homoserine lactones quorum-sensing signal molecules (Boteva et al., 2005; Stoitsova et al., 2003). LecA recognizes terminally placed Galα1-3Gal, Galα1-4Gal and Galα1-6Glc (melibiose) structures on glycolipids and glycoproteins (Blanchard et al., 2014; Chen et al., 1998; Lanne et al., 1994).

LecA is a homotetramer (12.7 kDa for a monomer) composed of four chains of 121 amino acids arranged around a 222 pseudo axis of symmetry. Each monomer adopts a jelly roll β-sandwich fold composed of two curved sheets, each consisting of four antiparallel β-strands (Figure 14).
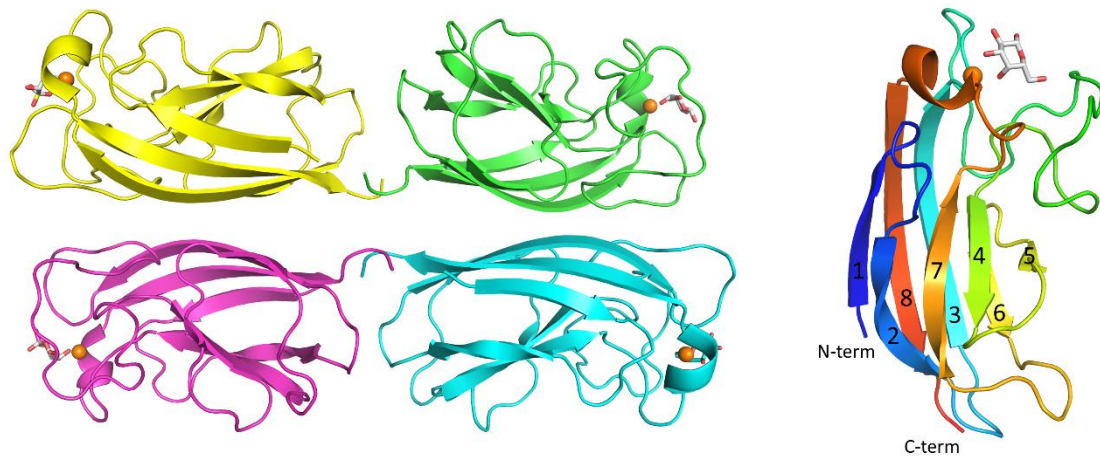
Figure 14: Structure of LecA/galactose complex (PDB: 1OKO). (A) Tetrameric arrangement of LecA with four subunits, each containing D-galactose ligand (grey sticks) and a calcium atom (orange sphere). (B) Monomer of LecA is composed of two curved sheets, each formed by 4 antiparallel β-strands.

In the ligand-bound form, each monomer contains a calcium ion and one galactose molecule (Figure 15). The calcium atom is directly coordinated by O3 and O4 atoms of galactose. All galactose hydroxyl groups but the anomeric one are involved in the hydrogen bonding with protein. The lectin-carbohydrate interaction is made by six direct hydrogen bonds between the sugar hydroxyl groups and protein amino acids. O4 of galactose forms a hydrogen bond with Asp100. Asn107 establishes hydrogen bonds with O2 and O3 while Thr104 and Val101 participate in the hydrophobic contacts with the apolar face of galactose. O6 forms a hydrogen bond with His50 and Gln53 with one water molecule making a bridge to the carbonyl oxygen of the protein backbone. Several water molecules make contact with the sugar hydroxyl groups. Interestingly, the calcium-mediated sugar recognition resembles the one in the C-type family of lectins despite the lack of fold and sequence similarities between LecA and C-type lectins. In all of them, O3 and O4 sugar hydroxyl groups coordinate one calcium atom and O4 additionally establishes a hydrogen bond with Asp residue that participates in the calcium coordination (Cioci et al., 2003).
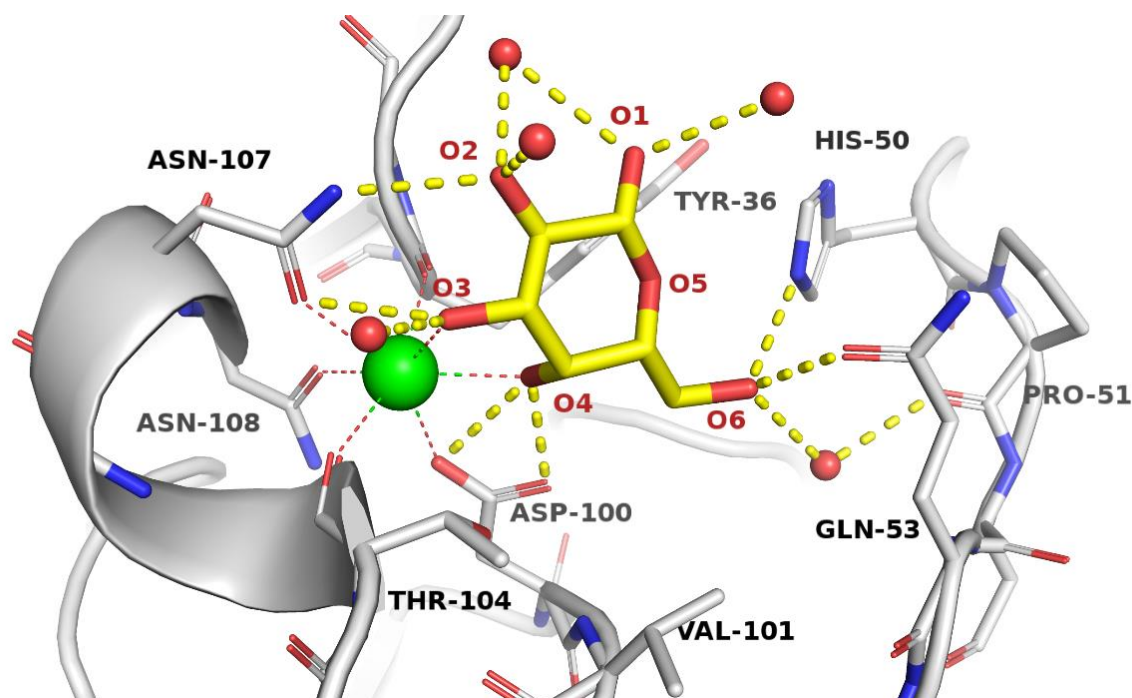
Figure 15: Cartoon representation of the calcium and galactose binding site of LecA. Galactose is represented as yellow sticks. Calcium atom and oxygen atoms (of water molecules) are represented as green and red spheres, respectively. Hydrogen bonds are depicted as yellow dashed lines and coordination contacts as red dashed lines. (PDB: 1OKO).

1.2.2.2   LecB (PA-IIL) lectin

LecB (11.7 kDa for a monomer) is a fucose-specific lectin with un unusually high affinity towards L-fucose with an association constant (K_a) of 1.5 x 10^6 M^{-1} (Garber et al., 1987). Hemagglutination-inhibition assays showed that LecB can also bind other monosaccharides but with lower affinity (strongest first) L-fucose > L-galactose > D-arabinose > D-fructose > D-mannose (Garber et al., 1987; Sabin et al., 2006). The reason why LecB can recognize all of these glycans is that they exhibit the same either L-galacto or D-manno configuration that contains a sequence of two equatorial and one axial hydroxyl groups when in the $^1C_4$ and $^4C_1$ pyranose conformation, respectively.

The overall structure of LecB is that of a homotetramer composed of four independent subunits of 114 amino acids arranged around a pseudo 222 axis of symmetry (Figure 16). Each subunit folds as a β-sandwich composed of nine

antiparallel strands. The first five strands form a Greek key structural motif, extended by strands 6-8 (Figure 16a). The N-terminal β-strand inserts between strands 5 and 6. The dimerization results from a head-to-tail association of two monomers (Figure 16b). Tetramerization occurs by the antiparallel association of strands 6 of each dimer to its corresponding counterparts in the other dimer. An important feature is the involvement of the C-terminal carboxyl group (Gly114) in the ligand-binding site of the neighbouring monomer.
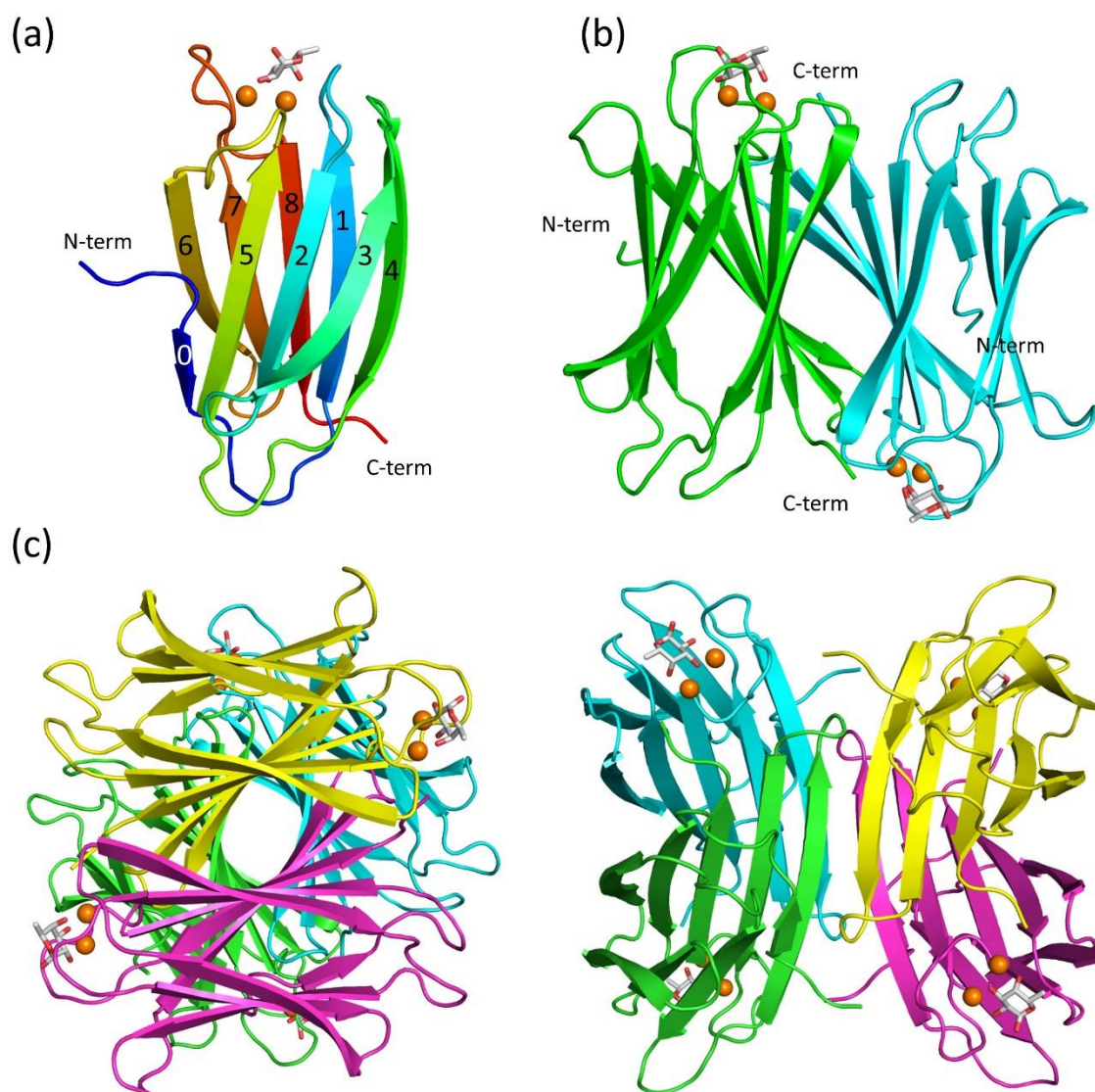
Figure 16: Three-dimensional structure of LecB/fucose complex. Fucose is in stick representation (grey) and calcium atoms are depicted as orange spheres. (a) Monomeric unit of LecB is composed of 9 antiparallel β-strands forming a curved β-sandwich. (b) Dimerization of chain A and chain B. (c) Two perpendicular views of LecB tetramer. (PDB: 1GZT).

In the LecB/fucose complex, sugar sits on a pair of calcium atoms that are directly involved in the binding (Figure 17). There is an apparent asymmetry in the ligand distribution of the calcium coordination. The site 1 calcium interacts with five carboxylate and two ligand oxygen atoms while site 2 calcium interacts with three carboxylate, two carbonyl and two ligand oxygen atoms. The average distance between the two calcium atoms is rather small, 3.75 Å, possibly due to the overcompensation of the positive charges by acidic residues. Three fucose hydroxyl groups coordinate the Ca^{2+} ions, O2 coordinates the first calcium, O3 both and O4 the second calcium. These three hydroxyl groups also participate in the hydrogen bond network with the acidic residues in the binding site. Moreover, ring oxygen O5 of fucose interacts with the main chain nitrogen atom. The fucose methyl group creates an additional hydrophobic contact with hydroxymethyl and methyl side chains of Ser23 and Thr45, respectively. A water molecule mediates an additional hydrogen bond between the fucose O2 atom and the backbone amide of Thr98.
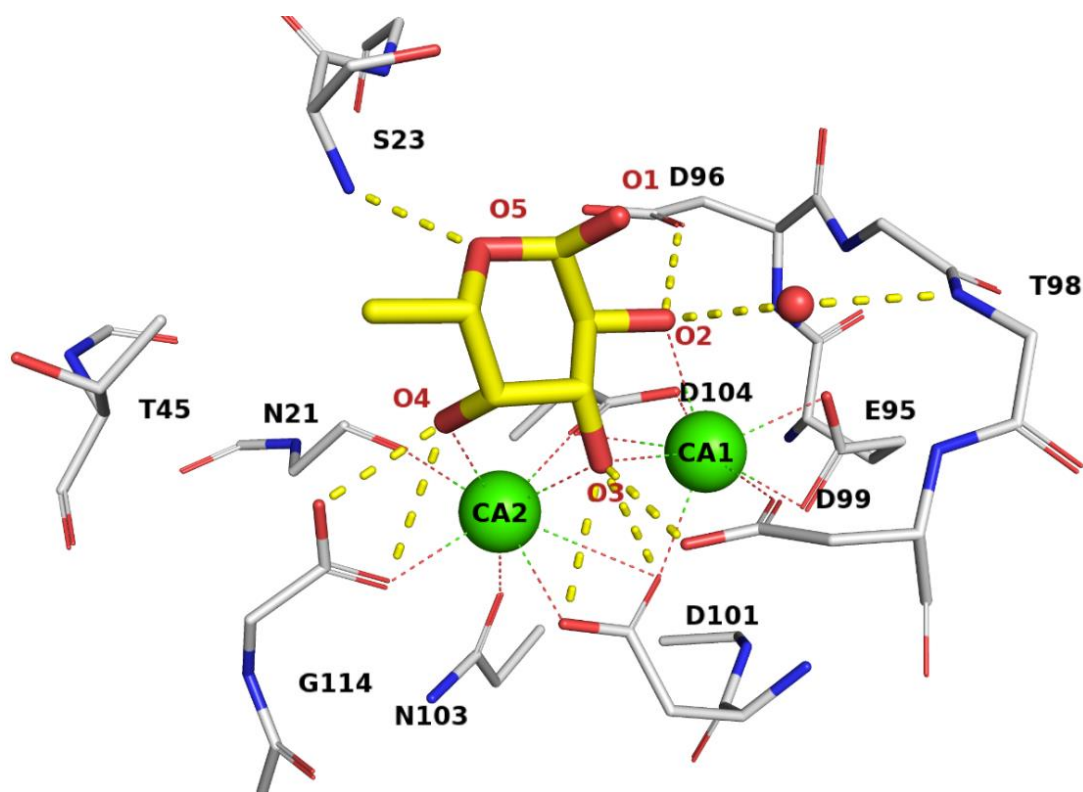
Figure 17: Stick representation of the calcium and fucose binding site in LecB. Fucose is represented by yellow sticks. Calcium and oxygen atoms (of water molecules) are represented as green and red spheres, respectively. Hydrogen bonds and calcium coordination are represented by thick yellow dashed lines and thin red dashed lines, respectively. (PDB: 1GZT).

LecB can also bind to human oligosaccharides bearing terminal L-fucose including A, B, H and Lewis blood groups (Gilboa-Garber et al., 1994; Lesman-Movshovich et al., 2003; Perret et al., 2005). The Lewis a antigen has been identified as the best natural ligand with the highest affinity (K_d) of 200 nM for LecB (Perret et al., 2005). This was rationalized by the X-ray crystal structure of the LecB/Le[a] complex that identified an additional hydrogen bond between GlcNAc-6-OH and the protein main chain.

1.2.3   Antiadhesive therapy

In 2017, The World Health Organization published a list of bacteria for which new antibiotics are needed and ranked *P. aeruginosa* a Priority 1 pathogen addressing the

urgent need for research and development of novel effective therapeutics against this multi-drug resistant bacterium (Tacconelli et al., 2018). Alongside traditional antibiotic-based chemotherapy, the aim of which is to suppress the growth or kill the microorganism, new approaches of anti-pseudomonal therapy emerge. The aim of the new therapy is to reduce the colonization and damage caused by bacteria by targeting biofilm formation, virulence factors, quorum sensing or adhesins. LecA and LecB lectins produced by *P. aeruginosa* are extensively investigated as targets for new drugs with anti-adhesive properties. The general approach of antiadhesive therapy is to design drugs that would bind to lectins with high specificity and affinity and thus inhibit the binding of bacteria to human cells and/or reduce the biofilm formation (Sommer et al., 2013).

One of the most studied targets for anti-adhesive therapy is FimH adhesin from uropathogenic *E. coli* that is able to bind highly mannosylated receptors of urothelial cells resulting in colonization and infection of the urinary tract (Sauer et al., 2019). Administration of D-mannose in the treatment of recurrent urinary tract infection (UTI) has been evaluated in clinical studies showing efficacy in reducing symptoms and the recurrence rate of UTIs (Domenici et al., 2016). Since FimH is a target of interest, various mono-and polyvalent FimH-inhibitors based on D-mannose ligand have been developed and analyzed over the past decades (Mydock-McGrane et al., 2017). Ligand-based drug design yielded several potent glycomimetic compounds with improved pharmacokinetic properties such as increased metabolic stability and bioavailability. D-mannose-based glycomimetic compounds bearing an aromatic aglycone part showed nanomolar affinities due to the binding to the lipophilic pocket called "tyrosine gate" of FimH (Schönemann et al., 2019). Although the *O*-linked mannopyranosides were found to be inadequate for therapeutic purposes due to the high susceptibility to glycohydrolysis, the *C*-glycosides with similar binding affinities to the *O*-linked mannosides show huge therapeutic potential with some of them in preclinical and clinical phase (Mousavifar et al., 2018).

Both LecA and LecB lectins from *P. aeruginosa* are potential targets for antiadhesive therapy and various carbohydrates have been tested to decrease bacterial adherence through competition with lectins (Chemani et al., 2009). Administration of simple sugars including mannose, fucose and galactose inhibited bacterial adhesion both *in vitro* and in the murine model and reduced the lung damage with further reduction when co-administrated with pseudomonal antibiotics (Bucior et al., 2013; Chemani et al., 2009). Human milk oligosaccharides also inhibited *P. aeruginosa* adhesion to human respiratory epithelial cells (Weichert et al., 2013).

Numerous monovalent and multivalent ligands have been synthesized although monovalent ligands showed little increase in binding affinities. The affinity of galactosides to LecA can be increased by adding an aromatic substituent on the anomeric position, for example, such as naphthyl 1-thio-β-D-galactopyranoside with 10-fold increase in binding affinity (Rodrigue et al., 2013). This is associated with an additional T-shaped CH/π interaction between His50 in LecA and the aromatic ring of the galactoside aglycone (Kadam et al., 2013). Recently synthesized highly potent divalent LecA inhibitors that can bind two binding sites per monomer showed the highest nanomolar affinities reported to date (Zahorska et al., 2020).

The design of monovalent inhibitors of LecB that would have higher affinity than those observed for Lewis a antigen has not been very successful. Moreover, inhibitors bearing terminal fucosides can interfere with other fucose-binding proteins of the host system such as DC-SIGN, MBL or selectins that play important roles in the immune system (Hauck et al., 2013). On the other hand, monovalent glycomimetic LecB inhibitors based on D-mannose scaffold, cinnamides and sulphonamides showed nanomolar potency against LecB with excellent pharmacokinetic properties on a murine *in vivo* model (Sommer et al., 2018). Moreover, these inhibitors effectively inhibited the biofilm formation without influencing the bacterial viability and have thus therapeutic potential.

Since both LecA and LecB lectins are homotetrameric proteins, they can be targeted by multivalent ligands such as glycoclusters, glycodendrimers, glycopolymers,

glyconanoparticles or neoglycoproteins that show nanomolar affinities (Cecioni et al., 2015; Grishin et al., 2015).

1.3 *Photorhabdus laumondii*

The genus *Photorhabdus* comprise gram-negative bacteria from the family *Morganellaceae* that live in a symbiotic association with nematodes from the genus *Heterorhabditis* (Clarke, 2020). *Photorhabdus* are motile, facultatively anaerobic rods and so far the only known bioluminescent terrestrial bacteria (the name comes from Latin and means "glowing rod"). Recent phylogenetic reconstruction based on the genome sequencing of *Photorhabdus* bacterial strains yielded 19 different species in *Photorhabdus* genus, out of which *Photorhabdus laumondii*, subsp. *laumondii* (formerly *Photorhabdus luminescens*, subsp. *laumondii*) is the most studied and best characterized (Machado et al., 2018).

All *Photorhabdus* strains are pathogenic to insects (entomopathogens) while they live exclusively in a mutualistic association with nematodes from the *Heterorhabditis* genus. In addition, one species, *Photorhabdus asymbiotica*, has also been identified as an emerging human pathogen causing local soft tissue infections with disseminating bacteremia (Gerrard et al., 2004).

Bacteria reside in the guts of a specialized stage of a nematode host called "infective juvenile" (IJ) (Waterfield et al., 2009). IJs live freely in the soil where they actively seek out insect prey, which is usually beetles, moths or butterflies in larval stages. The nematode can penetrate the body of the insect either through natural openings or directly through the cuticle. Once inside the insect, the IJ migrates into the insect haemolymph (fluid filled with cells and plasma analogous to blood in vertebrates) and regurgitates bacteria from its guts into the haemolymph. This is the start of the process called IJ recovery when the IJ develops into a self-fertile adult hermaphrodite nematode. Meanwhile, bacteria exponentially replicate and kill the insect prey by establishing septicaemia and toxaemia. All of the insect tissues and

organs are converted into the *Photorhabdus* biomass, which serves as a source of nutrients for the nematode growth and development. The adult nematode proceeds with laying eggs externally which develop through four larval stages into adult nematodes. Interestingly, in *Heterorhabditis bacteriophora*, it has been observed that the nematode can lay eggs also internally inside the nematode body in a process called *endotokia matricida* (Ciche et al., 2008). These eggs develop into IJs that eventually disrupt the parental nematode body cavity and emerge into the soil ready to infect a new insect host. During this complex cycle of events (Figure 18), *Photorhabdus* must kill the insect prey, support the nematode reproduction and development and colonize the new offspring. Photorhabdus thus represents a unique model system for the study of pathogenicity and symbiosis.
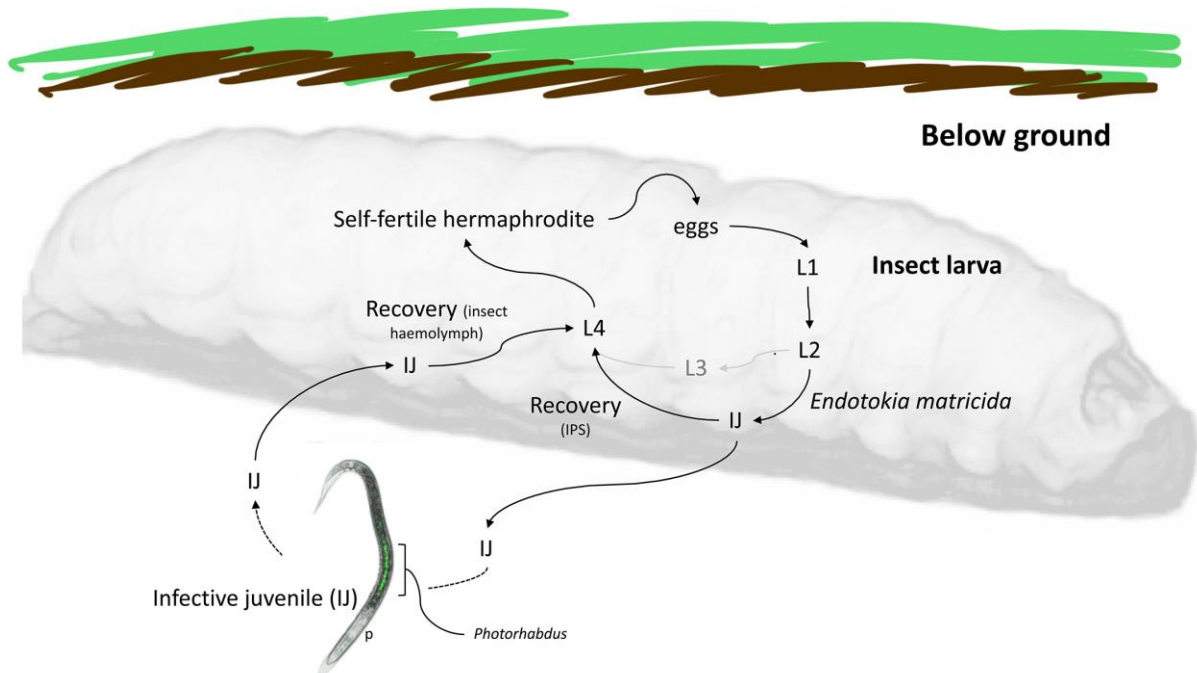
Figure 18: Life cycle of *Heterorhabditis* nematode. Bacteria reside in the guts of infective juveniles (IJ) of *Heterorhabditis* nematode. Nematodes live in soil where they actively search for insect prey. After penetration of the insect larva, they regurgitate bacteria into the insect haemolymph that eventually kill the insect host. The IJ develops into an adult hermaphrodite nematode in a process called IJ recovery. The adult self-fertile nematode then lays eggs externally that develop through four larval stages (L1-L4) into an adult nematode. In subsequent generations, the IJs lay eggs internally in a process called *endotokia matricida* that yield new IJs that can either develop into an adult nematode or leave the dead body and start a search for a new prey. Figure is taken from (Clarke, 2020).

1.3.1   *Photorhabdus* as a nematode symbiont

As described above, Photorhabdus interacts with its nematode partner to help in the IJ recovery, to support the nematode growth and development and to colonize the newly developed IJs. After regurgitation of bacteria into the insect haemolymph, IJ detects yet uncharacterized signals that start the process of IJ recovery at this early stage of infection. The nematode reproduction starts after the insect death and the recovery of the newly reproduced IJs is stimulated by a food signal produced by bacteria. It has been shown that part of this signal consists of a small bioactive molecule stilbene called 3,5-dihydroxy-4-isopropylstilbene (IPS), which is a major secondary metabolite produced by *Photorhabdus* (Strauch and Ehlers, 1998). Stilbenes are mostly associated with plants as a family of small molecules that are synthesized in response to various environmental or biotic stresses and *Photorhabdus* is the only non-plant organism reported to produce stilbene (Joyce et al., 2008). IPS is a multipotent molecule and plays various roles in the *Photorhabdus* life cycle as will be discussed later. It also acts as an antibiotic with antibacterial, antifungal and nematicidal properties which might protect the insect cadaver from other saprophytic organisms (Hapeshi et al., 2019). Even though nematodes need this molecule for IJ recovery, it is not sufficient and viable bacteria must also be present (Joyce et al., 2008).

Furthermore, it has been shown that *Photorhabdus*-free media do not support the nematode growth and development and that bacteria thus play an essential role in the nematode nutrition and reproduction (Clarke, 2020). The molecular mechanisms are still to be investigated but it seems that bacteria can provide the nematode with high amounts of amino acids that accumulate in the form of proteinaceous crystalline inclusion bodies (Bowen and Ensign, 2001).

The key event in the *Photorhabdus* life cycle is the transmission process where *Photorhabdus* bacteria colonize guts of the newly developed IJs. Transmission is a rather complex event and was originally described in *Heterorhabditis bacteriophora*

nematode (a symbiont partner of *Photorhabdus laumondii* TTO1) undergoing *endotokia matricida* (IJs develop inside the body cavity of the parental nematode) (Ciche et al., 2008). In the beginning of the process, *Photorhabdus* attaches to the gut cells of the parental nematode followed by an invasion of the neighbouring rectal gland cells. After lysis of the rectal gland cells, bacteria are released into the body cavity of the parental nematode.

The molecular interactions in this complicated process must involve cell-specific adhesion. One of the key determinants of bacterial adherence is Mad fimbriae, the expression of which is controlled by an invertible DNA switch referred to as a *madswitch* (Somvanshi et al., 2012, 2010). Bacteria can display different phenotypes based on the activation and deactivation of the *madswitch*. When the switch is in the ON orientation (Mad fimbriae are produced), bacteria are in the M-form (mutualistic) which allows the nematode colonization. When it is in the OFF orientation (Mad fimbriae are not produced), bacteria are in the P-form (pathogenic) and support the nematode growth.

1.3.2 *Photorhabdus* as an insect pathogen

After regurgitation of bacteria into the insect haemolymph, *Photorhabdus* faces big challenges from the insect's innate immune system. Photorhabdus is a very powerful insect pathogen. The insect immune system is composed of cellular and humoral branches (Eleftherianos et al., 2017). The cellular response is mediated by insect immune cells haemocytes circulating in the haemolymph that can induce phagocytosis and nodule formation (Ratcliffe and Gagen, 1977). The humoral response involves the production of antimicrobial peptides (AMPs) such as attacins and cecropins, as well as activation of the prophenoloxidase (PPO) into phenoloxidase (PO) by circulating serine proteases, expression of which is induced after recognition of the pathogen-associated molecular patterns (PAMPs) (Eleftherianos et al., 2017). The enzyme deposits melanin into bacteria, encapsulating them and introducing reactive oxygen species directly killing the invading organism.

Photorhabdus has evolved many mechanisms to avoid the insect immune response and eventually kill the insect host. Similarly to other gram-negative bacteria, *Photorhabdus* uses the PhoPQ two-component system that regulates the expression of genes involved in modifications of lipopolysaccharides on the bacterial cell envelope in response to AMPs produced by the insect (Derzelle et al., 2004). *Photorhabdus* can also inhibit both the humoral and cellular insect immune responses. As described earlier, *Photorhabdus* produces isopropylstilbene that not only acts as a signalling molecule and an antibiotic but also as an inhibitor of PO (Eleftherianos et al., 2007).

Furthermore, *Photorhabdus* produces a large number of adhesins, hemolysins and toxins. In fact, after sequencing *P. laumondii* TTO1 genome in 2003, more genes encoding toxins were predicted for *Photorhabdus* than for any other organism, for which the genome had been sequenced at that time (Duchaud et al., 2003). Several toxin families were identified and characterized. The Tc toxins are big assemblies composed of three components TcA, TcB and TcC that use a syringe-like mechanism to inject toxins into the cytoplasm of the cell (Meusch et al., 2014; Waterfield et al., 2001). The Mcf toxins cause a rapid loss of body turgor in the intoxicated insect as indicated by its name "makes caterpillars floppy" (Mcf) (Daborn et al., 2002). Finally, *Photorhabdus* virulence cassettes are phage tail-like protein complexes that resemble R-type pyocins and act like syringes that can deliver toxins into the targeted cells (Yang et al., 2006).

1.3.3   Entomopathogenic complex as a biopesticide

The remarkable life cycle of entomopathogenic complex (EPN) is not interesting only for studying how organisms switch between pathogenicity and symbiosis but also for its use in commercial applications. Since *Photorhabdus* is pathogenic to a wide variety of insect species, the entomopathogenic complex has been used as a biocontrol agent against insect pests in crops (Labaude and Griffin, 2018). EPN together with entomopathogenic fungi, nematicidal fungi, parasitoid wasps, *Bacillus thuringiensis* bacteria and baculoviruses are becoming employed as biopesticides to reduce the

negative impact of chemical pesticides (Kumar et al., 2019; Mbata et al., 2019). Even though the use of EPN has advantages such as a broad host range, fast speed of kill, active search for an insect host and safety; research still needs to be done on the large-scale production, storage and application methods. Some limitations also rise from the environmental requirements for the effectiveness of EPN such as moisture of the soil, sensitivity to UV radiation, temperature, presence of predators, pathogens, *etc.* (Labaude and Griffin, 2018). Understanding the biology and behaviour of EPN can help in the development of more effective and resistant entomopathogenic nematodes.

1.3.4 Lectins from *Photorhabdus* bacteria

Many pathogenic bacteria use lectins in targeting the host-cell glycans to establish the interactions. *Photorhabdus* interacts with both the insect prey and the nematode hosts and one type of proteins involved in such specific associations can be lectins.

To date, there are 5 lectins produced by *Photorhabdus* strains that have been reported. PLL (Kumar et al., 2016), PLL2 (Fujdiarová et al., 2020), PLL3 (Faltinek et al., 2019), PllA (Beshr et al., 2017) *from Photorhabdus laumondii* and PHL lectin (Jančaříková et al., 2017) from *Photorhabdus asymbiotica*. PLL lectin was the first experimentally characterized lectin from *P. laumondii*. PLL, PLL2, PLL3 and PHL all belong to the same class of lectins with a range of 62-81 % sequence identity. All of them fold as 7-bladed β-propellers in the monomeric form but they differ in oligomeric states, except PllA which is an α-galactoside-binding lectin that has 37 % sequence identity to LecA lectin from *Pseudomonas aeruginosa* (Beshr et al., 2017).

PLL is a homotetramer, PLL2 and PHL are homodimers and PLL3 folds as a monomer. All of them exhibit affinity towards ʟ-fucose and a rare *O*-methylated disaccharide 3,6-*O*-Me$_2$-ᴅ-Glcβ1-4(2,3-*O*-Me$_2$)-ʟ-Rhaα present in *Mycobacterium leprae* glycolipid PGL-1 (Hunter et al., 1982). Moreover, PHL lectin also exhibits affinity towards ᴅ-galactose and thus has two sets of unique binding sites within a monomeric unit.

It has been experimentally demonstrated that PLL lectin can bind to insect hemocytes and a nematode cuticle (Kumar et al., 2016) while PLL2 and PHL interfere with the insect immune system by activating the phenoloxidase system and modulate the reactive oxygen species levels in human blood (Eleftherianos et al., 2017; Fujdiarová et al., 2020). Specific functions of these lectins are still unknown but they have been suggested to play roles in overcoming the innate immune response at an early stage of infection.

The Unilectin database containing the PropLec module for the prediction of β-propeller lectins currently identifies 6 distinct classes based on conserved amino acid motifs. One of these classes (so-called PropLec7A) currently predicts 2981 PLL-like lectins in all species based on the conserved EVF amino acid sequence.

Structurally, PLL lectin is a seven-bladed β-propeller made of seven tandem repeats organized around a seven-fold pseudoaxis of symmetry. Each repeat (referred to as a W-motif) consists of a twisted 4-stranded (A-B-C-D) antiparallel β-sheet connected by short loops (Figure19a). Consecutive blades are connected by relatively long loops that emerge from the outside of the protein to the central cavity. The overall shape resembles a short cylinder with a diameter of approximately 45 Å and a height of 30 Å (Figure 19b). PLL folds as a homotetramer both in solution and in the crystal structure. The tetramer is generated by two pseudo-2-fold axes that create "side wall-to-side wall" and "bottom-to-bottom" associations (Figure 19c). Two disulphide bridges are involved in the tetrameric arrangement. One disulphide bridge formed by residues Cys-260 (β-sheet D of the W-motif 5 in all monomers) from adjacent monomers is involved in the "side wall-to-side wall" association (Figure 19c). The other disulphide bridge created by residues Cys-227 (located in the short loop between strands A and B of the W-motif 5 of all monomers) of adjacent monomers plays a role in the "bottom-to-bottom" association (Figure 19c).
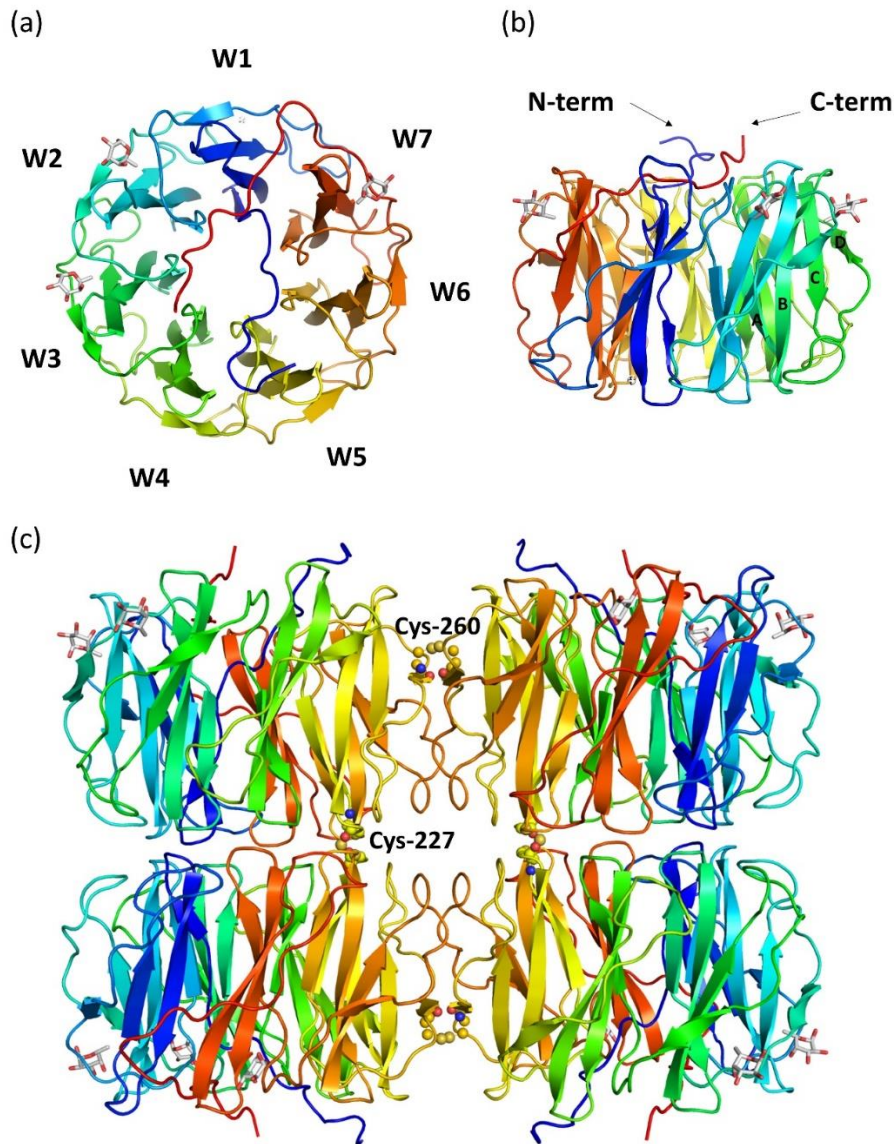
Figure 19: Cartoon representation of PLL lectin with fucose ligands. (a) PLL monomer, seven-bladed β-propeller with three fucose-binding sites. (b) Side view (c) PLL homotetramer with two intermolecular disulphide bridges formed by Cys-260 involved in the "side-to-side" association and Cys-227 involved in the "bottom-to-bottom" association. (PDB: 5C9P).

PLL lectin binds fucose with a millimolar affinity (Kumar et al., 2016). Structural comparison of all repeats revealed that PLL can have four binding sites. These data were also confirmed by isothermal titration calorimetry experiments where the binding stoichiometry was found to be 3.7 in (Kumar et al., 2016). In the crystal structure of the PLL/fucose complex (PDB: 5C9P), three fucose molecules were

observed in three binding sites, site II (between W-motifs 1 and 2), III (between W-motifs 2 and 3) and VII (between W-motifs 6 and 7) respectively (Figure 19a). Binding sites V and VI (between blades W4-W5 and W5-W6, respectively) are not available due to their involvement in the tetrameric association and site VII is not accessible due to the incorporation of the C-terminus between repeats W7 and W1. The fourth binding site (between repeats W3 and W4) has not been described structurally yet. In the crystal structure of the complex, the site was occupied by three water molecules and in the apo structure by trapped glycerol that was used as a cryoprotecting agent.

1.4 Neutrons in structural biology

Most of the macromolecular structures deposited in the Protein Data Bank (PDB) have been solved by X-ray diffraction and these currently account for more than 154 600 structures (March 2021, https://www.rcsb.org). X-ray crystallography has played an immense role in the determination of three-dimensional structures of small molecules to large macromolecular complexes that can be underlined by the highest number of Nobel Laureates with at least 25 Nobel Prizes that have been awarded for research involving crystallography (Galli, 2014). X-ray crystallography contributed to structural biology with a vast repertoire of 3D macromolecular structures, from DNA (Watson and Crick, 1953), photosynthetic reaction centre (Deisenhofer et al., 1985), ribosome (Ramakrishnan, 2002) to G-protein coupled receptors (Rosenbaum et al., 2009). In glycobiology, XMX has helped elucidate the structural basis of highly specific interactions between lectins and carbohydrates and thus plays an important role in structural glycobiology (Cordara and Krengel, 2013). According to recent reports dating to March 2021, the PDB database counted ~15 000 entries containing some carbohydrate residues, representing ~9.7% of all deposited structures (De Meirelles et al., 2020), most of them being glycoproteins with the partial structure of *N-* or *O-*glycans. The lectin structures from the PDB have been collected in the Unilectin3D database (https://unilectin.eu/unilectin3D/) (Bonnardel et al., 2019). Currently, there

are 2253 3D X-ray structures of lectins deposited in this database with over 60 % of them being complexed with a glycan. The key atoms in ligand-binding interactions are hydrogen atoms that form directional hydrogen bonds that are either direct or bridged by water molecules. Hydrogen atoms are rarely observable by X-ray crystallography and even then only at atomic resolution (~1 Å) (Ahmed et al., 2007; Fisher et al., 2012; Podjarny et al., 2004; Takaba et al., 2019). Since hydrogen atoms have only one electron, they scatter X-rays very weakly. Moreover, the visibility of mobile hydrogen atoms, which are often the most biologically relevant is observed only if high-resolution data is available. The visibility of hydrogen atoms drops off with increasing thermal atomic motion (B factors > 10 Å^2). It has been observed that even at an ultra-high resolution of 0.66 Å, only ~ 50% of all possible hydrogen atoms could be observed (Howard et al., 2004). This was linearly correlated with the temperature factors of atoms to which the hydrogen atoms were covalently bonded (Figure 20).

% of observed hydrogen atoms vs. B-factor

Linear fit. r=0.97
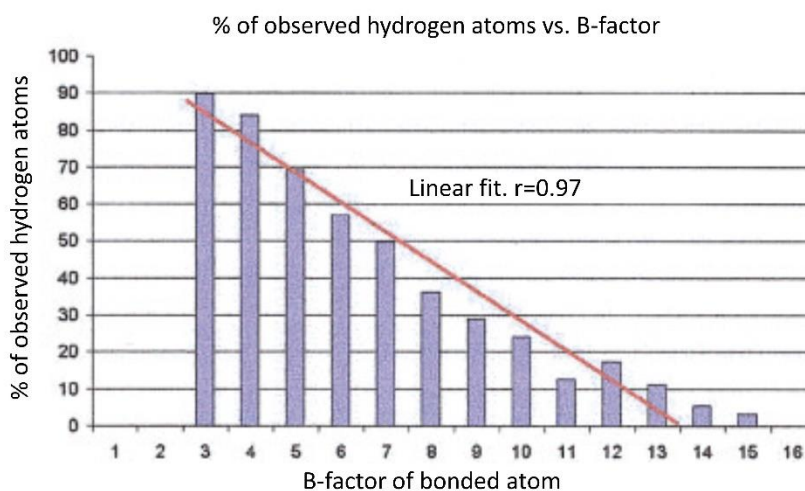
Figure 20: Histogram of percentage of observed hydrogen atoms as a function of the B-factor of the covalently linked heavy atom. Figure adapted from (Howard et al., 2004).

Neutron macromolecular crystallography (NMX), on the other hand, is an invaluable tool to directly locate hydrogen atoms in crystal structures and represents a complementary technique to X-ray crystallography (Niimura and Podjarny, 2011). To date, there have been only 177 macromolecular structures solved from neutron

diffraction data deposited in the PDB (March 2021). The reason behind such a small number of deposited structures lies, amongst other issues, in the fact that there is a limited number of neutron sources and instruments dedicated to NMX when compared to the synchrotron sources with several beamlines dedicated to XMX per synchrotron. Moreover, NMX is not used for the determination of novel structures because it takes weeks to months to complete structure determination (Niimura and Bau, 2008). It is rather used to answer specific biological questions including insights into enzyme mechanisms or ligand-binding where knowledge of the precise position of hydrogen atoms is of high importance.

The main issue is the much lower fluxes of neutron sources compared to synchrotrons. The effective neutron flux at a sample at a neutron diffractometer ranges from $\sim 10^6$ to 10^9 neutrons cm^{-2} s^{-1} at a high-flux reactor such as the Institut Laue-Langevin (ILL), while the X-ray flux from a rotating anode is $\sim 10^{10}$-10^{11} photons cm^{-2} s^{-1} and from synchrotrons is $\sim 10^{16}$ photons cm^{-2} s^{-1} (Meilleur et al., 2006). The low flux of neutrons must be compensated by a bigger sample size with typical crystal volumes of > 0.1-1 mm^3 and longer data collection times (days-weeks) when compared to the X-ray data collection at synchrotrons that can take less than a minute. The large crystal growth is a challenging process that requires a considerable amount of protein sample, time and dedication and is often the limiting step in NMX.

The fundamental difference between XMX and NMX lies in the way how X-rays and neutrons interact with matter. X-rays are scattered by electron clouds and the scattering power of atoms increases proportionally to the increasing atomic number and thus the number of electrons (Niimura and Podjarny, 2011). Neutrons are scattered by atomic nuclei and the neutron scattering power can be described by neutron scattering length, *b,* which varies between elements and even between isotopes of the same element. Since hydrogens have only one electron, they are weakly diffracted by X-rays. On the other hand, the coherent neutron scattering lengths *b* of hydrogen (^{1}H) and its isotope deuterium (^{2}H) are similar in magnitudes to other elements commonly found in proteins such as carbon, oxygen and nitrogen (Table 1). Hydrogens

and/or deuteriums can then be easily visualized in the neutron scattering length density maps at a typical resolution of 1.5-2.5 Å of most protein structures (Meilleur et al., 2006).

Table 1: Neutron coherent scattering length and neutron incoherent cross-section and X-ray scattering factors of elements found in biological macromolecules.

Isotope	Atomic number	Neutron incoherent cross-section (Barns)	Neutron coherent scattering length (10^{-12} cm)	X-ray scattering factors (10^{-12} cm)	
				$\sin \theta = 0$	$\sin \theta/\lambda = 0.5$ Å^{-1}
^{1}H	1	80.27	-0.374	0.28	0.02
^{2}H (D)	1	2.05	0.667	0.28	0.02
^{12}C	6	0.00	0.665	1.69	0.48
^{14}N	7	0.49	0.937	1.97	0.53
^{16}O	8	0.00	0.580	2.25	0.62
^{24}Mg	12	0.00	0.549	3.38	1.35
^{32}S	16	0.00	0.281	4.50	1.90
^{39}K	19	0.25	0.379	5.30	2.20
^{55}Mn	25	0.40	-0.373	7.00	3.10
^{56}Fe	26	0.00	1.010	7.30	3.30

Another advantage of neutrons is that unlike X-rays, which are a form of ionizing radiation that cause radiation damage to protein crystals, cold neutrons used for neutron diffraction experiments have millielectronvolt energies that are too small to induce radiation damage (Niimura and Podjarny, 2011). As a practical consequence, neutron data can be collected at ambient temperature ($\sim$ 20 °C) without the need to cryo-cool the protein crystal. Since cryo-cooling and temperature can both perturb the protein structure, room temperature data collection can be beneficial when drawing biologically relevant conclusions as it is closer to physiological conditions. Moreover, the same crystal used for the neutron diffraction experiment can be subsequently used for X-ray diffraction experiments.

The heavy isotope of hydrogen ^{2}H, deuterium, has neutron scattering characteristics that make it beneficial to exchange hydrogens for deuteriums in a

protein molecule (Helliwell, 2020). The inherently large neutron incoherent scattering cross-section of ^{1}H (80 barns) produces high background in diffraction images that complicates measurements of Bragg reflections at high resolution. Deuterium has 40 times smaller incoherent neutron scattering cross-section and thus replacement of hydrogen by deuterium reduces the background substantially (Figure 21).
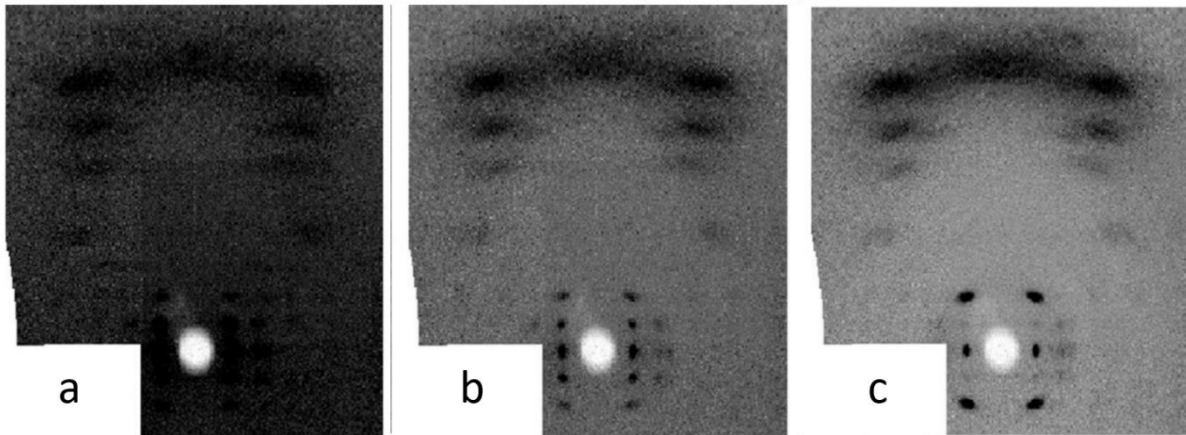
Figure 21: Neutron fiber diffraction pattern of A-DNA measured at D19 instrument at ILL. The humidity of the sample environment is changed from H_2O to D_2O in a series from a to c. The decrease in hydrogen incoherent scattering is evident with the improvement of the visibility of the coherent diffraction peaks. Figure adapted from (Haertlein et al., 2016).

Moreover, due to the negative sign of the neutron scattering length of ^{1}H (-3.74 fm), whose magnitude is about half of the neutron scattering length of C which is positive (6.65 fm), NMX (at the resolutions typically achieved of ~ 1.5-2.5 Å) cannot resolve individual scattering peaks from hydrogen and carbon atoms in -CH_2 groups. As a practical result, aliphatic carbon atoms of protein main chains and sugar rings often suffer from cancellation effects in the neutron scattering length density maps (Koruza et al., 2019), which complicates the analysis of neutron maps. To minimize the negative effects caused by ^{1}H, H/D replacement can be performed either partially by soaking the protein crystal in a D_2O-based solution or by the production of a fully deuterated (perdeuterated) sample (Haertlein et al., 2016).

1.4.1 Neutron structures of protein-ligand complexes

Neutron crystallography has deciphered many enzymatic mechanisms including those in carbohydrate-processing enzymes that could not be observable by X-ray crystallography (Fisher et al., 2011; Gerlits et al., 2017; Kovalevsky et al., 2020; Langan et al., 2014). NMX has been further employed to provide insights into ligand binding and solvent structures. The direct determination of the protonation state of amino acids involved in the active sites such as histidines, aspartic acid, glutamic acid and lysines is of crucial importance (Ahmed et al., 2007). Structural information on the ligand-binding site including hydrogen network, the directionality of water molecules and protonation state of amino acid is essential for rationalizing drug design of novel potent inhibitors. Here will be presented some examples of protein structures where NMX played an important role in answering specific biological questions.

1.4.1.1 Lysozyme

Lysozyme is an enzyme that catalysis hydrolysis of glycosidic bonds that link *N*-acetylneuraminic acid and *N*-acetylglucosamine in polysaccharides of the bacterial cell wall. Lysozyme has been extensively studied by X-ray and neutron crystallography (Niimura et al., 2004, 1997). The enzyme mechanism involves a proton transfer that was shown by neutron diffraction studies of lysozyme crystals grown at different acidities. At an acidic pD of 4.9, the carboxyl group of the catalytic residue Glu35 is protonated and the proton is then transferred to the oxygen atom of the glycosidic bond during the hydrolysis (Figure 22). It has been demonstrated that at higher pH of 8.0 the residue Glu35 is deprotonated which is in agreement with the reduced catalytic activity (Davies et al., 1969).
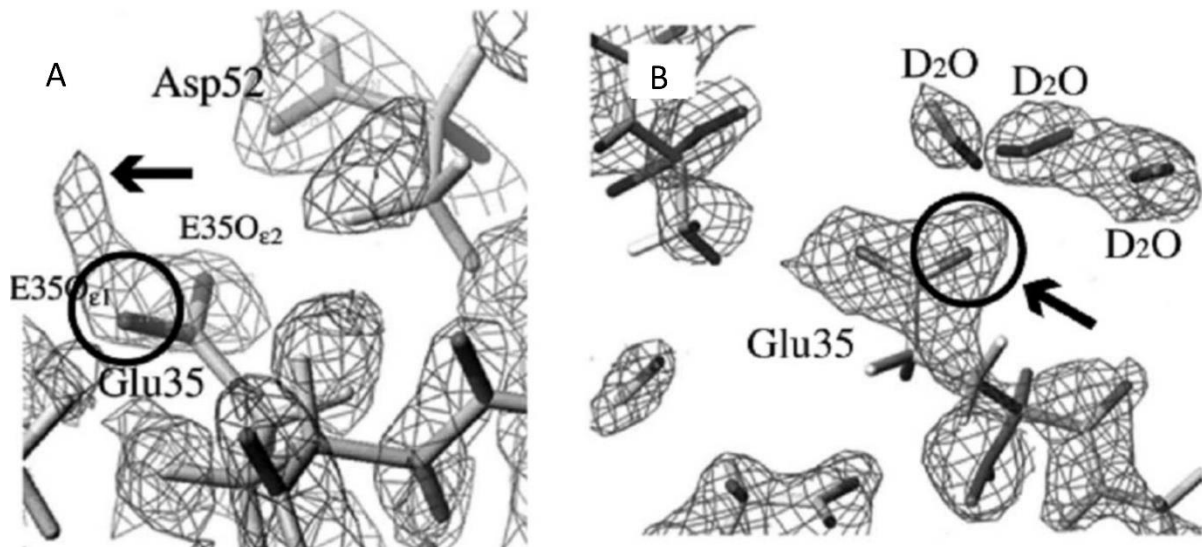
Figure 22: The active site of lysozyme. The $2F_o$-F_c nuclear scattering length density map around the carboxyl group of the catalytic residue Glu35 at pD of 4.9 (A) and 8.0 (B), respectively. The arrow in A shows a positive neutron density extending from atom E35O$_{\varepsilon 1}$ indicating its protonated form at active (acidic) conditions. Figure adapted from (Niimura and Bau, 2008).

1.4.1.2   Xylanase

Xylanases belong to the glycoside hydrolase family of enzymes that apply acid/base chemistry to enzymatically cleave complex polysaccharides. Xylanase enzymatically degrades xylan, one of the most abundant components of hemicellulose, to smaller oligosaccharides (Paës et al., 2012). Two catalytic glutamic acid residues play a role in catalysis, one acts as a general acid, initially protonated and the second is negatively charged and serves as a nucleophile. Based on the neutron structures of xylanase it has been demonstrated that the Glu general acid side chain is in a dynamic equilibrium between the upward and downward conformations. The downward conformation with higher pKa facilitates its protonation (Figure 23) and the upward conformation with lower pKa facilitates the delivery of the proton to the glycosidic oxygen (Wan et al., 2015).
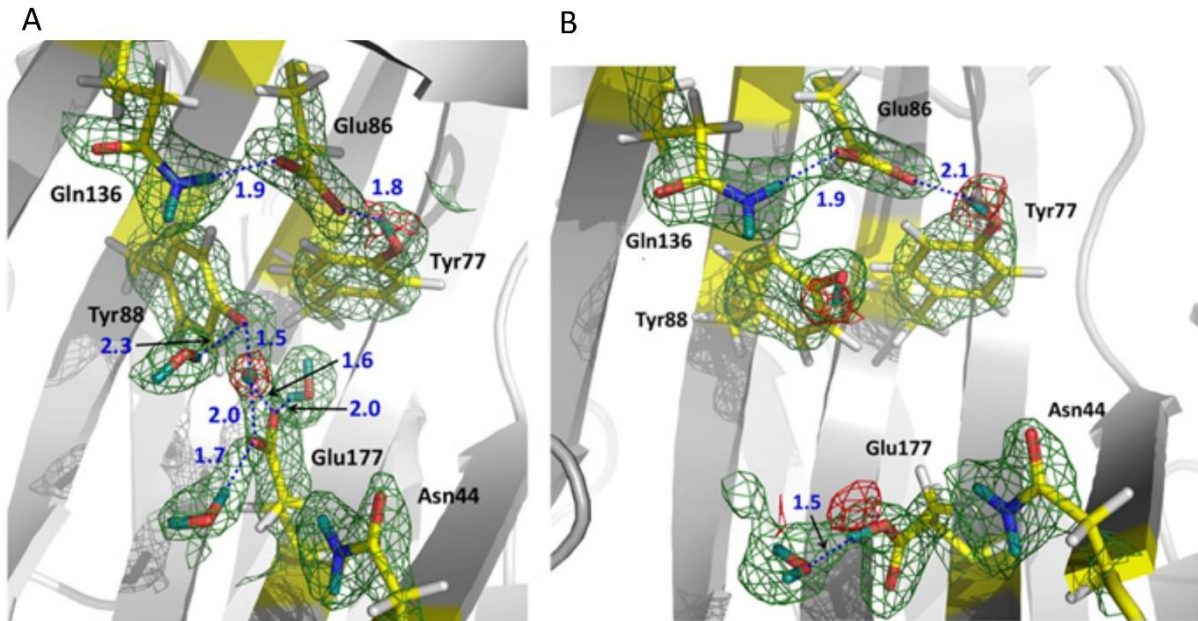
Figure 23: The active site of Xylanase II from *Trichoderma reesei*. The positive neutron scattering length density map (green mesh) is contoured at 1.2σ. F_o-F_c omit map (red mesh) is contoured 3.5σ (A) and 2.5σ (B). Hydrogen and deuterium atoms are grey and cyan respectively. Hydrogen bonds are shown as blue dotted lines. (A) Structure at a basic pD of 8.9 showing the deuterium atom participating in the hydrogen bond between Glu177 and Tyr88. (B) Under acidic conditions with pD of 4.8, Glu177 side chain rotates 90° about to assume a downward conformation. Figure adapted from (Wan et al., 2015).

1.4.1.3   Xylose isomerase

Xylose isomerase is a metalloenzyme that catalysis the conversion of aldose sugars D-xylose and D-glucose into keto sugars D-xylulose and D-fructose, respectively. It also catalyzes the slow conversion of some monosaccharides into their C-2 epimers (Langan et al., 2014). Xylose isomerase is studied thoroughly because of its industrial applications in the production of fructose corn syrup and biofuels. Four neutron structures of xylose isomerase, both in native and ligand-bound forms provided insights into the reaction mechanisms that occur in three stages, ring-opening, isomerization and ring-closure (Figure 24) (Glusker et al., 2010). Specifically, it was observed that the catalytic residue His54 is doubly protonated throughout the reaction and donates its hydrogen to the sugar O5 to promote the sugar opening. The metal-bound water is

deprotonated to OD⁻ during the isomerization and Lys289 residue is neutral before ring-opening and gains a proton afterwards becoming positively charged possibly helping in the proton transfer during isomerization. Moreover, the movement of one metal ion might facilitate the isomerization as it brings the substrate closer to the catalytic water (Kovalevsky et al., 2010). In another neutron study, it was demonstrated that L-arabinose isomerization and epimerization to L-ribulose and L-ribose, respectively occurs via similar mechanisms to the ones previously described for D-glucose and D-xylose (Langan et al., 2014).
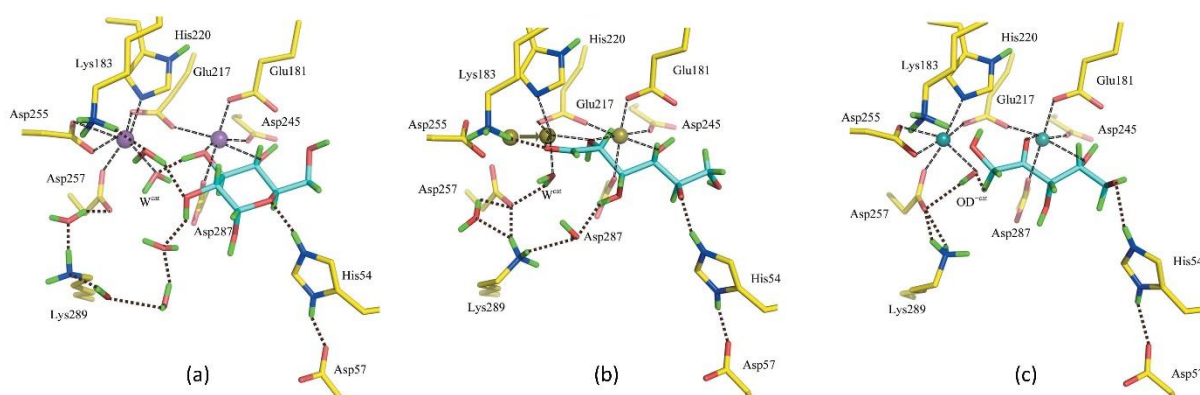
Figure 24: Intermediate stages along the D-glucose isomerization reaction catalyzed by xylose isomerase based on the neutron structures. (a) Cyclic D-glucose bound. (b) Linear D-glucose bound after the ring-opening. (c) Linear D-xylulose product after the isomerization step has occurred. The substrate is represented as cyan sticks and metal ions as spheres. Metal coordination is represented by black dash lines and hydrogen bonds are represented as dotted brown lines. Deuterium atoms are coloured green. Figure is taken and modified from (Glusker et al., 2010).

1.4.1.4   HIV-1 protease

HIV-1 protease catalyzes the hydrolysis of peptide bonds of Gag and Gag-Pol polyproteins during the maturation stage of the HIV-1 replication cycle and is a drug target for HIV/AIDS therapy (Lv et al., 2015). It is a homodimeric aspartic protease that utilizes two close aspartic residues in hydrolysis. One protonated Asp carboxylic side chain acts as a general acid and donates a proton to the carbonyl oxygen of the peptide bond. The other deprotonated Asp acts as a general base that abstracts

a proton from the catalytic H_2O. The room temperature neutron structure of the enzyme complexed with the clinical drug darunavir allowed observation of changes in H atom positions in the catalytic site and protonation states by altering pH in the crystal (Gerlits et al., 2016). The two-proton transfer from the inner to the other carboxylic oxygen of the aspartic pair was stable only at a lower pH of 4.3 (Figure 25). The proton transfer is triggered by the long-range electrostatic effects of four distant residues undergoing protonation at low pH (Kovalevsky et al., 2020).
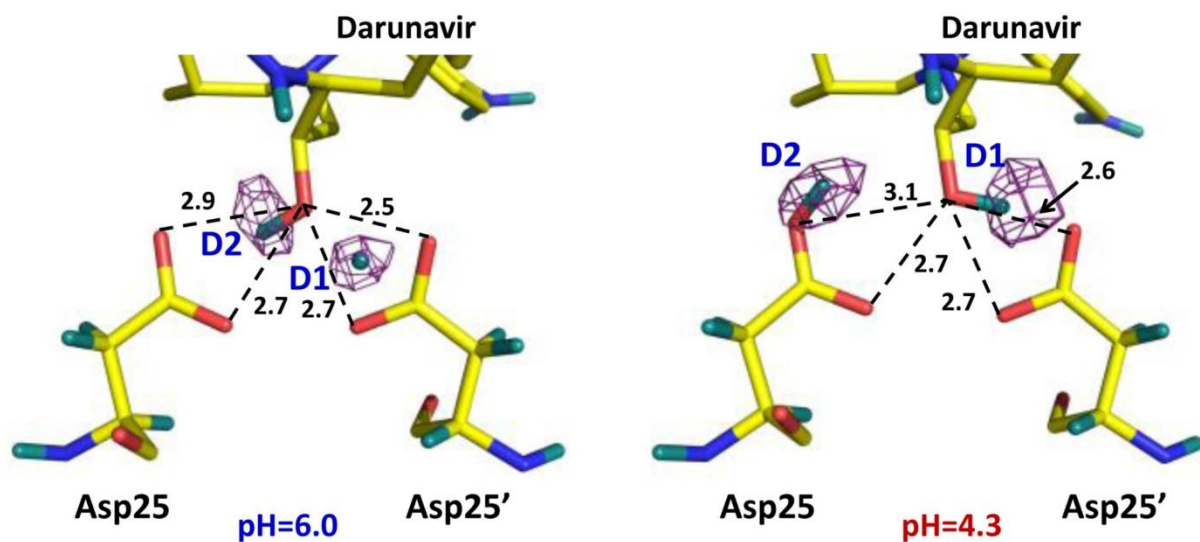
Figure 25: HIV-1 protease catalytic site with a clinical drug darunavir at pH 6.0 (left) and pH 4.3 (right). The omit F_o-F_c difference neutron map (purple mash) is contoured at 3σ. Deuterium atoms are coloured teal. The figure was taken and modified from (Kovalevsky et al., 2020).

1.4.2   Neutron structures of lectins

To date, there have been only two lectins studied by neutron protein crystallography, namely concanavalin A and galectin-3C.

1.4.2.1   Concanavalin A

Concanavalin A (ConA) is a 25 kDa legume lectin extracted from jack bean seeds with specificity towards mannose and glucose and that requires Ca^{2+} and Mn^{2+} ions for carbohydrate-binding (Hardman and Ainsworth, 1972).

Crystals of H/D-exchanged ConA were used for neutron diffraction studies. ConA was used as a model system in the development of neutron Laue data collection with one of the first neutron protein structures solved from Laue diffraction data obtained at the LADI diffractometer to 2.75 Å resolution (Habash et al., 1997). Already in 1997, Prof. Helliwell and co-workers proved that neutron Laue data can be collected from smaller crystals than the ones used for monochromatic neutron data collection (Habash et al., 1997) and in a shorter time. In this study, the authors collected neutron data from a 4.8 mm³ crystal that grew over a period of four years. They observed that Asp28 residue was protonated which was in agreement with their previous 0.94 Å X-ray crystal structure of the saccharide-free ConA. ConA is a lectin and should have no enzymatic activity, although it contains Asp28-Glu8 pair that is reminiscent of an Asp-Glu pair in a catalytic site of glycosyl hydrolases. It was thus suggested that this lectin could have a functional relationship with glycosyl hydrolases such as lysozyme. In the neutron structure, the proton on Asp28 was confirmed to be H and was not exchanged by D concluding that there was no role of this residue in the catalytic activity. Later report extended the resolution to 2.4 Å, which allowed observation of triangular-shaped contours of water molecules (Habash et al., 2000). The room temperature neutron structure of the apo ConA revealed positions of the water-bound D atoms in the carbohydrate-binding site and in the vicinity of the metal ion-binding site (Habash et al., 2000). Three water molecules occupied the carbohydrate-binding site establishing hydrogen bonds with the protein similar to those of hydroxyl groups in the ligand-bound complex. Four water molecules were observed in the metal-binding site pointing away from the metal ions, two coordinating the Ca^{2+} and two coordinating the Mn^{2+} atom (Figure 26).
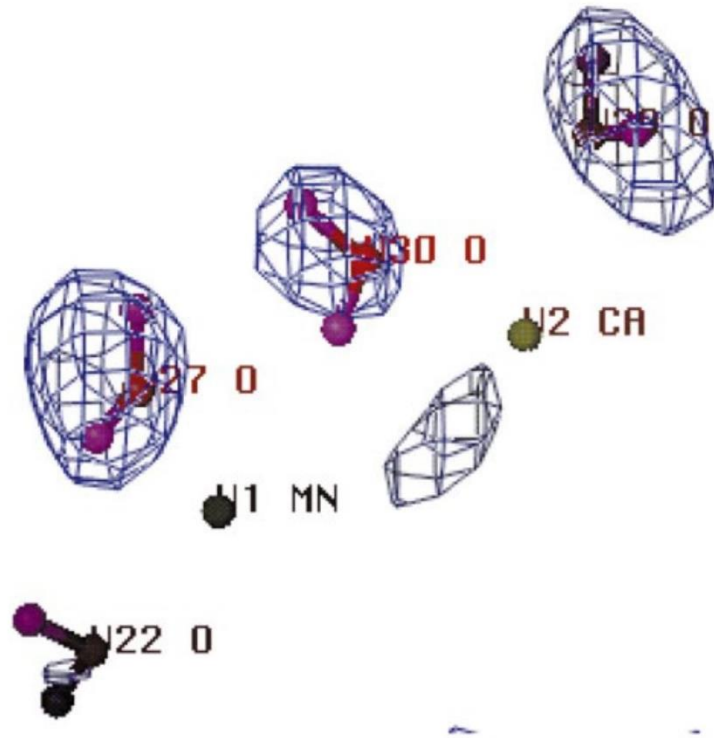
Figure 26: The F_o-F_c omit neutron map (blue mesh) contoured at 3σ showing four water molecules in the metal-binding site of ConA. Two water molecules coordinate the Mn^{2+} atom and the other two coordinate the Ca^{2+} atom. Figure is taken from (Habash et al., 2000).

The 15 K neutron structure of saccharide-free concanavalin A proved that it is possible to successfully cryo-cool large crystals (5.6 mm^3 and 1.6 mm^3) typical for NMX and opened the door to cryo-NMX that could be used for freeze-trapping structural studies (Blakeley et al., 2004). Furthermore, the authors observed that the crystal quality was the worst in the area where the crystal was first exposed to liquid nitrogen and improved with the distance from the point of contact (Blakeley et al., 2004). Moreover, the 15 K neutron structure (Blakeley et al., 2004) also provided a better defined hydrogen-bonding network of water molecules in the saccharide-binding site of the apo protein (Figure 27) compared to the room temperature neutron structure (Habash et al., 2000).
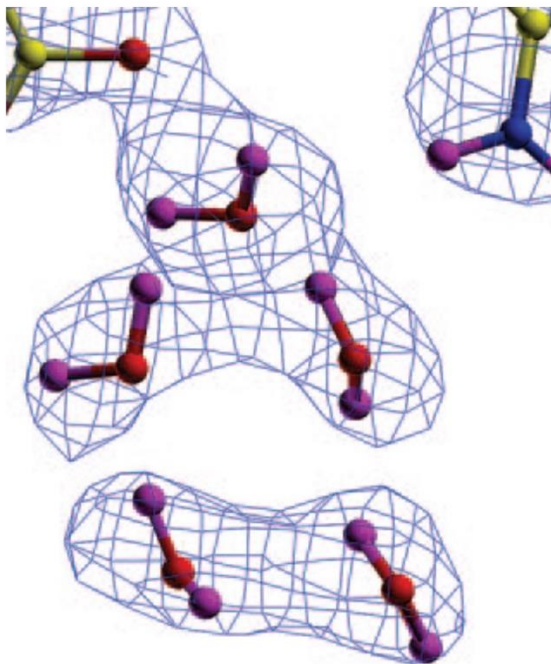
Figure 27: Five water molecules (oxygen and deuterium atoms in red and purple, respectively) identified in the 15 K neutron structure of the carbohydrate-binding site of saccharide-free Concanavalin A at 2.5 Å. The $2F_o$-F_c neutron map is shown as blue mesh and contoured at 1.5σ. Figure is taken from (Blakeley et al., 2004).

The researchers also attempted to collect neutron data on the ConA/methyl-α-D-glucopyranoside complex with a rather large unit cell with a cell edge of 167.8 Å (Kalb et al., 2001). The data collection was unfortunately interrupted by a reactor technical problem. A recent neutron study of concanavalin A/mannobiose (Manα1–2Man) complex has shown that the binding of the ligand altered both the protonation state and interactions of residues close and distant to the binding site (Gerlits et al., 2017). The binding of mannobiose to ConA also induces changes in the nature of a hydrogen bond between Asp28 and Glu8 with the proton being equidistant between two carboxylic oxygens, possibly participating in a low-barrier hydrogen bond that would not be visible by X-ray crystallography (Figure 28).
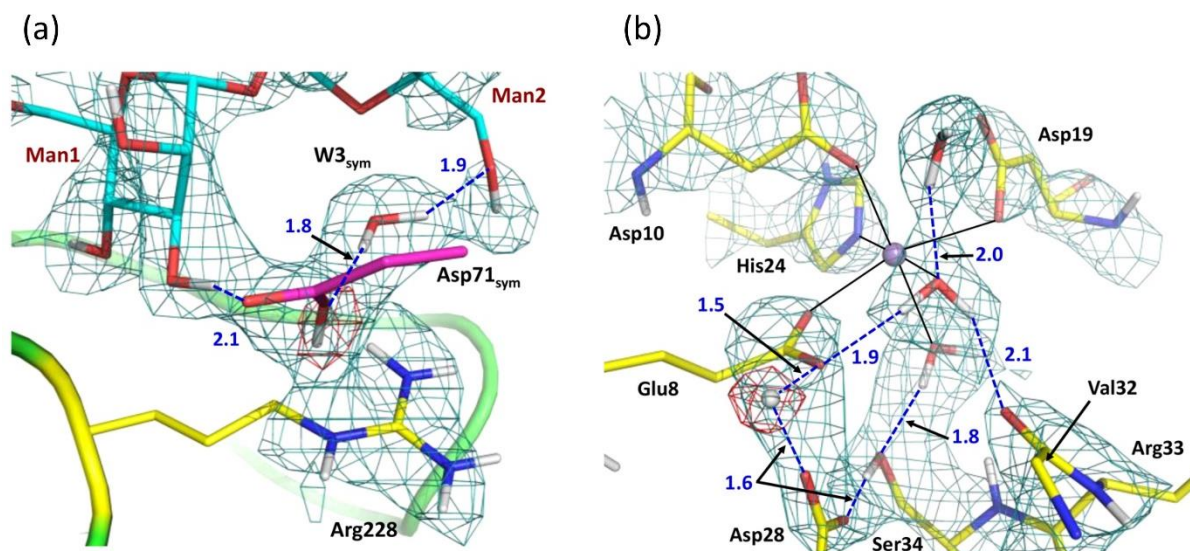
Figure 28: Neutron structure of the Concanavalin A/mannobiose complex. (a) Interactions between mannobiose and Asp71 of the symmetry-related ConA. The $2F_o$-F_c neutron map (teal mesh) is contoured at 1.5σ. The F_o-F_c omit neutron map (red mesh) is contoured at 3σ. Hydrogen bonds are shown as blue dashed lines. (b) Mn^{2+} (violet sphere)-binding site and the protonation of As28-Glu8 pair. Deuterium atom (grey sphere) in the low-barrier hydrogen bond between Asp28 and Glu8. The $2F_o$-F_c neutron map (teal mesh) is contoured at 1.2σ. The F_o-F_c omit neutron map for D atom (red mesh) is contoured at 3σ. Hydrogen bonds are shown as blue dashed lines. Figure is taken and modified from (Gerlits et al., 2017).

1.4.2.2   Galectin-3

Galectin-3 is a human β-galactoside-binding lectin that has important roles in cell adhesion, membrane trafficking, tumour cell adhesion, proliferation, metastasis and angiogenesis (Sciacchitano et al., 2018). The CBD of galectin-3 is solvent-exposed and binds lactose via numerous hydrogen bonds with a significant number of water molecules that participate in the binding (Saraboji et al., 2012). Galectin-3 is considered a drug target and many inhibitors have been developed that compete with its natural ligand (Collins et al., 2012; Delaine et al., 2016). The directionality of hydrogen bonds involved in the binding is of importance for future drug design.

Neutron crystallographic studies have been carried out on the C-terminal domain of galectin-3 (galectin-3C) responsible for carbohydrate recognition (Manzoni et al., 2018). Perdeuterated galectin-3C was produced in a high amount (20-50 mg per

liter of culture) using 100 % D_2O M9 minimal medium with deuterated glycerol-d_8 as a carbon source (Manzoni et al., 2016). Neutron data for galectin-3C in the apo form, lactose- and glycerol-bound complexes were collected at three different neutron sources (Manzoni et al., 2016). The authors used a feeding technique to grow large crystals of perdeuterated galectin-3C with volumes up to 1.8 mm^3 that improved the crystal quality over macro-seeding technique. The authors also successfully merged data sets collected at different sources which improved the data statistics (Manzoni et al., 2016). The room temperature X-ray/neutron jointly refined structure of the lactose complex revealed an extensive hydrogen-bonding network between the ligand and the protein (Manzoni et al., 2018). The directionality of three key hydroxyl groups of lactose (4-OH and 6-OH hydroxyl groups of galactose and the 3-OH hydroxyl group of glucose moiety) that participate in the binding could be unambiguously determined (Figure 29). In the saccharide-free neutron structure, these three hydroxyl groups are replaced by water molecules while in the glycerol-bound complex, three heavy atoms of perdeuterated glycerol mimic the inner part of the galactose moiety of lactose. The study highlighted the fine-tuning of hydrogen bonding interactions in the carbohydrate-binding site. The protonation of histidine residues could also be determined from the neutron structure analysis. Specifically, His158 was observed to be in the singly protonated ND1-tautomeric form crucial for the sugar recognition (Manzoni et al., 2018)
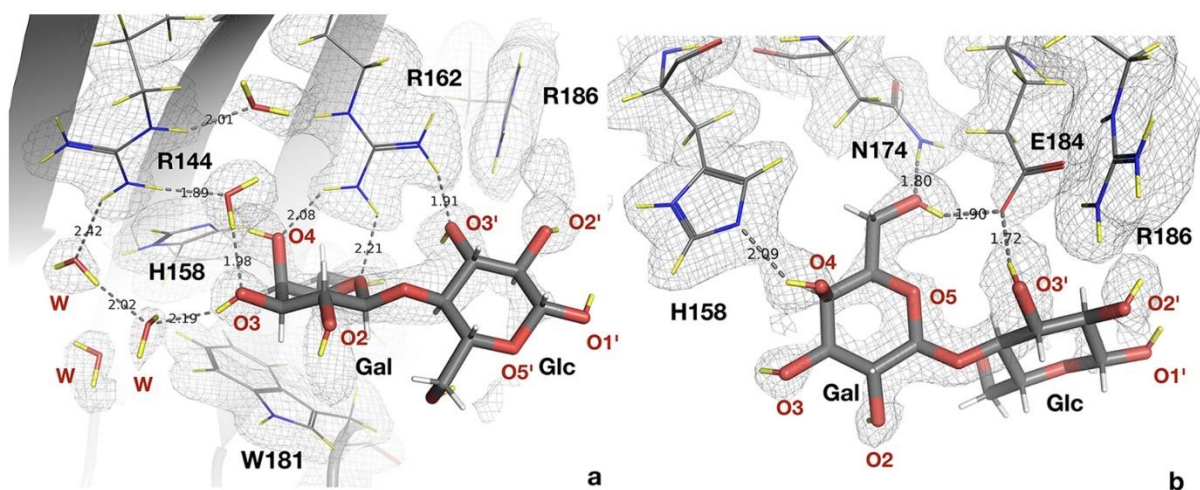
Figure 29: Neutron structure of the galectin-3C/lactose complex. The $2\mathrm{m}F_o$-DF_c neutron map is shown as grey mesh and contoured at 0.6σ. Lactose is depicted as thick grey sticks. Figure taken from (Manzoni et al., 2018).

In a recent study, a combination of X-ray and neutron refinement with quantum refinement of the neutron structure of the galectin-3C/lactose complex was carried out (Caldararu et al., 2019). Orientation of solvent molecules as well as deuterium atoms of the ligand were improved by quantum refinement to form more favourable hydrogen-bonding pattern resulting in an improved structure.

<div align="right">

2

Methodology

</div>

The objective of this chapter is to give a general idea of methods involved in the neutron protein crystallographic studies including perdeuteration, obtaining large crystals and neutron data collection. A detailed description of standard methods used in structural biology from molecular cloning, protein expression, purification and X-ray crystallography is avoided. Protocols with specific parameters of methods used in studies described in this thesis are given in Materials and methods section of each chapter.

2.1 Perdeuteration

Hydrogen atoms represent about 50% of all atoms in a protein and are also present in the surrounding solvent as H_2O. Moreover, out of all hydrogen atoms in a protein molecule, only ∼ 25% are on exchangeable sites (titratable groups), i.e. hydrogen atoms bound to nitrogen, oxygen and sulphur atoms. These can be exchanged either prior to crystallization by growing crystals of hydrogenated protein in deuterated mother liquor or by vapour diffusion in D_2O-containing solution after the large crystals have been obtained. To study the interactions between lectins and carbohydrates by NMX, it is highly advantageous to perdeuterate both the proteins and ligands, where all hydrogen atoms are replaced by deuteriums (Haertlein et al., 2016). The use of perdeuterated molecules in neutron diffraction experiments provides clearer visualization of nuclear scattering length density maps as well as eliminating the large incoherent neutron scattering cross-section by ^{1}H (80 barns), thus vastly reducing the background and enhancing the signal-to-noise ratio. In order to facilitate the large-scale production of perdeuterated biomolecules, dedicated facilities for macromolecular deuterium labelling

have been developed at neutron sources throughout the world. These include ISIS deuteration facility at ISIS (UK), the Deuteration laboratory at ILL (France), Jülich Centre for Neutron Science at Forschungszentrum Jülich institute (Germany), Australian National Deuteration Facility at ANSTO (Australian Nuclear Science and Technology Organisation), Deuteration laboratory at J-PARC MLF (Materials and Life Science Facility within Japan Proton Accelerator Research Complex) and Deuteration and Macromolecular Crystallisation (DEMAX) platform at ESS and Lund University (Sweden).

2.1.1 Adaptation to D_2O-based minimal medium and deuterated glycerol-d_8

The goal of the adaptation is to achieve good and consistent growth rates of *E. coli* cultures in a deuterated medium. The first step in the production of a perdeuterated protein or a perdeuterated sugar is the adaptation of the *E. coli* cells to the deuterium oxide-based minimal medium with a defined composition. Such medium is rather simple and consists of D_2O, mineral salts and a deuterated carbon source such as glycerol-d_8.

Firstly, a single colony grown on a solid Luria-Bertani (LB) medium supplemented with the appropriate antibiotic is used to inoculate an LB culture and is grown overnight at 37 °C at constant shaking. The overgrown LB culture is then used to inoculate the hydrogenated minimal medium with the hydrogenated glycerol and is grown overnight at 37 °C. Finally, the overgrown culture is used to inoculate the deuterated minimal medium. This step of passaging of the deuterated culture is repeated several times until the cells reach values as close as possible to the growth rate in the hydrogenated culture. At the end of the adaptation, the final culture of approximately 100-150 mL is used to inoculate the deuterated medium in a bioreactor used for fermentation (Figure 30).
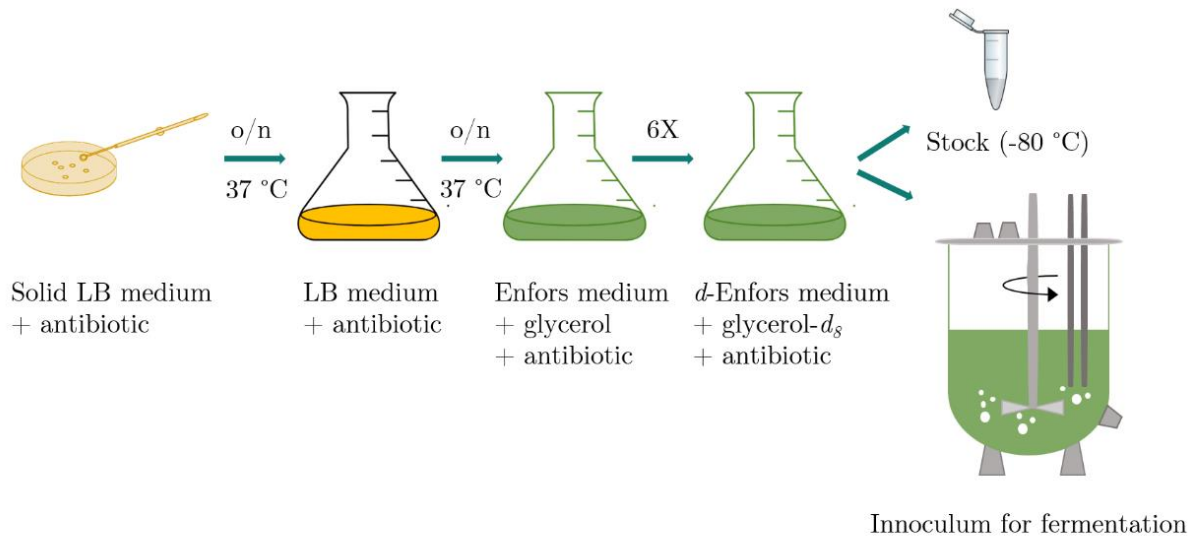
Figure 30: Schematic illustration of the steps involved in the adaptation of the *E. coli* cells to the deuterated minimal medium used for high cell-density fermentation.

2.1.2 High cell-density culture

High cell-density culture (HCDC) was carried out to maximize the production of the perdeuterated PLL and LecB lectins and the perdeuterated monosaccharide L-fucose.

Most fermentation strategies are based on a batch phase followed by a fed-batch phase. In a batch phase, cells are inoculated and allowed to grow consuming all nutrients and glycerol present in the culture medium. An important point is to use limiting amount of carbon source during growth in order to keep the growth rate below a critical value, thus limiting the amount of secreted products (such as acetate) which can inhibit cell growth. The fed-batch phase starts after exhaustion of the carbon source from the initial culture medium. Cells are then continuously fed with a feeding solution composed of 12 % (w/v) glycerol. The carbon source is supplied in a limiting way so that the cell-specific growth rate does not reach its maximal value.

Fermentation is carried out in a 3-L (2.3-L working volume) bioreactor (Infors) under control of the IRIS software (Figure 31). The deuterated minimal medium (around 1.2 L) is inoculated with a freshly prepared adapted deuterated preculture

(around 100-150 mL). Various parameters are monitored and controlled during the high cell-density culture such as pD, temperature, amount of dissolved oxygen (pO_2), gas flow rate of sterile air, stirring speed, addition of base solution (typically NaOD), addition of feeding solution and optical density at 600 nm. The protein expression is induced with IPTG and expression conditions such as the IPTG concentration and temperature are dependent on experiments carried out ahead of the fermentation with each construct. At the end of the fermentation when the carbon source is depleted, cells are harvested by centrifugation and the deuterated wet cell paste is weighed and stored at -80 °C. In the case of fucose production, the sugar molecule was secreted into the medium and the final supernatant containing the deuterated fucose was sterile filtered and stored at -20 °C.
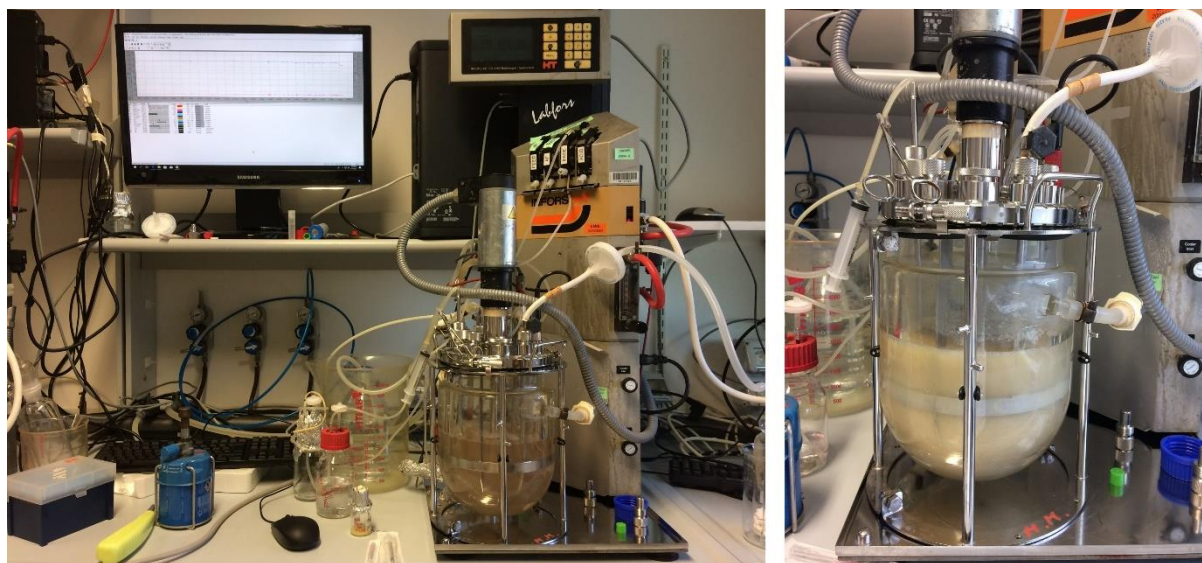
Figure 31: Equipment including the 2.3-L bioreactor (Infors) used for high-cell density culture.

2.1.3   Kanamycin resistance marker

The antibiotic resistance gene is an essential part of the expression vector and acts as a selection marker. A product of this gene is a protein that is able to modify or degrade the antibiotic molecule in a way that inhibits its effect. Cells with plasmids containing

the antibiotic resistance gene gain a survival advantage since they can grow in an antibiotic-containing culture medium. Historically, the most used antibiotic in bacterial expression studies of recombinant proteins is ampicillin. The ampicillin resistance gene codes for an enzyme β-lactamase that can hydrolyze the β-lactam ring of the ampicillin. Since β-lactamase is secreted into the periplasmic space of the bacterium, it can often leak into the culture medium and destroy all ampicillin. This is especially true when high cell-density growth in deuterated medium is carried out. Bacteria that lose the expression plasmid could replicate faster and out-populate bacteria containing the plasmid. In high cell-density cultures, such a situation can cause severe problems in production yields of the recombinant protein. Therefore, in HCDC and using deuterated medium, it is advised to use a kanamycin resistance gene coding for the aminoglycoside-3'-phosphotransferase that is expressed in the cytosol and does not cause similar problems.

2.2 Large crystal growth

Crystallogenesis is often the major bottleneck for NMX. Obtaining a crystal that is big enough for collecting good quality neutron diffraction data is a complicated process that requires a considerable amount of sample, patience and dedication. Macromolecular crystallization is a very complex process that is still rather empirical in nature and demands a lot of research, identification and optimization of the parameters that influence crystal formation and yield high-quality crystals. With increasing knowledge of the fundamental properties, we can build rational approaches that help in crystallogenesis. Fortunately, for most of the NMX experiments, the protein of interest has already been crystallized for X-ray crystallographic studies and the crystallization conditions are known. These can be used as a starting point to guide the crystallization trials. It has been observed though, that crystallization of perdeuterated proteins as well as the use of D_2O buffers with specific biophysical

properties, influence the crystallization process, therefore requiring optimization of the crystallization conditions (Koruza et al., 2018).

2.2.1   Phase diagram

A protein crystallization phase diagram describes a state of aqueous protein solution as a function of various conditions. The phase diagram is sharply divided into two regions, an undersaturated region and a supersaturated region that are separated by a solubility curve that defines the maximum solubility of the protein at a specific concentration of the crystallization agent. The undersaturated region represents a stable solution phase with protein and crystallization agent concentrations that are too low to induce nucleation. Crystals introduced into this region will dissolve. The supersaturated region can be divided into three zones, the metastable zone and the labile zones divided by a supersolubility curve and a precipitation zone (Figure 32).
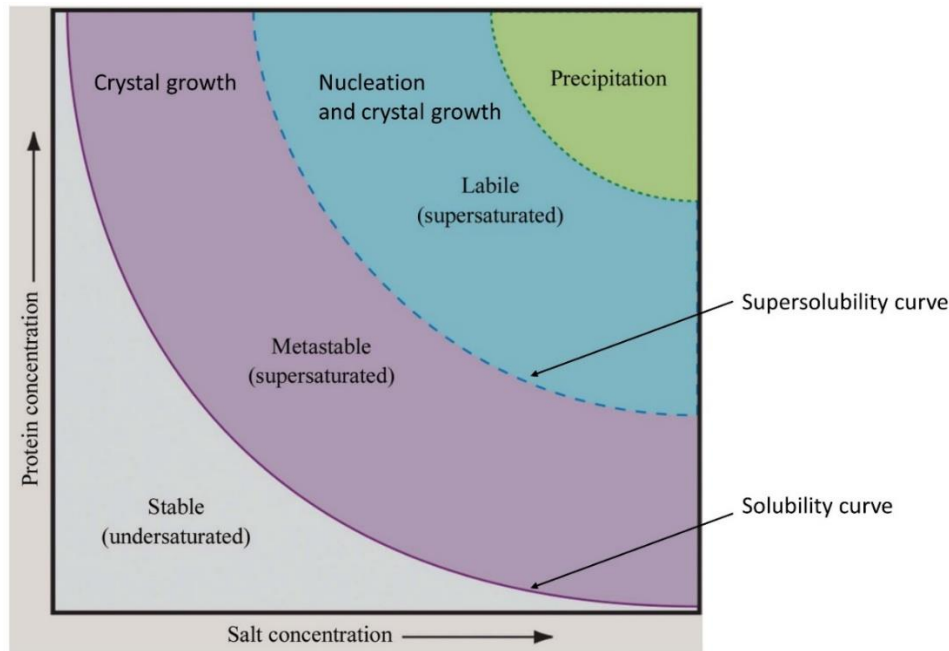
Figure 32: A typical phase diagram for crystallization of macromolecules. Figure is taken and modified from (McPherson and Gavira, 2014).

In the metastable zone, nuclei will develop into crystals that can grow but no nucleation occurs. In the labile zone, both nucleation and crystal growth are expected to happen. At very high supersaturation lies a precipitation zone where protein precipitates as an amorphous material and the formation of crystals is not favoured. A large single crystal grows in the metastable zone and if a seed is introduced into this zone, the protein molecules will concentrate onto the crystal and support its growth. There are several protein crystallization methods, the objective of which is to render the solution supersaturated and reduce the solubility of the protein so it comes out of solution in an ordered manner. The vapour diffusion method, batch method and the dialysis are among the most employed.

2.2.2   Vapour diffusion crystallization method

The most common protein crystallization technique is the vapour diffusion method that can be conducted in two set-ups, the hanging-drop and the sitting-drop (Figure 33). The basic principle of the hanging-drop vapour diffusion method is that a droplet containing the protein solution is mixed with the crystallization agent, usually salts or polymers, and suspended by surface tension on the cover slide. The slide with the droplet is placed above a mother liquor with a higher concentration of the crystallization agent and the droplet is allowed to equilibrate in a sealed chamber. The concentration gradient drives the water to evaporate from the drop to the reservoir. The concentration of the drop increases reaching a supersaturated state and at an optimal concentration of the protein and crystallization agent, the crystal formation occurs. The sitting-drop set-up works on the same principle but with a drop that can hold a larger volume and is thus better suitable for large crystal growth.
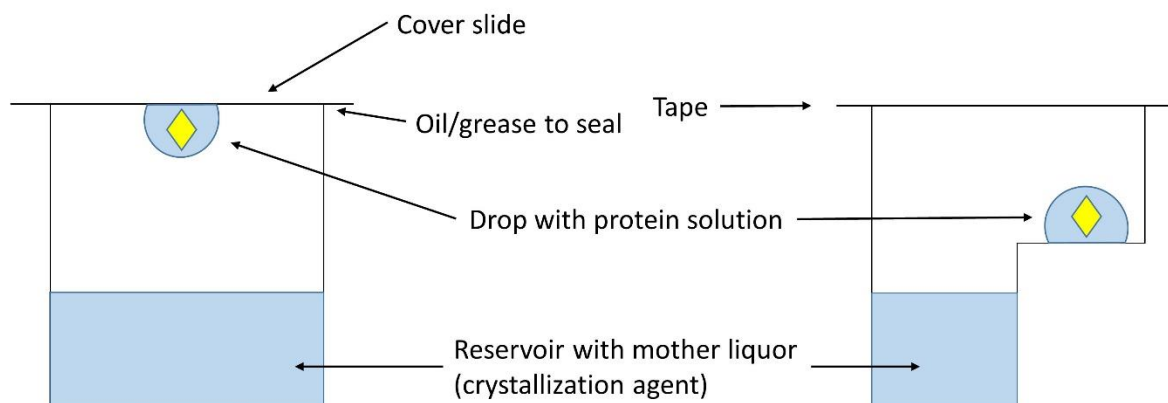
Figure 33: Vapour diffusion crystallization method with a hanging-drop set up (left) and a sitting-drop set up (right).

2.2.3   Improving crystal size by seeding and feeding

The seeding technique has been employed for the growth of diffraction-quality protein crystals for many structures. The main advantage is that it uncouples nucleation from growth. The previously nucleated crystals are used as seeds and can be introduced into new drops pre-equilibrated at a lower level of supersaturation. Generally, there are two types of seeding, the microseeding and the macroseeding.

Microseeding represents the easiest version that includes a transfer of microcrystals that are too small to be distinguished individually. During microseeding, one or more crystals are crushed into crystalline material using homogenizers, sonication, vortexing, glass beads or other utensils. This seed stock is then used to prepare dilution series with a different numbers of nuclei. The nuclei are then transferred into pre-equilibrated drop either by pipetting a small aliquot or by using a seeding wand dipped into the microseed mixture and streaked across the new drop. The drawback of microseding is that it is difficult to control the number of transferred seeds, and thus needs to be optimized experimentally using the dilution series.

Macroseeding, on the other hand, involves a transfer of a parental single crystal (10-100 µm) after multiple washes to a pre-equilibrated drop that is ideally in the

metastable zone of the phase diagram. Macroseeding requires a lot of crystal manipulations that increase the risk of physical damage to the crystal, unintentional showering, etc. It also requires a lot of optimization regarding the time needed for the pre-equilibration, the washing procedure of the parental crystal and the determination of the concentration of both the protein and the crystallization reagent needed to be in the metastable zone.

An effective way of growing large crystals is a feeding method. In feeding, a fresh protein solution is added onto a drop that contains a crystal. This has two effects, lowering the supersaturation by diluting the drop and increasing the amount of accessible protein molecules that can support the growth of the existing nuclei. Furthermore, it avoids handling the crystal and thus reduces the potential crystal damage. The repeated feeding method has been successfully applied to grow large crystals of galectin-3C for the neutron crystallographic studies (Figure 34).
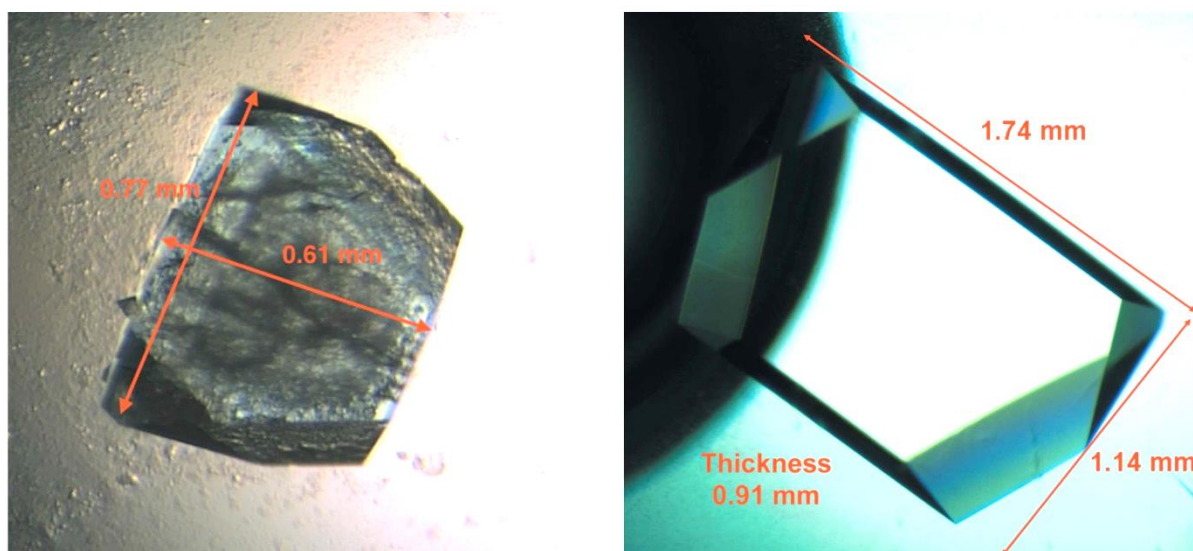
Figure 34: The improvement of crystal quality of galectin-3C grown by macroseeding (left) and repeated feeding (right) techniques. Figure is taken and modified from (Manzoni et al., 2016).

2.3 Neutron protein crystallography

2.3.1   Neutron data collection

The basic principle of single crystal neutron diffraction is similar to X-ray diffraction. X-rays are scattered by electron clouds while neutrons are scattered by atomic nuclei in the crystal lattice. The waves that are scattered coherently result in Brag diffraction spots (reflections) on the diffraction pattern. Each spot corresponds to the unique set of parallel planes of atoms and is assigned a particular set of values (*hkl*) also known as Miller indices. The intensity of the diffraction by a diffraction plane (*hkl*) is determined by the corresponding structure factor *F*(*hkl*) expressed as:

$$F(hkl) = \sum_j b_{c,j} \exp[2\pi(hx_j + ky_j + lz_j)]\exp[-B_j(sin^2\theta/\lambda^2) \qquad (1)$$

with $b_{c,j}$ and B_j representing the coherent neutron scattering length and the temperature factor of an atom j respectively.

The purpose of the diffraction experiment is to find positions of all atoms in the unit cell. These are derived from the computed nuclear scattering density $\rho(xyz)$, which is obtained by the Fourier transform of the structure factors *F*(*hkl*):

$$\rho(xyz) = \frac{1}{V}\sum\sum\sum_{h,k,l}\boldsymbol{F(hkl)}e^{-2\pi i(hx+ky+lz)} \qquad (2)$$

where V is the volume of the crystal unit cell and summation is performed over all values for the Miller indices (*hkl*).

A structure factor completely describes a diffracted ray including its amplitude and phase. The amplitude is proportional to the square root of the reflection intensity I_{hkl} measured from the diffraction experiment. The phase of the diffracted wave is not

directly obtainable which causes the so-called "phase problem" which will be discussed later.

The intensity of a Bragg diffraction spot with indices *hkl* is written as:

$$I \cong I_0 (F(hkl))^2 \frac{V_{cryst}}{V_{cell}^2} \tag{3}$$

where I_0 is the incident neutron intensity, $F(hkl)$ is its structure factor, V_{cryst} is the volume of the crystal and V_{cell} is the volume of the unit cell.

At reactor neutron sources both monochromatic and quasi-Laue neutron data collection modes have been developed. The quasi-Laue method uses a polychromatic beam (pink beam) of neutrons produced by a reactor that is passed through a bandpass wavelength filter in order to select the desired wavelengths. A pink beam of neutrons is used to compensate for the lower flux of neutrons and thus increasing the intensity of the diffraction spots and providing a rapid survey of reciprocal space since many reflections are recorded simultaneously. The Ewald's sphere construction helps to visualize the diffraction conditions (Figure 35).
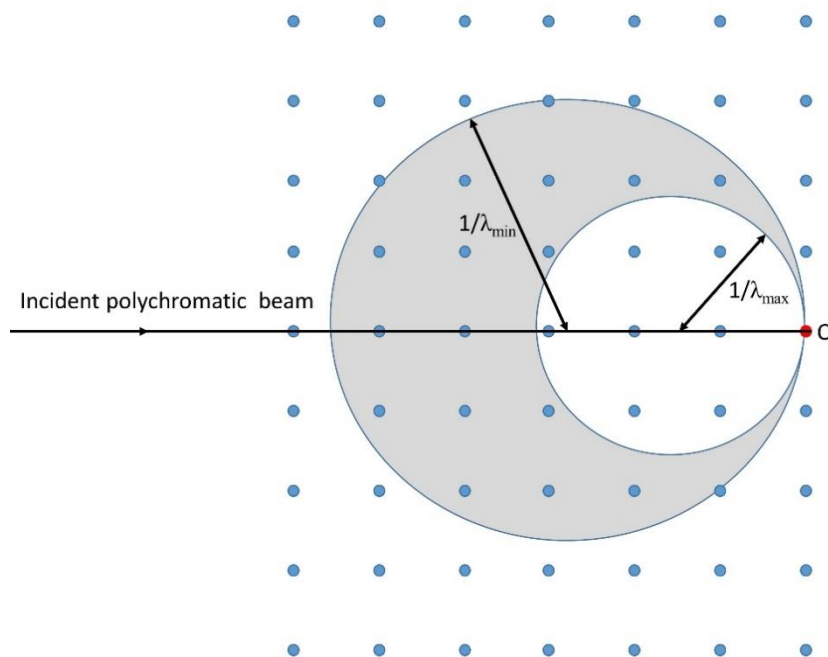
Figure 35: Ewald sphere construction for quasi-Laue diffraction. The crystal is exposed to a polychromatic neutron beam with neutron wavelengths of λ_{min}- λ_{max}. Any reciprocal lattice point (blue circles) laying in between the two circles (grey area) with radii $1/\lambda_{min}$ and $1/\lambda_{max}$ is recorded simultaneously.

The quasi-Laue diffraction method is used on the LADI-III diffractometer at the ILL in Grenoble (France) (Figure 36). The LADI-III instrument is dedicated to neutron protein crystallography at medium to high resolution (2.5-1.5 Å) and is installed on the cold neutron guide H143 at the ILL. LADI-III is equipped with a large cylindrical detector (400 mm in diameter and 450 mm in height) that completely surrounds the sample and provides a larger coverage of the reciprocal space. Bragg reflections are recorded on neutron-sensitive image plates (NIPs) that are bonded on the inner surface of the drum. The crystal in a quartz capillary is mounted on a goniometer head and inserted inside the cylindrical drum where it can be rotated around the cylinder axis. The Ni/Ti multilayer bandpass filter allows to select a wavelength range ($\Delta\lambda/\lambda$= 30 %) centred at 3.5 Å for the majority of systems. Data are collected using the quasi-Laue method, i.e., with a crystal that is held stationary

at a different ϕ (vertical rotational axis). The crystal is rotated by a ϕ step (usually 5-7 °) between successive images.
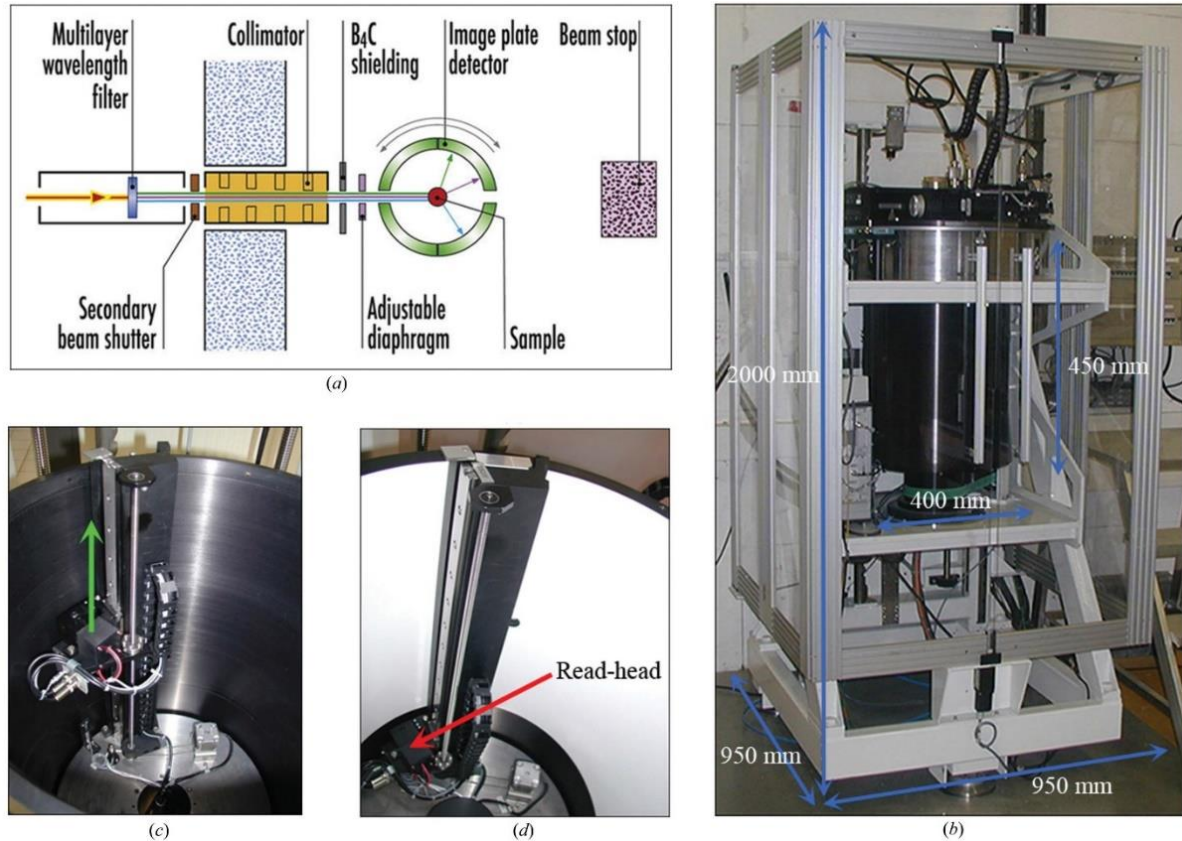
Figure 36: LADI-III instrument at the Institut Laue-Langevin. (a) Schematic layout of the LADI-III diffractometer. (b) Image of the LADI-III diffractometer with the instrument dimensions. (c) The inside of the drum with the read-out head. The green arrow indicates the direction of the reading head while the drum rotates. (d) The inside of the drum with the neutron-sensitive image plate. The red arrow indicates the position of the read-out head during exposure. Figure is taken from (Blakeley et al., 2010).

2.3.2   Neutron data processing

After the data collection, diffraction images are indexed and then integrated to obtain average intensities of all reflections (<I>) with an estimate of their standard uncertainties ($\sigma(I)$). Quasi-Laue diffraction data are indexed and integrated with the *LAUEGEN* suite (Campbell et al., 1998) and the *LSCALE* software (Arzt et al., 1999)

is used for wavelength normalization, which is based upon the intensities of symmetry-related reflections measured at different wavelengths.

As mentioned earlier, the phase information of each diffracted ray is needed for the structure determination. The most common approach is to retrieve the phases from a previously solved structure of a homologous or the same protein. This concept is called *molecular replacement* and was used in this thesis for the determination of all structures.

2.3.3   Structure refinement

Structure model refinement is a process that aims to minimize the differences between the observed structure factors (determined from the original intensity data) and the calculated structure factors (derived from the atomic model). This is done through the refinement of atomic coordinates in space (i.e. x, y and z), atomic displacement parameters (B-factors) and occupancy refinement.

The main goal of neutron protein crystallography is to locate hydrogen atoms that account for about 50% of all atoms in a protein molecule. Introducing these hydrogen atoms in the structure increases the number of parameters to be refined by a factor of two. In order to increase the data-to-parameter ratio, a joint refinement can be done simultaneously with both X-ray and neutron data against the same model. In this thesis, this approach was applied in the refinement of all protein structures for which the neutron data had been collected. *Phenix.refine* (Afonine et al., 2010) in the *Phenix* (Adams et al., 2010) package was used for the structural refinement.

Aim of the study

The aim of this thesis was to use single-crystal neutron diffraction experiments to provide novel insights into lectin-glycan recognition. Several crystal structures of sugar-binding proteins have been solved using neutron diffraction in recent years, including a few glycosylhydrolases illustrating the importance of hydrogen-bonding and water-mediated interactions for protein-carbohydrate recognition. Neutron data on lectins are limited and include available structures from concanavalin A and galectin-3C both in apo and ligand-bound forms. No data are available on the special role played by calcium ions that are present in numerous protein-carbohydrate interactions. Novel structures would therefore be of high interest for the basic understanding of the specificity and affinity involved in biological processes, for designing carbohydrate-based ligands or glycomimetics that are of therapeutical interest and finally for designing engineered neo-lectins for applications in synthetic glycobiology.

Two fucose-specific lectins that differ in sequences and in binding sites have been chosen for this study, namely PLL lectin from bacteria *Photorhabdus laumondii* and LecB lectin from bacteria *Pseudomonas aeruginosa*. PLL interacts with fucose via an interplay of hydrogen bonds, hydrophobic contacts and stacking interactions with tryptophan residues. Sugar binding by LecB involves bridging by two calcium ions, and by numerous water molecules. The presence of two calcium ions in the LecB binding site is a unique feature and is related to a very strong affinity of this lectin for fucosylated ligands.

The aim of this work was to produce these two perdeuterated lectins using high cell-density cultures, crystallize them, collect neutron diffraction data and solve the neutron structures. A necessary step was to produce perdeuterated ligand, L-fucose using a glyco-engineered strain of *Escherichia coli* bacteria to be used as a ligand for perdeuterated lectins during crystallization experiments.

The following chapters are based on the publications, each consisting of introduction, results, discussion and materials and methods. Each publication is preceded by a general introduction in which I explain the motivation behind the work, describe methodological procedures and discuss difficulties that I have encountered.

3

Production of perdeuterated fucose by engineered *E. coli*

3.1 Available methodology for the production of deuterated carbohydrates

Production and availability of perdeuterated glycans is more limited than production of perdeuterated proteins. More than 60 years ago, it was demonstrated that algae can be grown in 99.6% D_2O medium and used as a source of perdeuterated compounds including sugars such as glucose and mannose (Chorney et al., 1960; Gatley et al., 1986; Katz and Crespi, 1966).

Direct catalytic deuteration method using a deuterated Raney nickel catalyst in D_2O allows isotope exchange at carbon atoms that are bound to free hydroxyl groups such as those in carbohydrates. This method has been known for a long time and has been used for perdeuteration of trehalose and methyl pyranosides of D-glucose, D-galactose and D-mannose (Kent et al., 2014; Koch and Stuart, 1978). More recently, a new and efficient direct stereo- and regioselective deuterium labelling method has been applied to production of deuterated pyranosides and furanosides using a carbon-supported platinum group catalyst (Ru/C) that catalyzes the incorporation of deuterium atoms adjacent to hydroxyl groups under hydrogen atmosphere in D_2O (Sawama et al., 2012).

Multi-step biocatalytic synthesis of perdeuterated monosaccharides from low-molecular-weight deuterated precursors such as perdeuteriochlorobenzene and

perdeuteriodiol has been used to produce perdeuterated D-mannose and D-mannitol (Hudlicky et al., 1996).

Chemoenzymatic synthesis has been used for production of perdeuterated polysaccharide heparin from biosynthetically produced perdeuterated heparosan precursor (Cress et al., 2019). Perdeuterated bacterial cellulose has been obtained by growing D$_2$O-adapted strains of *Gluconacetobacter* and *Acetobacter* species to deuterated medium with deuterated glycerol as carbon source (O'Neill et al., 2015; Su et al., 2016). Only recently, recombinant *Escherichia coli* were used to produce perdeuterated anthocyanin, cyanidin 3-*O*-glucoside, by methabolic pathways from perdeuterated glycerol in a D$_2$O-based culture medium (Gupta et al., 2018).

3.2 Application of deuterated carbohydrates

Perdeuterated biomolecules including carbohydrates have wide applications in analytical methods such as the study of metabolism and biodistribution in animals (Shigeki Shimba et al., 1990). Perdeuterated monosaccharides have been used to monitor the nutrients uptake by bacteria using MeV ion beam analysis technique (Lowery et al., 2015). Raman scattering imaging technique has been developed to trace deuterium transfer along glucose metabolic pathways using a deuterated glucose precursor (Zhang et al., 2019). Deuterated biomolecules are gaining interest for the use in medicine and drug development. The isotope exchange has direct effects on pharmacokinetic and pharmacodynamic properties of perdeuterated molecules resulting in longer exposure, increased metabolic stability and reduced enzymatic cleavage (Liu et al., 2017). Recently, the first deuterated drug deutetrabenazine has been approved for use in treatment of involuntary movements in Huntington's disease (Dean and Sung, 2018).

Deuterated carbohydrates have been used to help stabilize membrane proteins such as bacteriorhodopsin during small-angle neutron scattering experiments (Midtgaard et al., 2018). Last but not least, several deuterated monosaccharides have

been used in the neutron protein crystallogaphy studies of carbohydrate-processing enzymes including perdeuterated D-xylulose and D-glucose (Kovalevsky et al., 2010, 2008; Langan et al., 2014). The importance of the use of the perdeuterated glycans in neutron crystallography is obvious as described earlier (see 1.4 Neutrons in stuctural biology). The progress in glycobiotechnology that focuses on glyco-engineering of microorganisms that produce mono- and oligosaccharides has come a long way. It is now routinely used for a high-scale production of glycans using high cell-density cultures. Recombinant *E. coli* are commonly used for the production of perdeuterated proteins and lipids and could become an efficient source of perdeuterated carbohydrates.

In collaboration with the deuteration laboratory of the ILL, we have decided to produce perdeuterated fucose using engineered *E. coli* in a bioreactor. Fucose is a target for several lectins from human pathogens including LecB lectin from *Pseudomonas aeruginosa* and was to be used as a ligand for lectins chosen for neutron diffraction studies.

3.3 Engineered *E. coli* for fucose production – methods developments

At the CERMAV institute (Centre de Recherches sur les Macromolécules Végétales) in Grenoble (France), Dr. Eric Samain and his team have been working on large-scale *in vivo* synthesis of various glycans including human oligosaccharides using metabolically engineered *Escherichia coli* cells (Priem et al., 2002). The FUC5 strain was designed by E. Samain to synthetize fucose using specific metabolic pathways with deletion and insertion of genes involved in the fucose synthesis (Figure 1 in Article I). Bacteria were grown on an alternative carbon source, glycerol. Lactose served as an acceptor for fucosyltransferase and was added to the medium during the fermentation.

The FUC5 strain was transferred from CERMAV to ILL and the first part of the project was to adapt it to deuterated medium for the production of perdeuterated compound. In order to optimize conditions for high cell-density cultivation of the FUC5 strain, we performed several fermentation experiments in two media, Enfors and Minimal media (Table 2). First, we compared the bacterial growth and product yield in hydrogenated fermentation both in batch and fed-batch set ups (Figure 37 and 38). By comparing the thin layer chromatography profiles after both fermentations and discussion with Eric Samain at CERMAV who performed other high cell-density cultivations in the Mineral media with fucose yields of over 20 g L-1, we decided to carry out a batch fermentation in the deuterated Mineral medium.

Table 2: Composition of Enfors and Minimal medium used for high cell-density culture of *E. coli* FUC5 strain.

	Enfors medium		Minimal medium	
	Components	g L^{-1}	Components	g L^{-1}
Minimal medium	$(NH_4)_2SO_4$	6.86	$NH_4H_2PO_4$	5.00
	KH_2PO_4	1.56	KH_2PO_4	5.00
	$Na_2HPO_4 \cdot 2H_2O$	6.48	KOH	1.65
	$(NH_4)_2HC_6H_5O_7$	0.49	NaOH	0.65
	$MgSO_4 \cdot 7H_2O$	0.25	$MgSO_4 \cdot 7H_2O$	1.00
	Glycerol	5.00	Citric acid	0.50
			Glycerol	45
Trace metals	$CaCl_2 \cdot 2H_2O$	0.50	$N(CH_2CO_2H)_3$ (nitrilotriacetic acid)	13.00
	$FeCl_3 \cdot 6H_2O$	16.70	KOH	7.00
	$ZnSO_4 \cdot 7H_2O$	0.18	$C_6H_5FeO_7$ (ferric citrate)	7.50
	$CuSO_4 \cdot 5H_2O$	0.16	$MnCl_2.4H_2O$	1.30
	$MnSO_4 \cdot 4H_2O$	0.15	$ZnSO_4.7H_2O$	1.20
	$CoCl_2 \cdot 6H_2O$	0.18	H_3BO_3	0.25
	Na-EDTA	20.10	$Na_2MoO_4.2H_2O$	0.15
			$CoCl_2. 6H_2O$	0.21
			$CuCl_2.2H_2O$	0.13
H_2O		Up to 1 L		Up to 1 L
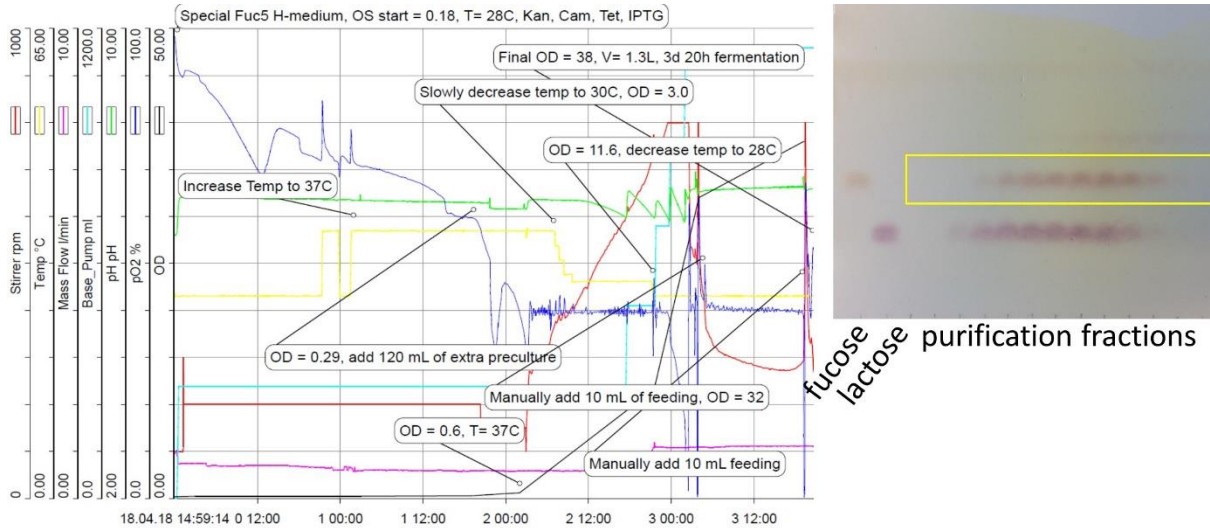
Figure 37: High cell-density culture profile of *E. coli* FUC5 strain for production of fucose in a hydrogenated Minimal medium using a batch fermentation. Hydrogenated Mineral medium was used for this fermentation. Lactose and isopropyl β-D-1-thiogalactopyranoside (IPTG) were added to the medium at the beginning of the fermentation to a final concentration of 0.5 g L^{-1} and 0.2 mM, respectively. The stirring speed (red line) was adjusted automatically to keep the pO_2 (blue line) level at 30 %. The pD (green line) of the culture medium was regulated at 6.8 by addition of base (light blue line, 30% NH_4OH). The temperature (yellow line) was monitored and was set to 28 °C towards the end of the fermentation. Bacterial growth (solid black line) was monitored off-online by regular sampling the cell culture and measuring OD_{600} (optical density at 600 nm). On the right of the Figure is a TLC chromatogram after the purification of the extracellular medium carried out after the fermentation.
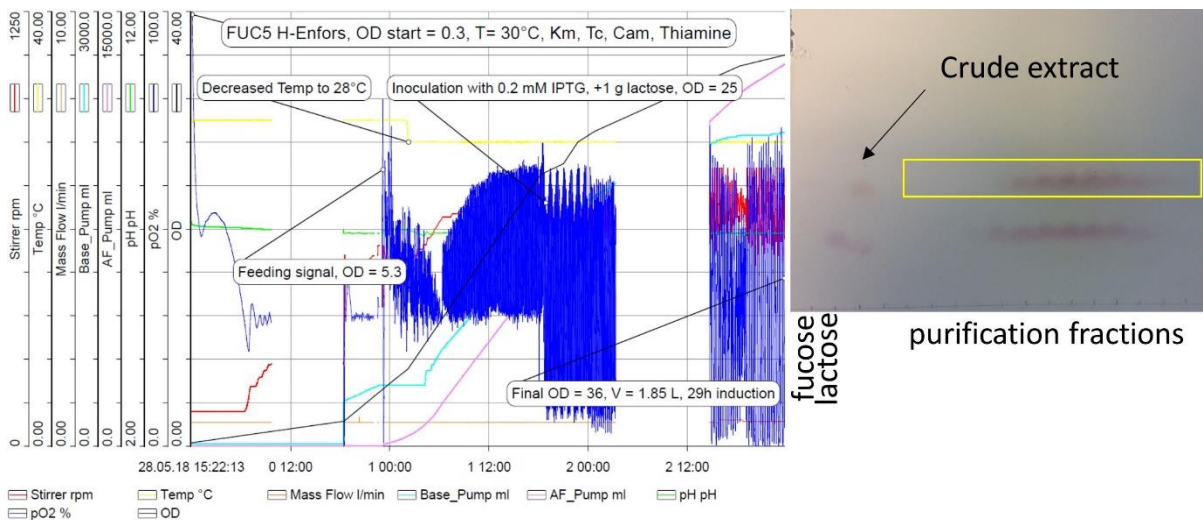
Figure 38: High cell-density culture profile of *E. coli* FUC5 strain for production of fucose in hydrogenated minimal medium using a batch/fed-batch fermentation. Hydrogenated Enfors minimal medium was used for this fermentation. After

79

consumption of the initial glycerol supply (5 g L^{-1}), cells were continuously fed with a feeding solution containing 12% glycerol (fed-batch phase). On the right of the Figure is a TLC chromatogram after the purification of the extracellular medium after fermentation.

Then, cells had to be adapted to the deuterated medium through several rounds of passaging to the fresh 100%-D$_2$O-based minimal medium with deuterated glycerol-d$_8$ as a carbon source. The final high cell-density culture was carried out in a 3-L bioreactor using a simple batch fermentation (Figure 39). During the fermentation process, cells exhibited a very slow growth in the lag phase. In order to boost the growth, we varied the temperature of the culture medium. During the exponential phase of the growth, we noticed sudden increases in pO$_2$ concentration indicating low metabolic rate of growing bacteria. A possible explanation could be the lack of nitrogen in the growth medium since the pD of the culture medium was adjusted by the addition of NaOD instead of ammonia that was used during the fermentation in hydrogenated conditions. The produced deuterated fucose was secreted into the extracellular medium throughout the high cell-density cultivation. At the end of the fermentation, cells were spun and the medium was filtered prior to purification.

The production of fucose was first followed by thin layer chromatography, which indicated the amount of produced fucose in a semi-quantitative way. The principle of the detection is based on the colorimetric reaction between carbohydrates and orcinol (poly-hydric phenol) heated in the presence of sulphuric acid (Bruckner, 1955). In order to precisely estimate the amount of produced fucose we used a commercially available kit for detection of free L-fucose in biological samples, called K-Fucose (Megazyme). This simple and rapid way of detection is based on the enzymatic conversion of L-fucose to L-fucono-1,5-lactone by fucose dehydrogenase in the presence of NADP. The amount of fucose is stoichiometric to the amount of NADPH produced during the reaction and that is measured by the increase of absorbance at 340 nm.
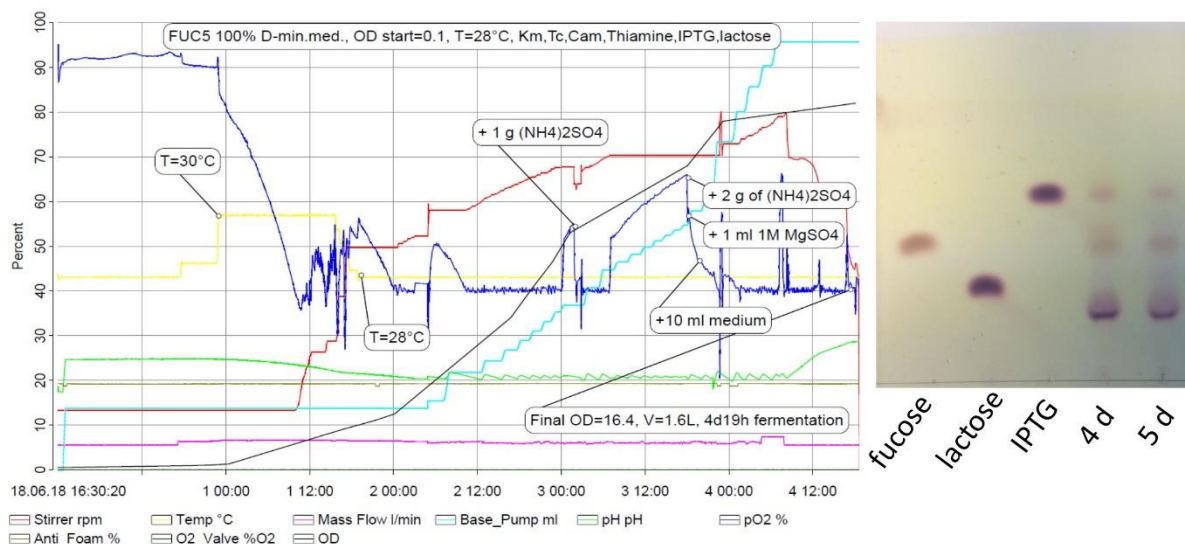
Figure 39: High cell-density culture profile of *E. coli* FUC5 strain for production of perdeuterated fucose in a deuterated minimal medium using a batch fermentation. Deuterated Mineral medium was used for this fermentation with deuterated glycerol-d_8 as carbon source. On the right of the Figure is a TLC chromatogram after the fermentation.

The purification step was crucial for obtaining a pure product and was also the most demanding and time-consuming part. After the initial ion-exchange chromatography purification, we observed insufficient purity based on the thin-layer chromatogram that revealed a considerable amount of lactose and IPTG that were co-eluted with fucose. In order to achieve a more efficient separation of these sugar molecules, we used the CERMAV Common Chromatography Service platform (with help of Laurine Buon) equipped with a size-exclusion chromatography systems for separation of mono- and oligosaccharides. A TOYOPEARL® HW-40 column filled with beads made of a hydroxylated methacrylic polymer was used in order to fractionate the mixture by size-exclusion chromatography. Compounds could be nicely separated and eluted as single peaks as shown by the refractive index detector (Figure 3 bottom in Article I).

After obtaining a pure lyophilized deuterated fucose, we proceeded with biophysical analysis. In order to characterize the perdeuterated product, we carried out mass spectrometry and nuclear magnetic resonance spectroscopy measurements.

81

Mass spectrometry experiments using electrospray ionization were first performed on the lyophilized sample dissolved in D_2O. The MS spectra were very complex (Figure 40) so in order to confirm the presence of perdeuterated fucose we carried out further MS experiments with sample dissolved in H_2O/methanol solution. Here, four exchangeable protons on fucose hydroxyl groups were back-exchanged from deuteriums to hydrogens. The presence of sodium, potassium and chloride adduct peaks agreed with the theoretical molecular weight of fucose molecule after the H/D exchange (Figure 41).
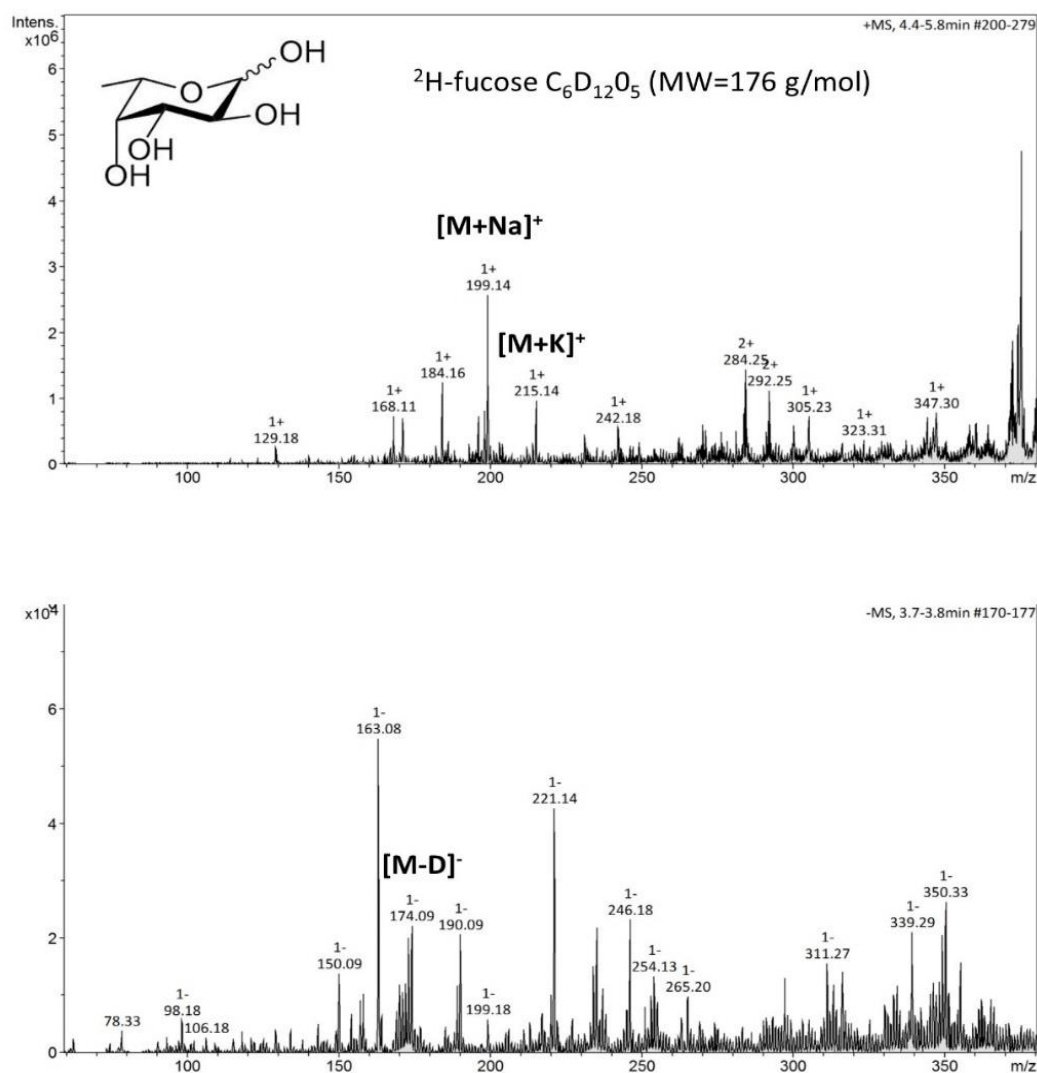
Figure 40: Deconvoluted ESI mass spectra of perdeuterated fucose dissolved in D_2O.
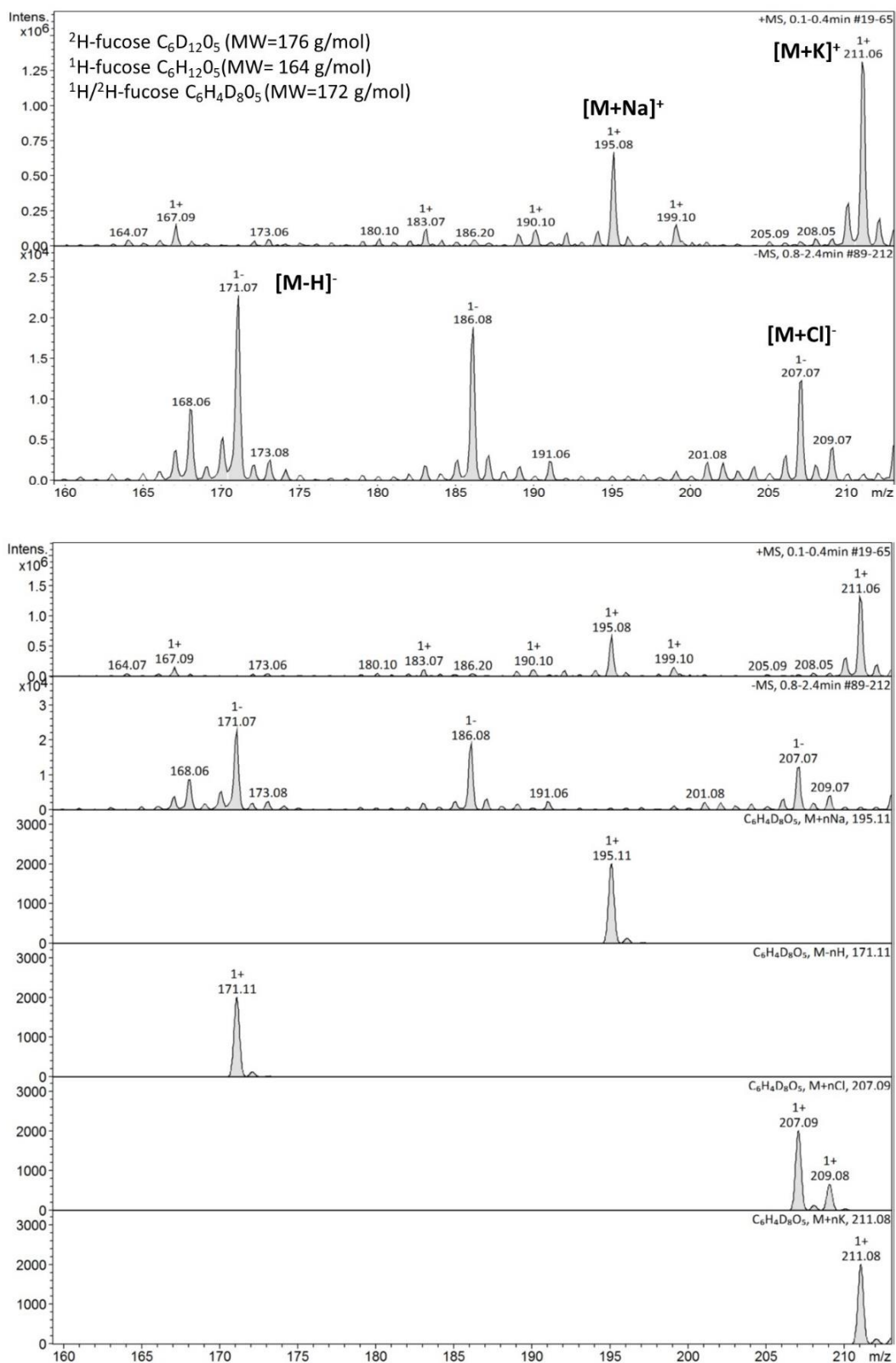
Figure 41: Deconvoluted ESI mass spectra of perdeuterated fucose dissolved in H_2O/methanol solution.

Proton and deuterium NMR showed full deuteration with decrease in signal intensity in the ^{1}H spectrum and increase of signal intensities in the ^{2}H spectrum (Figure 4 in Article I). Both techniques thus confirmed that the final product was fully deuterated L-fucose-d$_{12}$. The final amount of the product was estimated using the commercial kit (K-Enzyme) to be 0.2 g and confirmed by weighing the lyophilized powder.

Full results are described in the article below.

3.4 Article I

Résumé

Le L-fucose et les polysaccharides, glycoprotéines et glycolipides contenant du L-fucose
jouent un rôle important dans une variété de processus biologiques. Les glycoconjugués
contenant du L-fucose sont impliqués dans de nombreuses maladies, notamment le
cancer et la polyarthrite rhumatoïde. L'intérêt pour le fucose et ses dérivés est croissant
dans la recherche sur le cancer, la glyco-immunologie et l'étude des interactions entre
l'hôte et les agents pathogènes. Le L-Fucose peut être extrait de polysaccharides
bactériens et d'algues, ou produit (bio)synthétiquement. Alors que le glucose et le
galactose deutérés sont disponibles et présentent un grand intérêt pour les études
métaboliques et biophysiques, le fucose deutéré n'est pas facilement accessible.

Nous décrivons ici la production de L-fucose perdeutéré dans un bioréacteur, en
utilisant une souche ingéniériée de *Escherichia coli* avec l'utilisation d'un milieu de
croissance à base de D_2O et une source de carbone deutérée. Le rendement final était
de 0,2 g L^{-1} de sucre deutéré, ensuite caractérisé par la spectrométrie de masse et la
spectroscopie de résonance magnétique nucléaire. Nous prévoyons que le fucose
perdeutéré ainsi produit aura de nombreuses applications en biologie structurale où
des techniques telles que la RMN, la diffusion des neutrons en solution et la
cristallographie neutronique sont largement utilisées. Dans le cas de la cristallographie
macromoléculaire aux neutrons, la disponibilité du fucose perdeutéré peut être
exploitée pour identifier les détails de son interaction avec les récepteurs de protéines
et notamment le réseau de liaison hydrogène dans les sites de liaison des glucides.
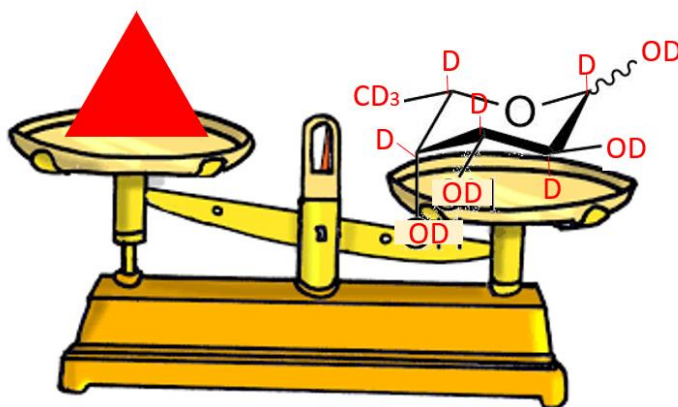
Production of perdeuterated fucose from glyco-engineered bacteria

**Lukas Gajdos[a,b,c], V. Trevor Forsyth[a,b,d], Matthew P. Blakeley[e], Michael Haertlein[a,b],
Anne Imberty[c]*, Eric Samain[c]* & Juliette M. Devos[a,b]***

[a] Life Sciences Group, Institut Laue-Langevin, 71 Avenue des Martyrs, 38000 Grenoble,
France
[b] Partnership for Structural Biology (PSB), 71 Avenue des Martyrs, 38000 Grenoble, France
[c] Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France
[d] Faculty of Natural Sciences, Keele University, ST5 5BG Staffordshire, UK
[e] Large Scale Structures Group, Institut Laue-Langevin, 71 Avenue des Martyrs, 38000
Grenoble, France.

* Authors for correspondence: Anne Imberty (anne.imberty@cermav.cnrs.fr), Eric Samain
(eric.samain@cermav.cnrs.fr), Juliette Devos (devosj@ill.fr).

Supplementary data: Figure S1, Figure S2

Abstract

L-Fucose and L-fucose-containing polysaccharides, glycoproteins or glycolipids play an
important role in a variety of biological processes. L-Fucose-containing glycoconjugates have
been implicated in many diseases including cancer and rheumatoid arthritis. Interest in fucose
and its derivatives is growing in cancer research, glyco-immunology, and the study of host-
pathogen interactions. L-Fucose can be extracted from bacterial and algal polysaccharides, or
produced (bio)synthetically. While deuterated glucose and galactose are available, and are of
high interest for metabolic studies and biophysical studies, deuterated fucose is not easily
available.

Here, we describe the production of perdeuterated L-fucose, using glyco-engineered
Escherichia coli in a bioreactor with the use of a deuterium oxide-based growth medium and a
deuterated carbon source. The final yield was 0.2 g L^{-1} of deuterated sugar, which was fully
characterized by mass spectrometry and nuclear magnetic resonance spectroscopy. We
anticipate that the perdeuterated fucose produced in this way will have numerous applications
in structural biology where techniques such as NMR, solution neutron scattering and neutron
crystallography are widely used. In the case of neutron macromolecular crystallography, the
availability of perdeuterated fucose can be exploited in identifying the details of its interaction
with protein receptors and notably the hydrogen bonding network around the carbohydrate
binding site.

Introduction

The deuteration of biomolecules by stable
isotope labelling is of high interest for

nuclear magnetic resonance (NMR) spectroscopy and neutron scattering studies of the structure and dynamics of biological macromolecules and is now widely used (Haertlein et al. 2016, Blakeley and Podjarny 2018). Tailor-designed deuteration can be used very effectively in studies of multi- component systems by small-angle neutron solution scattering (Laux et al. 2008, Cuypers et al. 2013b, Dunne et al. 2017, Josts et al. 2018, Maric et al. 2019) and neutron reflection (Grage et al. 2011, Hellstrand et al. 2013, Waldie et al. 2018). In Laue neutron crystallography, the use of perdeuterated protein imparts major benefits in terms of data quality and interpretation (Haupt et al. 2014, Dajnowicz et al. 2017, Yee et al. 2019, Kwon et al. 2020) and the same is true in the case of monochromatic neutron crystallography (Cuypers et al. 2013a, Cuypers et al. 2016). Deuteration of small biomolecules has wide application in analytical methods, with growing development for the study of metabolism (Shimba et al. 1990) and in living cells imaging techniques such as Raman microscopy (Wei et al. 2013). MeV ion beam analysis techniques such as ^{3}He Nuclear Reaction Analysis (NRA), where high sensitivity to deuterium was utilized to monitor the uptake of deuterated sugars in bacteria (Lowery et al. 2015). More recently, it has been demonstrated that deuteration of key hydrogen atoms at positions where drugs are enzymatically metabolised for elimination has direct effects on pharmacological properties of these substances by improving their stability, and therefore their half-life, and by decreasing their toxicities (Schmidt 2017). The basis of the increased stability of these "heavy hydrogen drugs" is related to the stronger deuterium-carbon bond making the drug more resistant to enzymatic modification and elimination; deuteration of known drugs, referred to as the "deuterium switch", resulted in a boost of patenting of labelled drugs, with several

of them now in clinical trials (Timmins 2017). The first approved deuterated drug for therapeutic use was deutetrabenazine, and it has been used for the treatment of chorea in Huntington's disease (Dean and Sung 2018).

Proteins and lipids are generally obtained in their perdeuterated form through the use of recombinant bacteria or yeasts (Meilleur et al. 2009, Haertlein et al. 2016, Moulin et al. 2018). The availability of perdeuterated, and well characterized glycans is more limited. For selective deuteration, chemical synthesis of perdeuterated monosaccharides has previously been described. Through the use of appropriate protecting groups, direct H–D exchange reactions in the presence of an activated carbon-supported platinum group catalyst were used to produce a variety of deuterated monosaccharides (Koch and Stuart 1978, Sawama et al. 2012). Using biotechnology approaches, it has been demonstrated more than 50 years ago that certain algae can be adapted to grow in 99.6% D_2O, resulting in the large scale production of perdeuterated glucose and mannose (Crespi et al. 1959, Chorney et al. 1960). It is only recently that cyanidin 3-*O*-glucoside was produced with a perdeuterated glucose moiety using recombinant *Escherichia coli* (Gupta et al. 2018). Polysaccharides have been produced in bacteria, in the form of perdeuterated cellulose (O'Neill et al. 2015) and perdeuterated heparin (Cress et al. 2019).

The current work focuses on L-fucose (6-deoxy-L-galactose), which is a non-classical monosaccharide with an L-configuration correlated to unusual ring conformation, and a methyl group at C6. L-fucose (Fuc) is present in all living kingdoms, from giant viruses (De Castro et al. 2013) to bacterial polysaccharides and plant cell walls. In mammals, it is a common component of glycolipids and glycoproteins (Vanhooren and Vandamme 1999, Schneider et al. 2017). Terminal fucosylation of oligosaccharides create

bioactive epitopes, such as the histo- blood group oligosaccharides (ABO and Lewis epitopes) (Painter et al. 1965). Human milk oligosaccharides are also rich in terminal fucose, which serve as receptor analogues for pathogens, and modulator of immune responses (Bode 2015). Fucose is therefore of interest as a biomarker for some cancers as well as a target for pathogen receptors (Heggelund et al. 2017).

The production of perdeuterated fucose, L-fucose-d_{12} (Fuc-d_{12}) is therefore of high interest. The chemical approach has been described but is complex and time-consuming (Koch and Stuart 1978, Sawama et al. 2012). An alternative approach is based on the production of "recombinant" oligosaccharides in microorganisms, by engineering of their metabolic pathways. This area of research was initiated at the end of the last century with the production of chito-oligosaccharides in *Escherichia coli* (Samain et al. 1997). During the last decades, a large number of oligosaccharides have been produced using this approach, including human milk oligosaccharides, that are of high commercial interest for improving infant formula milk (Priem et al. 2002). For example, α-2'-fucosyllactose, a major trisaccharide of human milk, has been obtained by introducing the FUT2 gene from *Helicobacter pylori* into *E. coli* (Albermann et al. 2001, Drouillard et al. 2006, Yu et al. 2018). This trisaccharide can serve as a substrate for a fucosidase. With appropriate engineering of microorganisms, fucose can be recombinantly produced, as demonstrated in *Saccharomyces cerevisiae* (Liu et al. 2018) and then in *E. coli* (Liu et al. 2019).
 In this work, the production of perdeuterated Fuc-d_{12} (per-*C*-deuterated Fuc-d_8 when dissolved in H_2O) is described. Based on the previous construct for production of α-2'- fucosyllactose (Drouillard et al. 2006), the metabolic pathways of the previously described *E.*

coli strain were further engineered with the addition of the appropriate hydrolase and removal of sugar transporter. After adaptation of the strains for cultivation in an appropriate deuterated medium, a large amount of perdeuterated fucose was produced and fully characterized.

Results
The fucose-producing strain
A schematic diagram of the production of L- fucose using the engineered FUC5 strain is shown in Figure 1. FUC5 is a derivative of the *E. coli* K12 strain DH1, with genetic modifications for overexpression of genes involved in GDP-fucose production, introduction of 2'fucosyltransferase gene from *Helicobacter pylori* and fucosidase gene from *Bifidobacterium bifidum* strain JCM1254, and deletion of genes involved in the metabolism of lactose. In addition, several genes have been knocked out to optimize the production of fucose: the *fucI* gene (coding for fucose isomerase) to prevent isomerisation to fuculose and the *fucP* gene (fucose permease) to prevent internalization of extracellular fucose. A detailed list of the plasmids, genes and *E. coli* strains described in this work can be found in Table I.

88

Figure 1. Schematic representation of the metabolic network for fucose production using the engineered *E. coli* FUC5 strain. Inserted genes are shown in italics and knockout genes are crossed out.

Table I. Genes, plasmids and *E. coli* strains used in this study

	Description	Reference or source
Genes		
futC	α-1,2-fucosyltransferase from *Helicobacter pylori* strain 26695	GenBank KY499613
afcA	α-1,2-fucosidase from *Bifidobacterium bifidum* strain JCM1254	GenBank AY303700
gmd, wcaG, manC, manB	*E. coli* genes coding GDP-Man dehydratase, fucose synthase, GDP-Man pyrophosphorylase and phosphomannomutase respectively	GenBank U38473
Plasmids		
pEXT20-futC	pEXT20 carrying *futC*	(Drouillard et al. 2006)
pBP-futC	Derived from pEX20-futC	This study
pBBRGAB	pBBR1MCS-3 derivative carrying *gmd*, *wcaG*, *manC* and *manB*	(Dumon et al. 2006)
pSU-fucase	pSU2718 derivative carrying active α-1,2-fucosidase domain of *afcA* from *Bifidobacterium bifidum*	This study
Strains		
DC	DH1 *lacZ lacA*	(Dumon et al. 2006)
FUC	DC *fucI fucP*	This study
FUC5	FUC (pBP-futC, pBBRGAB, pSU-fucase)	This study

Adaptation to D_2O and batch fermentation
In order to compare fucose production in different conditions, the FUC5 strain was grown in both unlabelled and deuterated minimal medium supplemented with hydrogenated glycerol and deuterated glycerol-d8, respectively. Its growth was stepwise adapted to deuterated media as previously described (Haertlein et al. 2016). Both hydrogenated and deuterated fucose

were produced in simple batch-fermentations. The hydrogenated culture showed a higher growth rate in the exponential phase, reaching a maximum OD_{600} of 26, 73 hours after inoculation. The deuterated culture showed a similar lag-phase profile but slower growth in the exponential phase with the highest OD_{600} being 17, 112 hours after the start of the fermentation (Figure 2A).
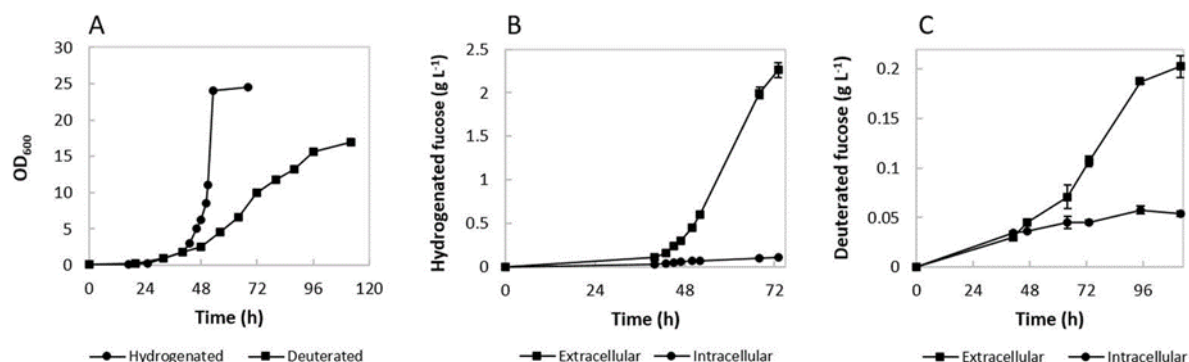
Figure 2: (A) Growth curves of hydrogenated and deuterated *E. coli* FUC5 in high cell-density cultures. The optical density was measured at 600 nm using a spectrophotometer and was plotted against time (hours). **(B, C)** Production of hydrogenated (B) and deuterated (C) fucose using the genetically modified *E. coli* FUC5. The concentrations of extracellular and intracellular fucose were measured spectrophotometrically using the K-Fucose kit (Megazyme). Symbols represent a mean value of two independent measurements. Error bars indicate the standard deviation and some are smaller than the symbol size used in the panels.

The bacterial growth and the deuterated fucose concentration showed a cell-density-associated production pattern (Figure 2). A maximum Fuc-d_{12} concentration of 0.20 g L^{-1} was obtained by the end of the fermentation process. The yield of fucose produced per gram of carbon source in the growth medium was 10 fold lower for the deuterated culture (0.2 vs. 2.1 g L^{-1}, see Figure 2B and Figure 2C) when supplied with the same quantity of carbon source, *ie* 30 g of glycerol used per liter of growth medium as indicated in the experimental section. Analysis by thin layer chromatography (TLC) showed gradually increasing concentrations of intracellular lactose transported from the medium to the

cells by lactose permease. The TLC also revealed the increase in concentration of extracellular fucose over time (Figures S1 and S2).

Optimisation of purification method
Deuterated fucose was purified directly from the crude extracellular fraction. After the first ion exchange chromatography purification step, TLC analysis showed the presence of lactose and isopropyl-β-D-thiogalactopyranoside (IPTG) that were added to the minimal medium during high cell-density culture. Further fractionation of these compounds was achieved by size-exclusion chromatography (SEC) using a

TOYOPEARL HW-40 column made of

hydroxylated methacrylic polymer (Figure 3

bottom). This technique was used as a polishing step to efficiently separate lactose, fucose and IPTG, three molecules of similar sizes that could be eluted as separate single peaks. After purification, 220 mg of lyophilized Fuc-d_{12} were obtained from a culture using 1.5 L of D_2O and 45 g of deuterated glycerol.
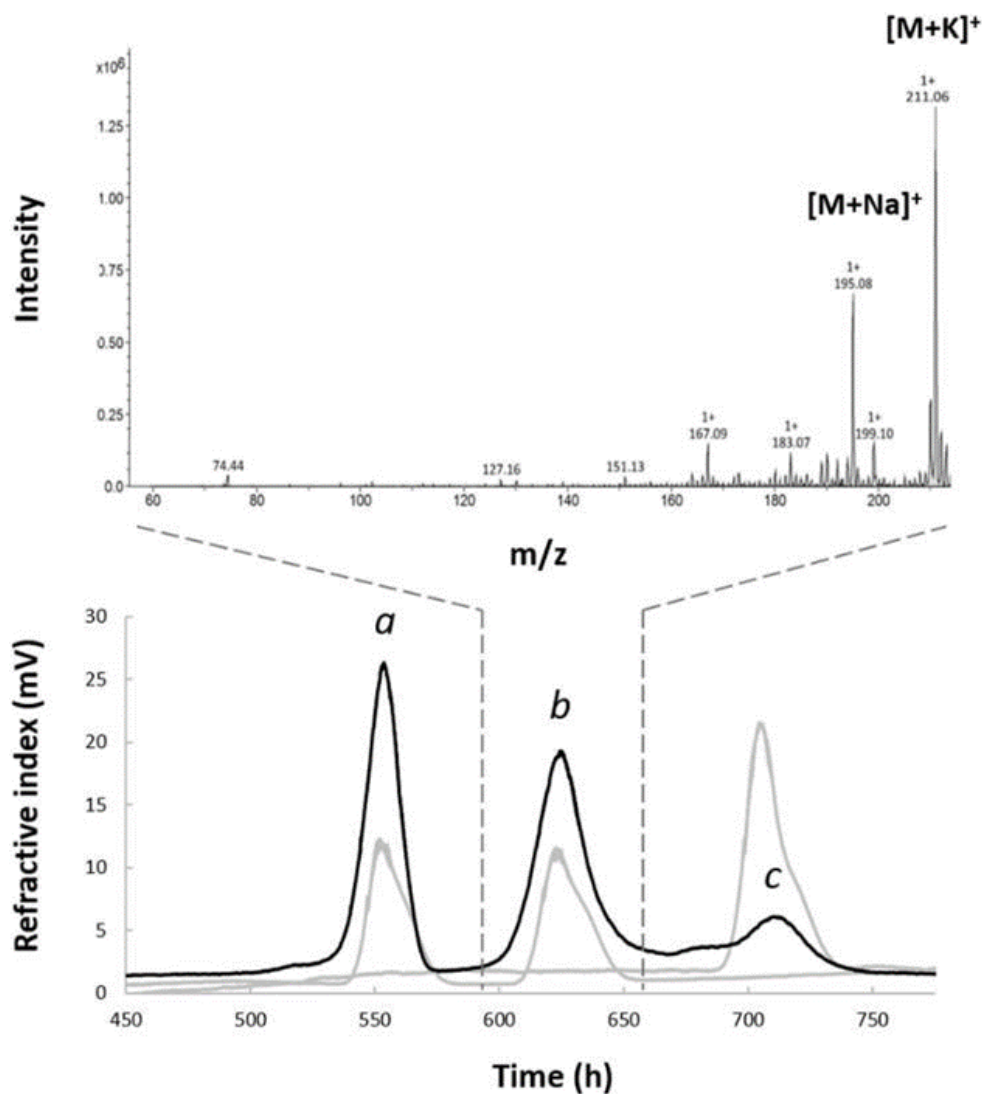
Figure 3: (Top) Mass spectrometry and **(Bottom)** chromatography analyses of per-C-deuterated fucose produced. Analyses were carried out in H_2O and thus correspond to Fuc-d_8. Bottom: Size-exclusion chromatography (SEC) profiles of the saccharides present in the extracellular fraction of the fucose-producing strain after batch-fermentation in deuterated media: (a) lactose, (b) deuterated fucose, and (c) isopropyl-β-D-thiogalactopyranoside (IPTG) in the black line, superimposed with a SEC profile of standard solutions in the grey line. Top: Only the positive electrospray ionization (ESI) mass spectrum of the deuterated fucose (Fuc-d_8) is displayed with sodium and potassium adduct ions at *m/z* 196 and 211 respectively.

Deuterated fucose characterization

The metabolic incorporation of deuterium in the fucose molecule was examined and demonstrated by electrospray ionization (ESI) mass spectrometry, and quantified by ^{1}H NMR. The results (Figure 3 top) showed that the purified product dissolved in a water/methanol solution generated two major peaks corresponding to the exact masses of sodium adduct ions $[M+Na]^+$ and potassium adduct ions $[M+K]^+$ at *m/z* values of 195 and 211, respectively, confirming the presence of per-*C*-deuterated L-fucose (Fuc-d_8).
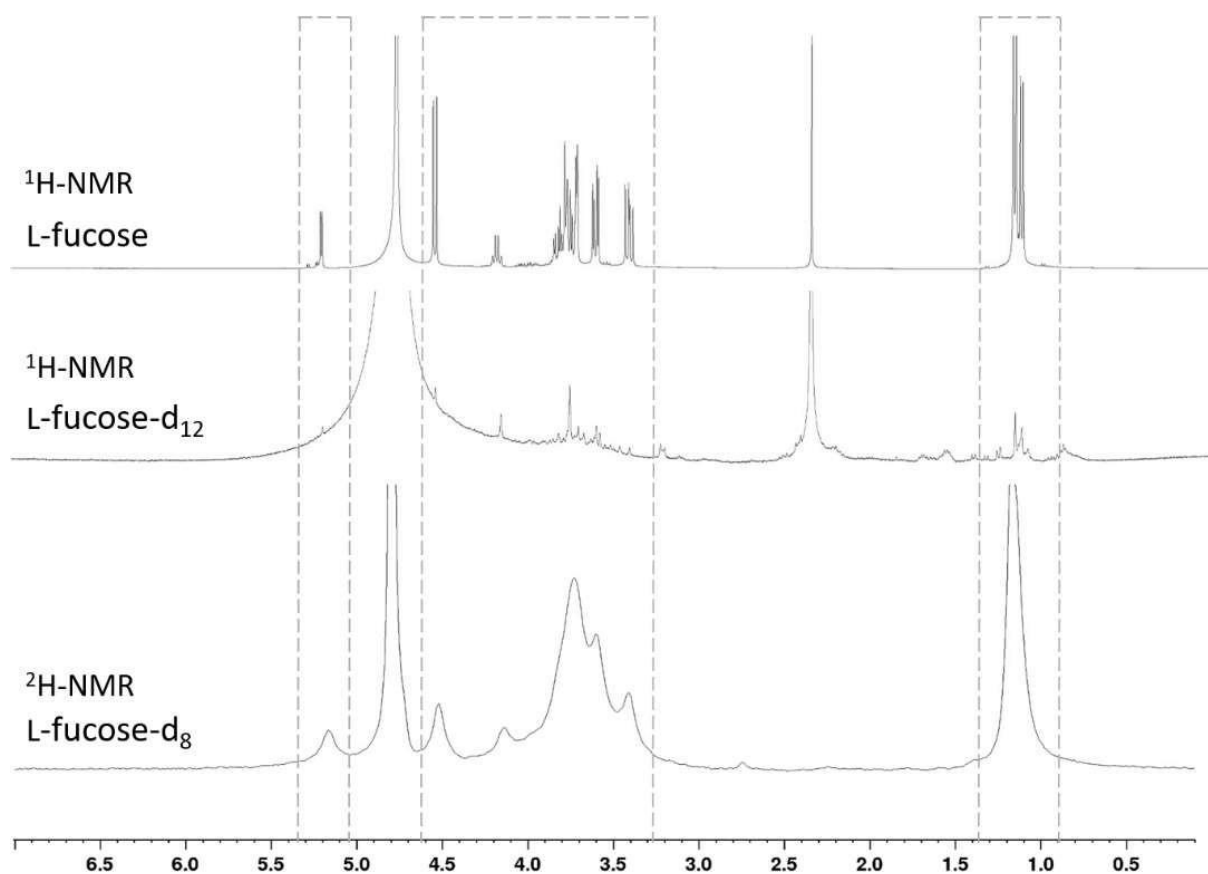
Figure 4: [1]H **(Top and Middle)** and ^{2}H **(Bottom)** NMR spectra comparing the signal intensity for hydrogenated L-fucose and deuterated L-fucose. Peaks in the dashed boxes correspond to the hydrogen (^{1}H spectra) and deuterium (^{2}H spectrum) atoms of the fucose molecule. A signal of sodium succinate is observed at 2.4 ppm for the ^{1}H spectra.

The extent of deuteration of the fucose ring and the methyl carbon was also verified by ^{1}H and ^{2}H NMR. Proton NMR spectra were measured on samples of commercially available hydrogenated L-fucose and on the deuterated L-fucose produced in this work, both solubilized in D_2O. The deuterium NMR spectrum was measured on the deuterated fucose solubilized in water. The ^{1}H spectrum of the purified compound exhibited the loss of intensity of all proton signals compared to the non-labelled standard (Figure 4). Moreover, the deuterium signals observed in the ^{2}H spectrum of the perdeuterated product demonstrated that deuterium replacement has taken place succesfully on all of the carbon atoms of the fucose molecule.

Discussion
In this study, glyco-engineered *E. coli* has been used to produce deuterated fucose from deuterated glycerol and D_2O via biosynthetic pathways. For the production of deuterated fucose-d$_{12}$, deuterated glycerol-d$_8$ was used as the sole carbon source. Even though glucose is a preferred carbon source for *E. coli*, several studies showed that engineered *E. coli* strains producing fucosyllactose and fucose give higher titers of fucose when grown on glycerol (Jung et al. 2019, Liu et al. 2019). Compared to the hydrogenated culture, the perdeuterated high cell-density culture grew slower in the exponential phase and gave lower titers of fucose. Similar behaviour has been observed in deuteration of organic molecules using enginereed organisms (Meilleur et al. 2009, Moulin et al. 2018).
The engineered FUC5 strain produced 0.2 g L^{-1} of deuterated fucose from a limited amount of glycerol in a simple batch-fermentation. The fully deuterated fucose was characterized by mass spectrometry, proton and deuterium NMR, confirming full deuteration of the final product. It should be noted that the final yield of both the hydrogenated and the perdeuterated monosaccharides in this study was strongly dependent on the amount of glycerol used during the fermentation process. Due to the cost of the deuterated carbon source, this amount was limited to 45 g in both the H- and the D-study. Under optimal hydrogenated conditions where the quantity of carbon source is not limited, the highest titer of fucose obtained was 20 g L^{-1} (data not shown).
Perdeuteration using engineered *E. coli* is therefore a convenient and effective method for a high-quantity production of fully deuterated simple and more complex sugars for use in neutron scattering experiments as well as in pharmacokinetic studies. In a first application, Fuc-d$_{12}$ will be used as a ligand for fucose-specific lectins in single crystal neutron diffraction experiments.
Due to the larger and positive coherent scattering of deuterium compared to hydrogen in neutron scattering experiments, there is a better visibility of deuterium atoms versus hydrogen. In this study, the ^{1}H atoms changed for deuterium (^{2}H) atoms in the fucose molecule will allow clearer visualisation of crucial interactions between the sugar and bacterial lectins, as well as an optimisation of the overall quality of the neutron data through the reduction of hydrogen incoherent scattering.

Materials and methods
Strains and plasmid construction
The fucose-producing strain FUC5 was obtained by transforming the host strain FUC with the three following plasmids: pBP-*futC* which contains the α-1,2-fucosyltransferase gene *futC* from *H. pylori* strain 26695; pBBRGAB which contains the four *E. coli gmd*, *wcaG*, *manC*, and *manB* genes coding GDP-Man dehydratase, fucose synthase, GDP-Man pyrophosphorylase and phosphomannomutase respectively; pSU-fucase which contains the sequence of the

active α-1,2-fucosidase domain of the *afcA* gene from *Bifidobacterium bifidum* strain JCM1254 (Katayama et al. 2004).

The host strain FUC was designed from strain DC (Dumon et al. 2006) by knocking out the genes *fucI* and *fucP* encoding fucose isomerase and fucose permease respectively. Strain DC was a *lacA lacZ* null mutant derived from the *Escherichia coli* K12 strain DH1 (DSM 4235). In order to knockout the *fucPI* genes, a 1.435 kb segment located between nucleotides 913 of *fucP* and 998 of *fucI* was deleted and replaced by the 5'AAGCTT sequence as follows: two DNA segments flanking the deleted sequence were amplified by PCR. The upstream 0.873 kb segment was amplified with primers 5'GGATCCGTAGATAAAGATGCAGGGCAAAGCAGAAG and 5'AAGCTTGGTTCCGGTTAAATAGTTAGCGGCAAAG and the downstream 0.806 kb segment was amplified with primers 5'AAGCTTCGTGGCGACCGAAAACGACAG and 5'CTCGAGACCGGGCATCACATCAGGGAG. The two amplified fragments were ligated at their terminal *HindIII* restriction site and cloned together into the *BamHI SalI* sites of the suicide vector pKO3. The deletion was then carried out according to the pKO3 gene replacement protocol (Link et al. 1997).

The plasmid pBP-*futC* was designed from plasmid pEXT20-futC (Drouillard et al. 2006) by removing the ampicillin gene using *DraI* digestion and by inserting via blunt-end ligation a kanamycin cassette, obtained by *PstI* digestion of the pUC4K vector.

The design of pSU-fucase plasmid was carried out as follows: a 2.728 kb DNA fragment containing the sequence of the active α-1,2-fucosidase domain of the *afcA* gene was amplified by PCR using genomic DNA of *Bifidobacterium bifidum* strain JCM1254 as a template and the following primers:5'CTCGAGTAAGGAGGTAATA TAATGGTCATCGCCAGTGTCGAGGACG and 5'AAGCTTAGGCGCTCGCCTTCTTCGTGATCGTGTAC. The amplified fragment was first cloned into pCR4Blunt-TOPO vector (Invitrogen) and then sub-cloned into the *XhoI* and *HindIII* sites of the pSU2718 expression vector to form pSU-fucase (Martinez et al. 1988).

Construction of pBBRGAB plasmid was previously described (Dumon et al. 2006).

Production of deuterated fucose in batch-fermentation

Precultures of the fucose-producing strain were first grown in LB medium. All culture media were supplemented with 15 μg mL^{-1} tetracycline, 20 μg mL^{-1} chloramphenicol, 50 μg mL^{-1} kanamycin and 4 μg mL^{-1} thiamine. All cultures were grown at 28 °C with shaking at 160 rpm. The strain was adapted to deuterated minimal medium with the following composition: 5 g L^{-1} NH$_4$H$_2$PO$_4$, 5 g L^{-1} KH$_2$PO$_4$, 0.5 g L^{-1} C$_6$H$_8$O$_7$ (citric acid), 1.65 g L^{-1} KOH, 0.65 g L^{-1} NaOH, 7.5 mL trace mineral solution [13 g L^{-1} N(CH$_2$CO$_2$H)$_3$ (nitrilotriacetic acid), 7 g L^{-1} KOH, 7.5 g L^{-1} C$_6$H$_5$FeO$_7$ (ferric citrate), 1.3 g L^{-1} MnCl$_2$.4H$_2$O, 1.2 g L^{-1} ZnSO$_4$.7H$_2$O, 0.25 g L^{-1} H$_3$BO$_3$, 0.15 g L^{-1} Na$_2$MoO$_4$.2H$_2$O, 0.21 g L^{-1} CoCl$_2$. 6H$_2$O, 0.13 g L^{-1} CuCl$_2$.2H$_2$O], 45 g glycerol-d$_8$ (Eurisotop). A single colony of *E. coli* FUC5 cells containing pBP-*futC*, pBBRGAB and pSU-fucase plasmids grown overnight on LB agar plates supplemented with the 3 antibiotics was used to inoculate 15 mL of minimal medium and was grown overnight at 28 °C. The culture was then used to inoculate 15 mL of 100% D$_2$O minimal medium (with deuterated glycerol-d$_8$) at OD$_{600}$ of 0.1 and was grown overnight. This step was repeated five times until the doubling time for *E. coli* reached values similar to those for hydrogenated cultures. The last deuterated preculture of 200 ml was used to inoculate the 1.5 L of

deuterated minimal medium in a 3 L bioreactor used for batch-fermentation. The pD of the culture medium was regulated at 7.2 by addition of 4% NaOD. The temperature was maintained at 28 °C. IPTG inducer and lactose acceptor were added to a final concentration of 0.2 mM and 0.5 g L^{-1} respectively at the inoculation time. The fermentation was stopped after consumption of the deuterated glycerol-d_8 from the culture medium.

Analytical methods
Cell growth was monitored by measuring turbidity at 600 nm (OD_{600}) using a spectrophotometer. Culture aliquots (1 mL) were taken at different time points throughout the high cell-density culture. Samples were centrifuged and the supernatants were treated as the extracellular fraction. The pellets were resuspended in 1 mL of Milli-Q water and boiled for 20 min at 100 °C. After centrifugation, the supernatants were treated as the intracellular fraction. Samples were analysed using thin layer chromatography on silica gels with an eluent composed of 1-propanol: acetic acid: water in 2:1:1 ratio. After dipping the plate in orcinol- sulphuric acid reagent, the spots were visualized by heating at 100 °C.

Purification of deuterated fucose
At the end of the fermentation, the bacterial cells were recovered by centrifugation (8000 g for 1 h at 8 °C). The supernatant containing the deuterated fucose is referred to as extracellular fraction. The purification was performed in two steps. The first step was carried out using ion-exchange chromatography. The crude extracellular fraction was treated with a strongly acidic cation-exchanger resin (Amberlite IR120 hydrogen form, Sigma-Aldrich). The resin was added progressively until the pH dropped below 3. After decantation, precipitated proteins were eliminated by centrifugation (16 000 g for 1 h at 8 °C). The supernatant was filtered and loaded

onto a column with a strongly acidic cation-exchanger resin (DOWEX 50WX4 hydrogen form, Sigma-Aldrich) and immediately neutralized by passing through a weakly basic anion-exchanger resin (DOWEX 66 free base, Sigma-Aldrich). The final eluate was colourless with neutral pH. The purity and amount of fucose were analyzed by thin layer chromatography. Fractions containing fucose were pooled together, sterile filtered and lyophilized. The resulting mix was further fractionated by size-exclusion chromatography using a semi- preparative TOYOPEARL® HW-40 column (Tosoh Bioscience GmbH) equipped with a refractive index detector (Knauer). The chromatography was performed at ambient temperature with water as a mobile phase and a flow rate of 120 mL h^{-1}. Fractions containing pure product were pooled together and submitted to three cycles of resuspension in D_2O and freeze-drying.

Quantification of hydrogenated and deuterated L-fucose
Quantification of the intracellular and extracellular fucose was carried out on the cell paste and crude supernatant samples, respectively, using an enzymatic L-fucose assay kit (K-FUCOSE, Megazyme). The assay is based on the oxidation of L-fucose to L-fucono-1,5-lactone by L-fucose dehydrogenase in the presence of nicotinamide adenine dinucleotide phosphate (NADP+). The amount of L-fucose is stoichiometric to the production of reduced nicotinamide adenine dinucleotide phosphate (NADPH), measured by an increase of absorbance at 340 nm. The final amount of the deuterated fucose obtained was quantified by weighing the lyophilized powder after purification.

Structural analysis of deuterated fucose
Positive and negative-ion ESImass spectrometry data were recorded on the Bruker amazon speed ion-trap mass spectrometer. A sample of deuterated

fucose was solubilized in D_2O to a final concentration of 50 μM.

The proton and deuterium NMR spectra were recorded with a Bruker Avance™ III NMR spectrometer operating at a frequency of 400.13 MHz for ^{1}H and 61.42 MHz for [2]H. For the ^{1}H NMR experiments, samples of hydrogenated and deuterated fucose were solubilized in D_2O with 1.82 g L^{-1} sodium succinate as a standard. Residual signal of the solvent was used as the internal standard: HOD at 4.8 ppm at 298 K. The ^{1}H spectra were recorded at 298 K, with a 4006 Hz spectral width, 32 768 data points, 4.096 s acquisition times, 10 s relaxation delays and 16 scans.

The ^{2}H NMR spectrum was measured on a sample of 7 mg of deuterated fucose dissolved in H_2O at pH 6.7. The ^{2}H spectrum was acquired at 298 K, without broadband [2]H-decoupling, using a 90° pulse, 10 015 Hz sweep width, 3030 number of data points, 1 s acquisition time and 1024 number of scans.

Abbreviations
ESI, electrospray ionization; Fuc, fucose; FUT2, fucosyltransferase 2; IPTG, isopropyl-β-D-thiogalactopyranoside; LB, Luria Bertani medium; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; SEC, size-exclusion chromatography; TLC, thin layer chromatography.

References

Albermann C, Piepersberg W, Wehmeier UF. 2001. Synthesis of the milk oligosaccharide 2'-fucosyllactose using recombinant bacterial enzymes. *Carbohydr Res*. 334:97-103.

Blakeley MP, Podjarny AD. 2018. Neutron macromolecular crystallography. *Emerging Topics in Life Sciences*. 2:39-55.

Bode L. 2015. The functional biology of human milk oligosaccharides. *Early Hum Dev*. 91:619-622.

Chorney W, Scully NJ, Crespi HL, Katz JJ. 1960. The growth of algae in deuterium oxide. *Biochim Biophys Acta*. 37:280-287.

Crespi HL, Archer SM, Katz JJ. 1959. Culture of algae and other micro-organisms in deuterium oxide. *Nature*. 184(Suppl 10):729-730.

Cress BF, Bhaskar U, Vaidyanathan D, Williams A, Cai C, Liu X, Fu L, V MC, Zhang F, Mousa SA, *et al.* 2019. Heavy Heparin: A Stable Isotope-Enriched, Chemoenzymatically-Synthesized, Poly-Component Drug. *Angew Chem Int Ed Engl*. 58:5962-5966.

Cuypers MG, Mason SA, Blakeley MP, Mitchell EP, Haertlein M, Forsyth VT. 2013a. Near-atomic resolution neutron crystallography on perdeuterated *Pyrococcus furiosus* rubredoxin: implication of hydronium ions and protonation state equilibria in redox changes. *Angew Chem Int Ed Engl*. 52:1022-1025.

Cuypers MG, Mason SA, Mossou E, Haertlein M, Forsyth VT, Mitchell EP. 2016 Macromolecular structure phasing by neutron anomalous diffraction. *Sci. Rep*. 6:31487.

Cuypers MG, Trubitsyna M, Callow P, Forsyth VT, Richardson JM. 2013b. Solution conformations of early intermediates in Mos1 transposition. *Nucleic Acids Res*. 41:2020-2033.

Dajnowicz S, Johnston RC, Parks JM, Blakeley MP,

Keen DA, Weiss KL, Gerlits O, Kovalevsky A, Mueser TC. 2017. Direct visualization of critical hydrogen atoms in a pyridoxal 5'-phosphate enzyme. *Nat Commun*. 8:955.

De Castro C, Molinaro A, Piacente F, Gurnon JR, Sturiale L, Palmigiano A, Lanzetta R, Parrilli M, Garozzo D, Tonetti MG, *et al.* 2013. Structure of N-linked oligosaccharides attached to chlorovirus PBCV-1 major capsid protein reveals unusual class of complex N- glycans. *Proc Natl Acad Sci U S A*. 110:13956-13960.

Dean M, Sung VW. 2018. Review of deutetrabenazine: a novel treatment for chorea associated with Huntington's disease. *Drug Des Devel Ther*. 12:313-319.

Drouillard S, Driguez H, Samain E. 2006. Large-scale synthesis of H-antigen oligosaccharides by expressing *Helicobacter pylori* alpha1,2-fucosyltransferase in metabolically engineered *Escherichia coli* cells. *Angew Chem Int Ed Engl*. 45:1778-1780.

Dumon C, Bosso C, Utille JP, Heyraud A, Samain E. 2006. Production of Lewis x tetrasaccharides by metabolically engineered *Escherichia coli*. *Chembiochem*. 7:359- 365.

Dunne O, Weidenhaupt M, Callow P, Martel A, Moulin M, Perkins SJ, Haertlein M, Forsyth VT. 2017. Matchout deuterium labelling of proteins for small-angle neutron scattering studies using prokaryotic and eukaryotic expression systems and high cell-density cultures. *Eur Biophys J*. 46:425-432.

Grage SL, Keleshian AM, Turdzeladze T, Battle AR, Tay WC, May RP, Holt SA, Contera SA, Haertlein M, Moulin M, *et al.* 2011. Bilayer-mediated clustering and functional interaction of MscL channels. *Biophys J*. 100:1252-1260.

Gupta M, Zha J, Zhang X, Jung GY, Linhardt RJ, Koffas MAG. 2018. Production of Deuterated Cyanidin 3-O-Glucoside from Recombinant *Escherichia coli*. *ACS Omega*. 3:11643-11648.

Haertlein M, Moulin M, Devos JM, Laux V, Dunne O, Forsyth VT. 2016. Biomolecular Deuteration for Neutron Structural Biology and Dynamics. *Methods Enzymol*. 566:113- 157.

Haupt M, Blakeley MP, Fisher SJ, Mason SA, Cooper JB, Mitchell EP, Forsyth VT. 2014. Binding site asymmetry in human transthyretin: insights from a joint neutron and X- ray crystallographic analysis using perdeuterated protein. *IUCrJ*. 1:429-438.

Heggelund JE, Varrot A, Imberty A, Krengel U. 2017. Histo-blood group antigens as mediators of infections. *Curr Opin Struct Biol*. 44:190-200.

Hellstrand E, Grey M, Ainalem ML, Ankner J, Forsyth VT, Fragneto G, Haertlein M,
Dauvergne MT, Nilsson H, Brundin P, *et al.* 2013. Adsorption of alpha-synuclein to supported lipid bilayers: positioning and role of electrostatics. *ACS Chem Neurosci*. 4:1339-1351.

Josts I, Nitsche J, Maric S, Mertens HD, Moulin M, Haertlein M, Prevost S, Svergun DI, Busch S, Forsyth VT, *et al.* 2018. Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs. *Structure*. 26:1072-1079 e1074.

Jung SM, Park YC, Seo JH. 2019. Production of 3-Fucosyllactose in Engineered *Escherichia coli* with alpha-1,3-Fucosyltransferase from *Helicobacter pylori*. *Biotechnol J*. 14:e1800498.

Katayama T, Sakuma A, Kimura T, Makimura Y, Hiratake J, Sakata K, Yamanoi T, Kumagai H, Yamamoto K. 2004. Molecular cloning and characterization of *Bifidobacterium bifidum* 1,2-alpha-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95). *J Bacteriol*. 186:4885-4893.

Koch HJ, Stuart RS. 1978. The synthesis of per-*C*-deuterated D-glucose. *Carbohydrate Research*. 64:127-134.

Kwon H, Basran J, Devos JM, Suardiaz R, van der Kamp MW, Mulholland AJ, Schrader TE, Ostermann A, Blakeley MP, Moody PCE, *et al.* 2020. Visualizing the protons in a metalloenzyme electron proton transfer pathway. *Proc Natl Acad Sci U S A*. 117:6484- 6490.

Laux V, Callow P, Svergun DI, Timmins PA, Forsyth VT, Haertlein M. 2008. Selective deuteration of tryptophan and methionine residues in maltose binding protein: a model system for neutron scattering. *Eur Biophys J*. 37:815-822.

Link AJ, Phillips D, Church GM. 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol*. 179:6228-6237.

Liu JJ, Kwak S, Pathanibul P, Lee JW, Yu S, Yun EJ, Lim H, Kim KH, Jin YS. 2018. Biosynthesis of a Functional Human Milk Oligosaccharide, 2'-Fucosyllactose, and l- Fucose Using Engineered *Saccharomyces cerevisiae*. *ACS Synth Biol*. 7:2529-2536.

Liu JJ, Lee JW, Yun EJ, Jung SM, Seo JH, Jin YS. 2019. L-Fucose production by engineered *Escherichia coli*. *Biotechnol Bioeng*. 116:904-911.

Lowery R, Gibson MI, Thompson RL, Fullam E. 2015. Deuterated carbohydrate probes as 'label-free' substrates for probing nutrient uptake in

mycobacteria by nuclear reaction analysis. *Chem Commun (Camb)*. 51:4838-4841.

Maric S, Lind TK, Raida MR, Bengtsson E, Fredrikson GN, Rogers S, Moulin M, Haertlein M, Forsyth VT, Wenk MR, *et al.* 2019. Time-resolved small-angle neutron scattering as a probe for the dynamics of lipid exchange between human lipoproteins and naturally derived membranes. *Sci Rep*. 9:7591.

Martinez E, Bartolome B, de la Cruz F. 1988. pACYC184-derived cloning vectors containing the multiple cloning site and lacZ alpha reporter gene of pUC8/9 and pUC18/19 plasmids. *Gene*. 68:159-162.

Meilleur F, Weiss KL, Myles DA. 2009. Deuterium labeling for neutron structure-function-dynamics analysis. *Methods Mol Biol*. 544:281-292.

Moulin M, Strohmeier GA, Hirz M, Thompson KC, Rennie AR, Campbell RA, Pichler H, Maric S, Forsyth VT, Haertlein M. 2018. Perdeuteration of cholesterol for neutron scattering applications using recombinant *Pichia pastoris*. *Chem Phys Lipids*. 212:80-87.

O'Neill H, Shah R, Evans BR, He J, Pingali SV, Chundawat SP, Jones AD, Langan P, Davison BH, Urban V. 2015. Production of bacterial cellulose with controlled deuterium-hydrogen substitution for neutron scattering studies. *Methods Enzymol*. 565:123-146.

Painter TJ, Watkins WM, Morgan WT. 1965. Serologically active fucose-containing oligosaccharides isolated from human blood-group A and B substances. *Nature*. 206:594-597.

Priem B, Gilbert M, Wakarchuk WW, Heyraud A, Samain E. 2002. A new fermentation process allows large-scale production of human milk oligosaccharides by metabolically engineered bacteria. *Glycobiology*. 12:235-240.

Samain E, Drouillard S, Heyraud A, Driguez H, Geremia RA. 1997. Gram-scale synthesis of recombinant chitooligosaccharides in *Escherichia coli*. *Carbohydr Res*. 302:35-42.

Sawama Y, Yabe Y, Iwata H, Fujiwara Y, Monguchi Y, Sajiki H. 2012. Stereo- and regioselective direct multi-deuterium-labeling methods for sugars. *Chemistry*. 18:16436-16442.

Schmidt C. 2017. First deuterated drug approved. *Nat Biotechnol*. 35:493-494.

Schneider M, Al-Shareffi E, Haltiwanger RS. 2017. Biological functions of fucose in mammals. *Glycobiology*. 27:601-618.

Shimba S, Unno K, Okada S. 1990. Study on the biodistribution of deuterated biomolecules in mice aiming at new diagnostic radio-imaging agents. *Chem Pharm Bull (Tokyo)*. 38:2610-2613.

Timmins GS. 2017. Deuterated drugs; updates and obviousness analysis. *Expert Opin Ther Pat*. 27:1353-1361.

Vanhooren PT, Vandamme EJ. 1999. L-Fucose: occurrence, physiological role, chemical, enzymatic and microbial synthesis. *Journal of Chemical Technology & Biotechnology*. 74:479-497.

Waldie S, Lind TK, Browning K, Moulin M, Haertlein M, Forsyth VT, Luchini A, Strohmeier GA, Pichler H, Maric S, *et al.* 2018. Localization of Cholesterol within Supported Lipid Bilayers Made of a Natural Extract of Tailor-Deuterated Phosphatidylcholine. *Langmuir*. 34:472-479.

Wei L, Yu Y, Shen Y, Wang MC, Min W. 2013. Vibrational imaging of newly synthesized proteins in live cells by stimulated Raman scattering microscopy. *Proc Natl Acad Sci U S A*. 110:11226-11231.

Yee AW, Aldeghi M, Blakeley MP, Ostermann A, Mas PJ, Moulin M, de Sanctis D, Bowler MW, Mueller-Dieckmann C, Mitchell EP, *et al.* 2019. A molecular mechanism for transthyretin amyloidogenesis. *Nat Commun*. 10:925.

Yu S, Liu JJ, Yun EJ, Kwak S, Kim KH, Jin YS. 2018. Production of a human milk oligosaccharide 2'-fucosyllactose by metabolically engineered *Saccharomyces cerevisiae*. *Microb Cell Fact*. 17:101.
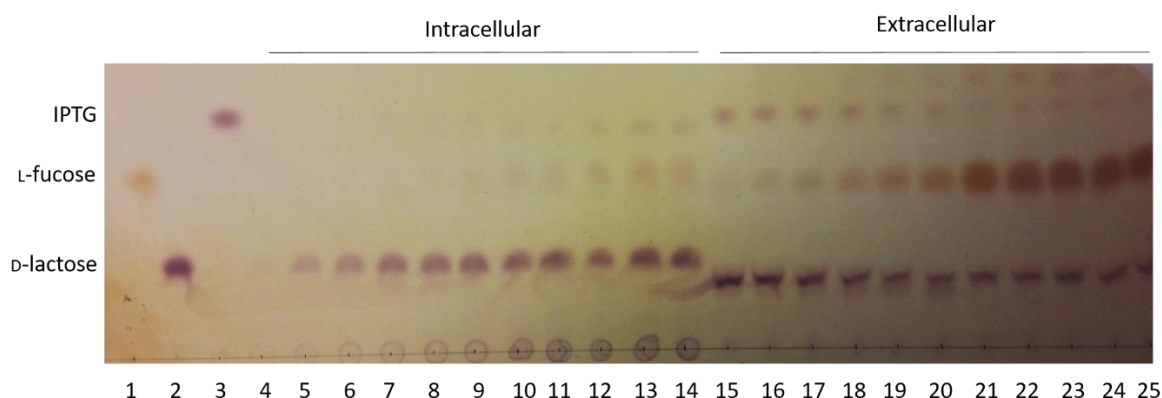
Figure S1: TLC analysis of intracellular and extracellular fractions of the hydrogenated high cell-density culture of *E. coli* strain FUC5 in a bioreactor. Lanes 1-3: 1 mg mL^{-1} standard solutions of L-fucose, D-lactose and IPTG respectively. Lanes 4-14: intracellular fractions withdrawn 40, 43, 45, 47, 50, 52, 64, 68, 70, 72 and 73 hours after the start of the fermentation. Lanes 15-25: extracellular fractions withdrawn 40, 43, 45, 47, 50, 52, 64, 68, 70, 72 and 73 hours after the start of the fermentation.
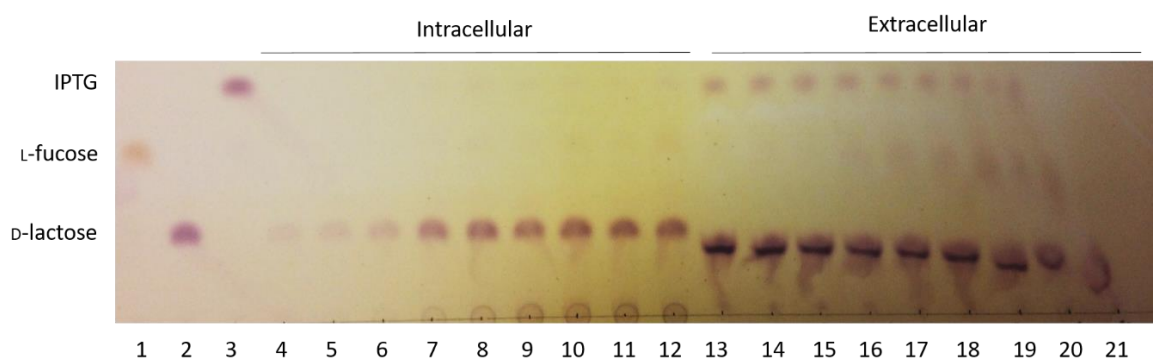
Figure S2: TLC analysis of intracellular and extracellular fractions of the deuterated high cell-density culture of *E. coli* strain FUC5 in a bioreactor. Lanes 1-3: 1 mg mL^{-1} standard solutions of L-fucose, D-lactose and IPTG respectively. Lanes 4-12: intracellular fractions withdrawn 40, 48, 56, 64, 72, 80, 88, 96 and 112 hours after the start of the fermentation. Lanes 13-21: extracellular fractions withdrawn 40, 48, 56, 64, 72, 80, 88, 96 and 112 hours after the start of the fermentation.

4

Neutron structures of apo PLL lectin and PLL/fucose complex

4.1 PLL as a model system for protein-carbohydrate interactions

Neutron protein crystallography plays an important role in structural biology. Direct visualization of hydrogen atoms in lectin-sugar complexes is of crucial importance for a better understanding of interactions that are often involved in pathological processes such as host-pathogen interactions. The fucose-specific PLL lectin from the insect pathogen *Photorhabdus laumondii* (formerly *Photorhabdus luminescens*) was selected as a model system. In the framework of a previous collaboration with Prof. Michaela Wimmerová (CEITEC, Brno, Czech Republic), it was observed that large crystals of PLL (with edge dimensions of more than 1 mm) could be reproducibly obtained. The X-ray structure of the complex was solved demonstrating that PLL binds fucose via direct hydrogen bonds between sugar hydroxyl groups and polar amino acids as well as stacking interactions with aromatic amino acids. Water molecules were also observed to bridge the interactions.

4.2 Methodology for neutron studies of PLL

The *E. coli* cells (containing plasmid pET29a-*pll* with a 6xHis-tag on the C-terminus) were adapted to D_2O and glycerol-d_8 using the following strategy. First, cells grown in LB medium were adapted to hydrogenated Enfors minimal medium. The expression of

PLL was optimized by varying expression conditions such as temperature, concentration of the IPTG inducer and induction time. Cells were subsequently adapted to the fully-deuterated Enfors minimal medium. After adaptation and optimization of the final expression conditions, we proceeded with the fermentation under deuterated conditions. Deuterated PLL lectin was produced in large amount (~180 mg per fermenter) using high cell-density culture (Figure 42), which resulted in about 40 g of wet cell paste. Both H- and D-PLL were purified by one-step affinity chromatography using the same protocol with hydrogenated buffers (Figure 43 and 44).
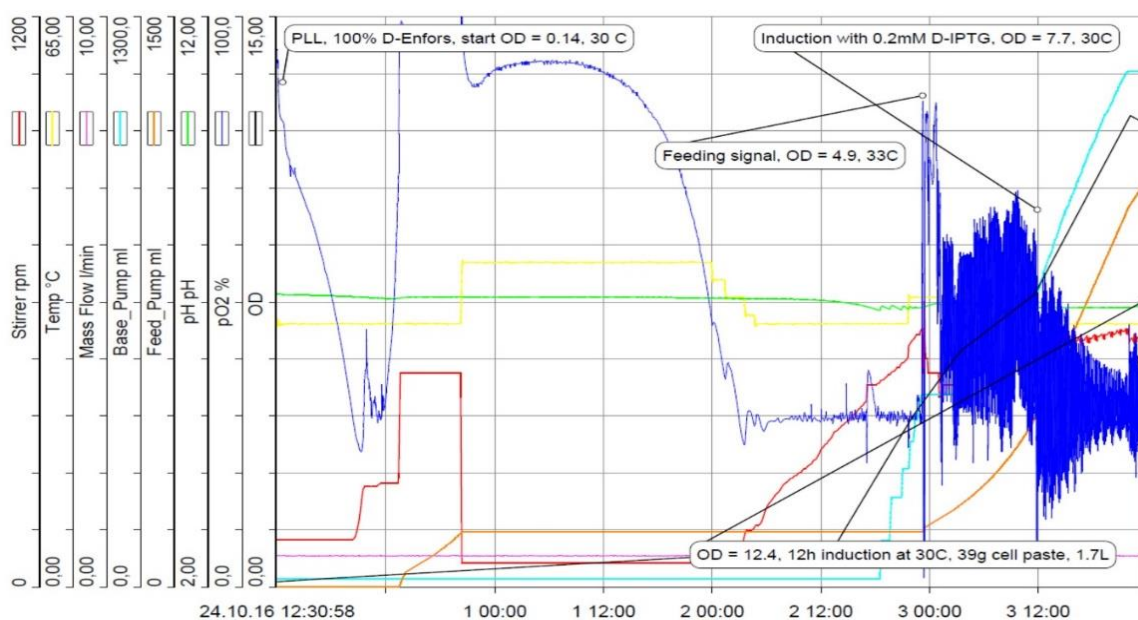
Figure 42: High cell-density culture profile of the batch/fed-batch fermentation process to produce perdeuterated PLL lectin. Deuterated Enfors minimal medium was used for the fermentation. Details of the fermentation process are described in Materials and methods in Article II.
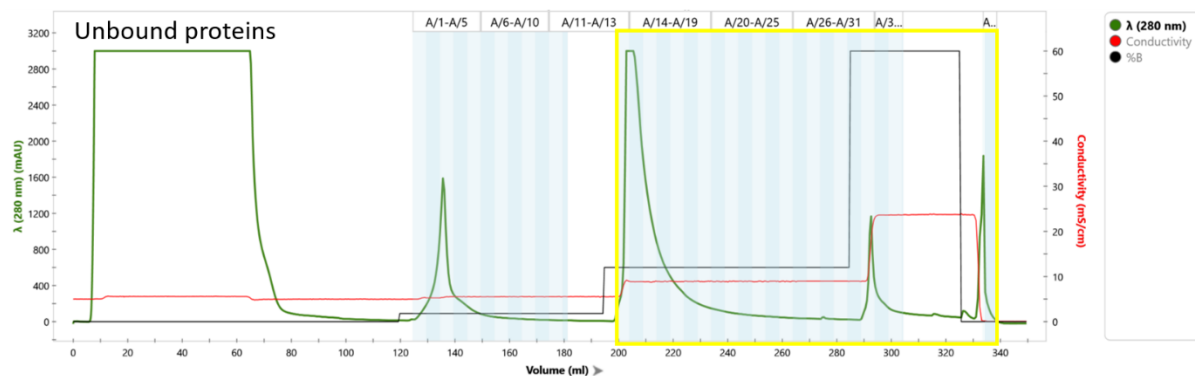
Figure 43: A typical chromatogram of the purification of recombinant D-PLL lectin (containing a polyhistidine tag on the C-terminus) using a Ni^{2+}-Sepharose affinity chromatography. Protein was eluted with a buffer (20 mM potassium phosphate, pH 7.5, 2 mM trehalose) containing 500 mM imidazole solution in a stepwise manner. The first peak (which corresponds to a wash step) was eluted with 3% imidazole, the second peak with 20% and the third one with 100% imidazole. The final peak corresponds to the elution with water.
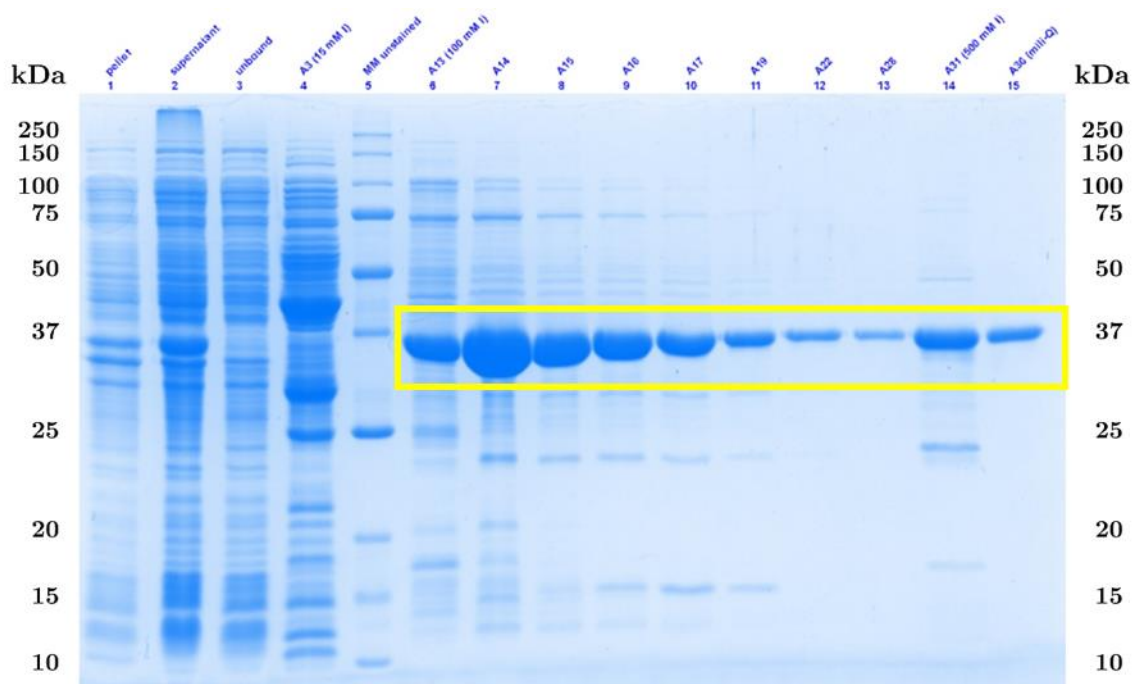
Figure 44: SDS-PAGE (12% Tris-glycine gel) of expression and purification of D-PLL lectin. Lane 1 (pellet, insoluble protein), lane 2 (supernatant, soluble protein), lane 3 (unbound proteins), lane 4 (elution with 3% imidazole), lane 5 (molecular weight marker), lane 6-13 (elution with 20% imidazole), lane 14 (elution with 100% imidazole), lane 15 (elution with mili-Q). Yellow box indicates bands that correspond to the of D-PLL protein.

More than 1000 sitting drops have been set up manually during crystallization experiments of hydrogenated and deuterated PLL (H- and D-PLL respectively) based on crystallization conditions screened at the High Throughput Crystallization Facility at EMBL (Grenoble, France). The protein crystallized easily in various crystallization conditions (Figure 45). The crystallization of H-PLL was rather straightforward and did not require specific optimization other that slight variations in concentration of crystallization reagents. Exchangeable hydrogen atoms were exchanged for deuterium atoms by vapour diffusion. Originally, the drop containing the crystal was exchanged for a fully-deuterated condition. This resulted in a crystal damage with visible cracking. In order to diminish the stress on the crystal, we increased the concentration of D_2O in a stepwise manner. The reservoir solution was replaced by 25%, 50%, 75% and finally 100% D_2O-based crystallization condition with exchanges being carried out about 1 week apart. The final 100% deuterated condition was changed three times prior to mounting the crystal into a quartz capillary. On the other hand, the crystallization of D-PLL had to be optimized since protein crystallized in showers of small crystals. In order to grow larger crystals, macro-seeding and feeding techniques have been applied. Using these techniques, crystals with volumes of 0.7 mm^3 could be grown within several (2-4) months (Figure 46).

0.1 M sodium acetate, pH 4.6
8% (w/v) PEG 4K

0.1 M sodium acetate, pH 4.6
8% (w/v) PEG 4K

0.4 M Na/K tartrate

0.4 M Na/K tartrate

0.4 M Na/K tartrate

0.2 M Na/K tartrate

0.4 M Na/K tartrate

0.2 M Na/K tartrate

0.4 M Na/K tartrate

0.2 M Na/K tartrate

0.1 M Tris/HCl, pH 8.5
4% (w/v) PEG 8K

0.4 M Na/K tartrate
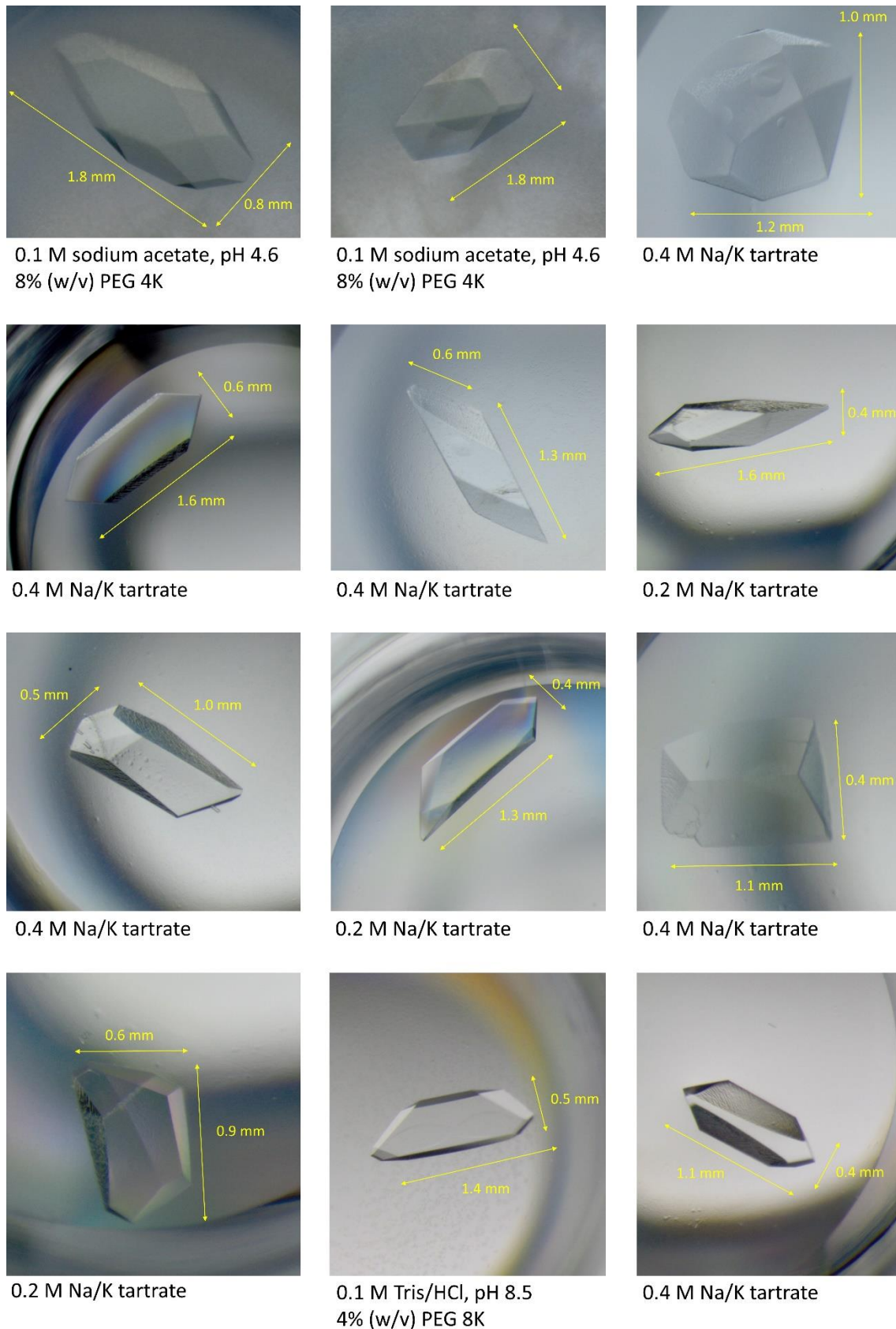
Figure 45: Crystals of H-PLL grown using a sitting-drop vapour-diffusion method.

0.2 M Na/K tartrate   0.05 M Na/K tartrate   0.1 M Na/K tartrate
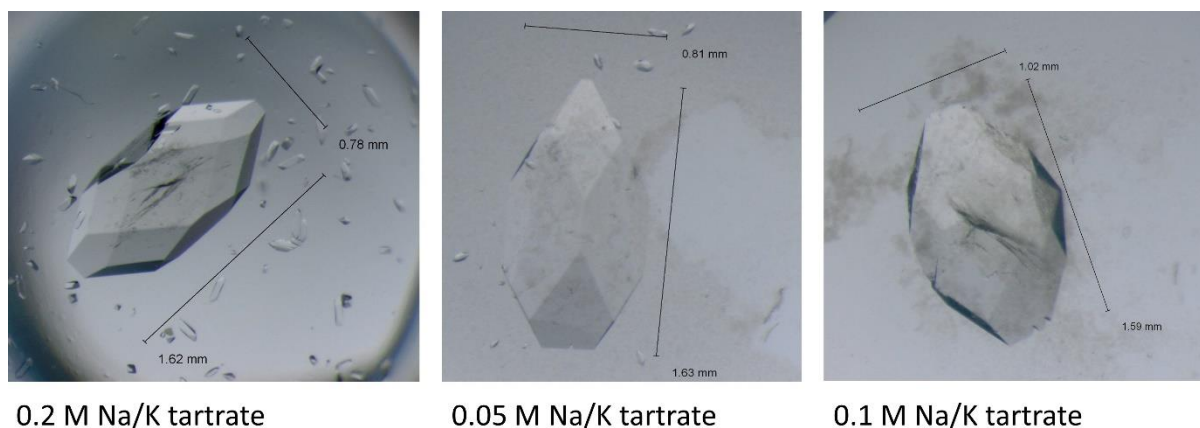
Figure 46: Crystals of D-PLL. Crystals were grown using macro-seeding and feeding techniques (in a sitting-drop vapour-diffusion set-up).

The mounting the crystals into quartz capillaries was very problematic for H/D-exchanged PLL that crystallized in the sitting-drops and needed a period of several months to reach their final size. Almost all of these crystals sank to the bottom of the crystallization well and were stuck to it. Unfortunately, it was not possible to separate the crystals from the bottom of the wells without damaging them during the harvesting procedure. In order to avoid this problem in the next crystallization trials, the sitting wells were coated with parafilm prior to crystallization. Harvesting D-PLL crystals was less problematic since crystals had been already transferred during macro-seeding and regularly checked to prevent that they adhere to the plastic surface

Total of three neutron beamtimes have been allocated to collect neutron diffraction data from PLL lectin. The first neutron data collection was from a 0.7 mm³ crystal of a buffer H/D-exchanged apo PLL lectin. The crystal was mounted in a quartz capillary with D_2O-based buffer surrounding the crystal from both sides of the capillary. After about five days from the start of the data collection, the crystal stopped diffracting. We noticed that the crystal was surrounded by a large drop of liquid that appeared after several days into the data collection (Figure 47) and induced the crystal to dissolve. Water condensation in the form of droplets was observed on the inner wall of the capillary (Figure 47). Such situation had not been noticed before. Possible explanations were suggested involving possible temperature change generated

by the instrument which led to water condensation inside the capillary resulting in the dissolution of the crystal. The neutron data collection resulted in an incomplete data (7 images) set that was completed in the next data collections (see below).

Figure 47: The crystal of H/D-exchanged apo PLL used for the neutron data collection. The crystal dissolved inside the quartz capillary during the data collection after collecting seven diffraction images of 18 h exposures per image.

In the second neutron data collection we tried to reduce the chances of possible failure by optimizing the crystal mounting process. Crystals of D-PLL/Fuc-d$_{12}$ complex were mounted at 19 °C (instead of in a room at a temperature of about 23 °C in the previous experiment). The buffer surrounding the crystal in the quartz capillary was added only on one side of the capillary, downstream of the crystal (the capillary is held in a vertical position inside the diffractometer). All capillaries were properly sealed using bee wax. After about two days, the data collection was aborted due to the loss of the diffraction from the crystal. The crystal turned visibly opaque possibly due to the dehydration (Figure 48). After about a week, the crystal was observed to be transparent again, possibly because of the rehydration that occurred in the sealed capillary although the crystal did not diffract neutrons anymore.

During this data collection, a crystal of H/D-exchanged apo PLL (grown from the same crystallization condition as the first crystal described above) was mounted and 3 images of 18 h exposure could be collected. Two partial neutron datasets of the H/D-exchanged apo PLL lectin could be later merged together resulting in a dataset of a sufficient quality for the data analysis.
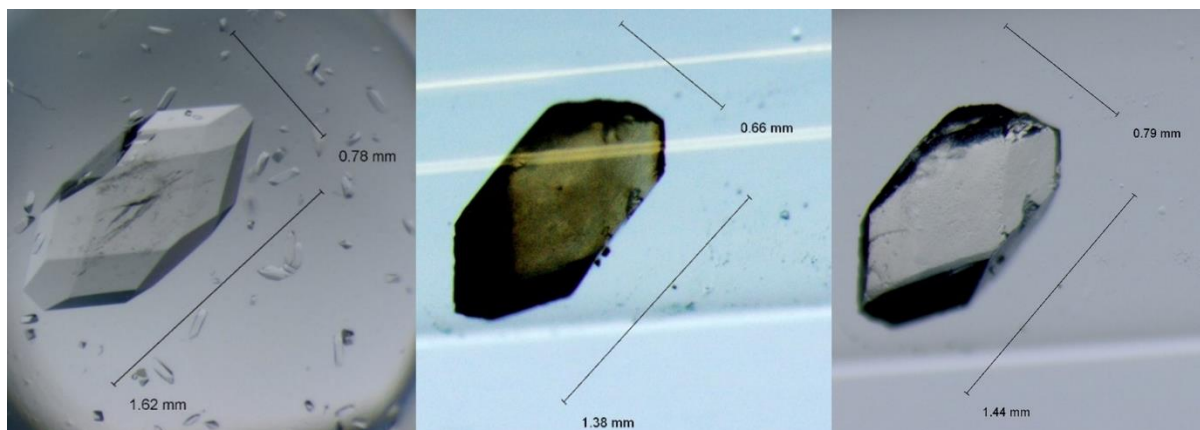
Figure 48: Crystal of D-PLL/Fuc-d_{12} complex used for the neutron data collection. The crystal stopped diffracting after two days from the start of the data collection and turned opaque (brownish under the optical microscope) possibly due to the dehydration caused by an increase of temperature during the data collection. The crystal was kept in the sealed capillary at 19 °C and rehydrated in a week.

Finally, the third neutron data collection was from a crystal of the D-PLL/Fuc-d_{12} complex. This time, several crystals were available and tested. Diffraction patterns of the first five crystals exhibited streaked diffraction spots indicating larger mosaic spread (Figure 49). The sixth crystal showed nice single diffraction spots and was used for the full neutron data collection.
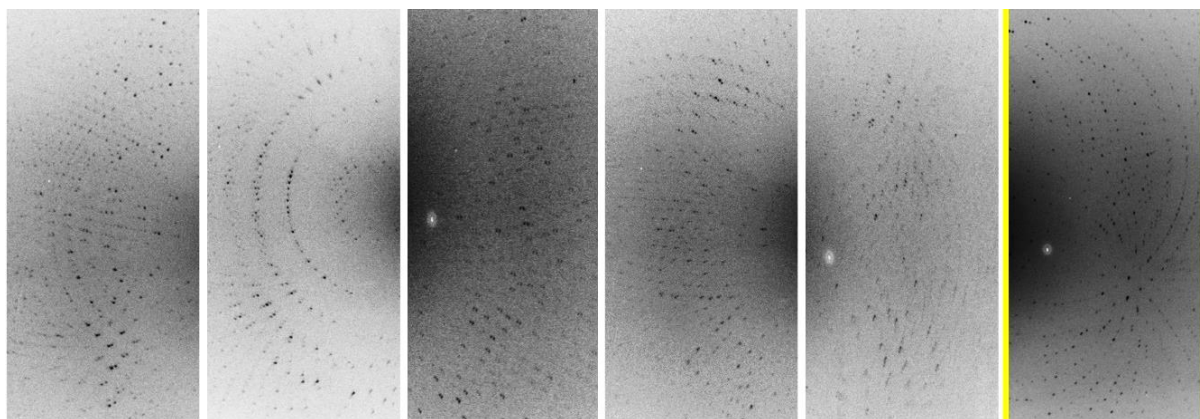
Figure 49: quasi-Laue neutron diffraction patterns collected from crystals of the D-PLL/Fuc-d_{12} complex. The first five crystals showed streaked diffraction spots. The sixth crystal showed single-lattice diffraction spots and was used for the neutron data collection.

To sum up, two room temperature neutron diffraction datasets have been collected from H/D-exchanged apo PLL and D-PLL/Fuc-d_{12} complex on the LADI-III diffractometer at the Institut Laue-Langevin in Grenoble (Figure 50). The room temperature X-ray data have been collected from the same crystals used for the neutron data collection. In addition, 100 K and RT X-ray data have been collected from H- and D-PLL crystals for comparisons (Figure 52).
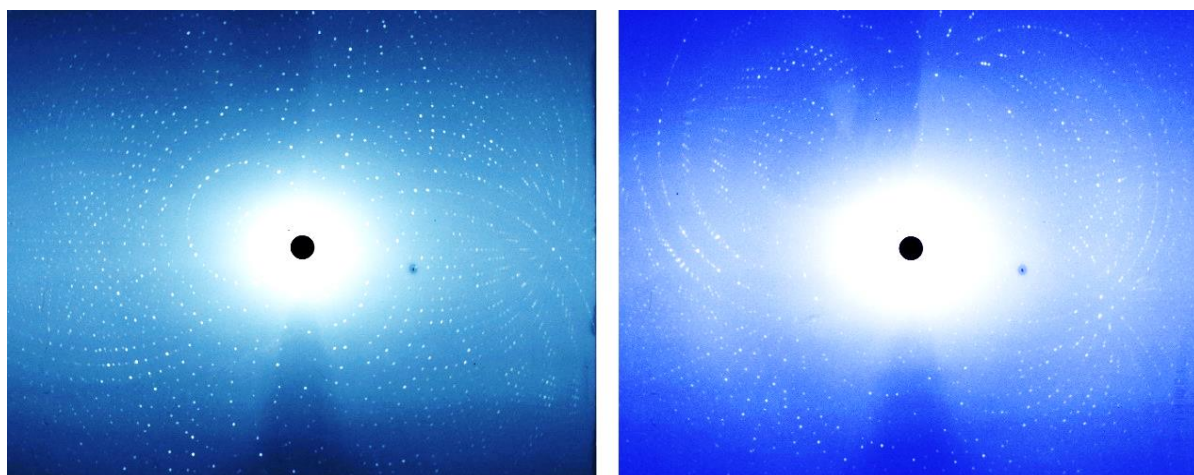
Figure 50: quasi-Laue neutron diffraction patterns of the H/D-exchanged apo PLL (left) and the perdeuterated D-PLL/Fuc-d_{12} complex, both collected on LADI-III diffractometer at ILL (Grenoble, France).

Both H/D-exchanged apo PLL and D-PLL/Fuc-d_{12} structures have been jointly refined against X-ray and neutron data using the *phenix.refine* in the *PHENIX* package and protonation state of histidine residues could be determined (Figure 51). A limitation of *phenix.refine* has arisen during the structure refinement of H/D-exchanged apo PLL. Individual occupancies of H and D atoms on exchangeable sites have been refined to sum up to 1.0. PLL contains more than 20 amino acids that could be refined with two or more conformations. In these, the partially exchanged H/D sites could not be refined individually to sum to 1.0 since *phenix.refine* was giving an error message and those were thus refined using deuterium only on the reasonable assumption that a full H/D-exchange occurred.

H/D-PLL apo   **D-PLL/Fuc-d$_{12}$**   **H/D-PLL apo**   **D-PLL/Fuc-d$_{12}$**

His51

His176

His80

His196

His100

His203

His148

His244

His155

His339
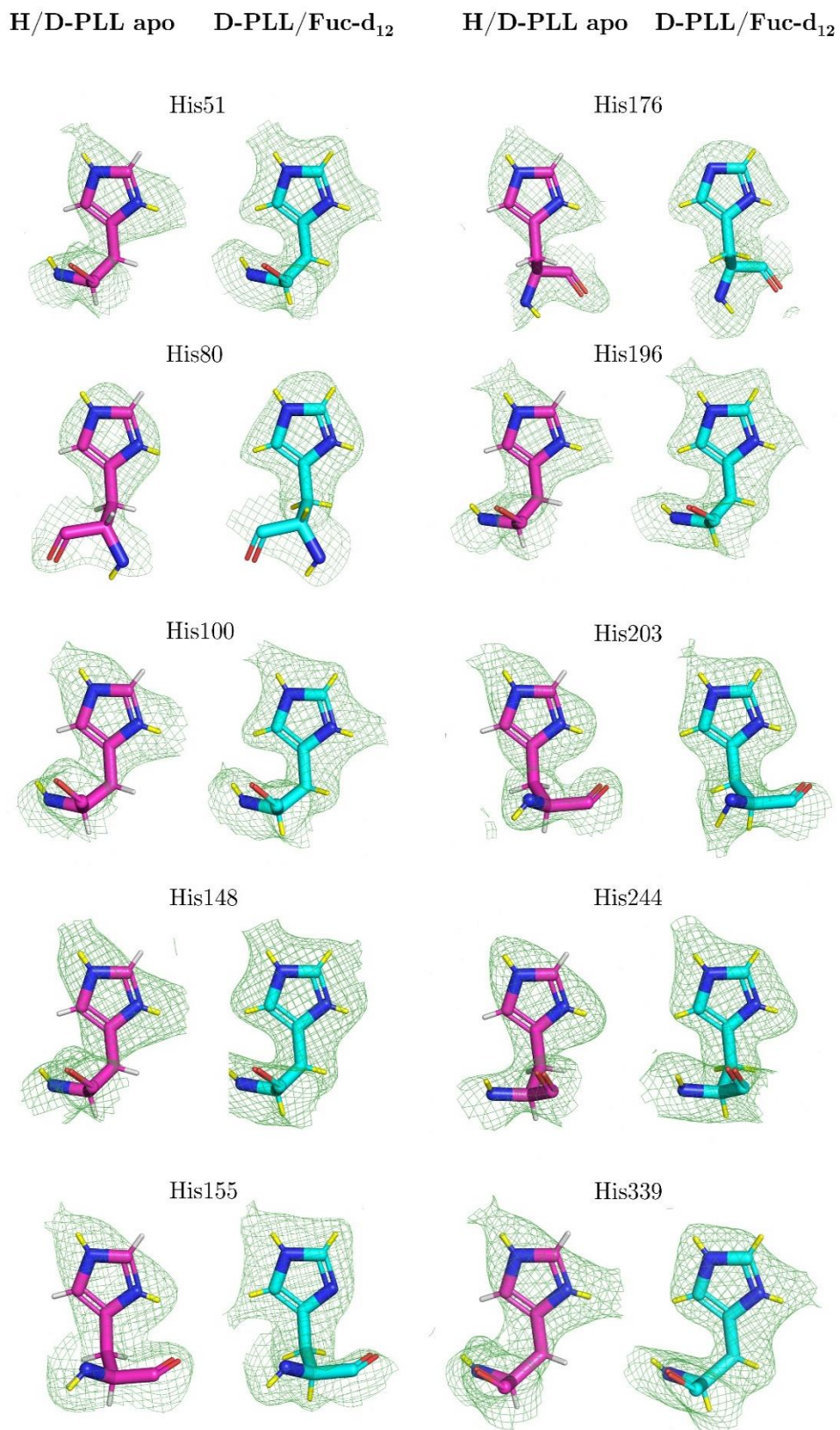
Figure 51: Protonation state of histidine residues and the associated 2mF_{o}-DF_{c} neutron scattering length density map (green mesh) contoured at 1σ. Hydrogen and deuterium atoms are shown as grey and yellow sticks respectively.

X-ray/neutron	T	Resolution [Å]	Completeness [%]	<I/σ(I)>	Unit cell [a, b, c, α, β, γ]	Spacegroup	Protein form	Crystallization form	Ligand	Crystallization condition	Cryoprotectant	Beamline	PDB
Neutron	RT	2.2	76.6 (61.1)	4.6 (3.2)	71.65, 88.00, 156.93, 90, 90, 90	I222	hydrogenated	deuterated	apo	0.1M NaAc, pD 4.6, 8% PEG 4K	-	LADI-III	7BBI
Neutron	RT	2.2	84.6 (73.6)	7.5 (4.1)	71.65, 87.91, 157.08, 90, 90, 90	I222	deuterated	deuterated	deuterated fucose soaked	0.05M Na/K tartrate	-	LADI-III	7BBC
X-ray	RT	1.6	99.9 (100)	25.3	72.63, 89.28, 159.22, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M NaAc, pD 4.6, 6% PEG 4K	-	ESRF FIP (BM30A)	7B7E
X-ray	RT	1.55	99.5 (99.9)	20.6	72.62, 89.07, 159.20, 90, 90, 90	I222	deuterated	deuterated	hydrogenated fucose	0.1M Tris/DCl, pD 8.5, 8% PEG 8K	-	ESRF FIP (BM30A)	7B7C
X-ray	RT	1.7	98.6 (99.6)	12.4	72.68, 89.34, 159.38, 90, 90, 90	I222	hydrogenated	deuterated	apo	0.1M NaAc, pD 4.6, 8% PEG 4K	-	ESRF FIP (BM30A)	7BBI
X-ray	RT	1.6	98.7 (99.2)	16.1	72.60, 89.24, 159.28, 90, 90, 90	I222	hydrogenated	deuterated	hydrogenated fucose	0.4M Na/K tartrate	-	ESRF FIP (BM30A)	7B7F
X-ray	cryo	1.53	98.3 (99)	7.5	71.29, 87.21, 158.00, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Na/K tartrate	30% PEG 400	ESRF ID30B	
X-ray	cryo	1.2	98.3 (91.1)	8.5	71.56, 87.64, 158.45, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.17	96.6 (89.8)	8.2 (1.2)	71.79, 87.75, 158.25, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.18	97.8 (97.2)	8.3 (1.4)	71.58, 87.38, 158.22, 90, 90, 90	I222	deuterated	deuterated	apo	0.18M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.15	98.8 (97)	5.9 (1.2)	71.67, 87.55, 158.14, 90, 90, 90	I222	deuterated	deuterated	apo	0.4M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	2.16	93.6 (84.2)	3.6 (1.1)	70.56, 86.63, 155.55, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Tris/DCl, pD 8.5, 6% PEG 8K	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.39	95.6 (97.4)	5.8 (1)	71.69, 87.96, 158.28, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Tris/DCl, pD 8.5, 15% EtOH	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.08	94.3 (85.9)	6.9 (1.2)	71.72, 87.57, 158.12, 90, 90, 90	I222	deuterated	deuterated	apo	5% EtOH, 0.75M NaCl	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.23	98.5 (98.4)	5.5 (1.1)	71.70, 87.71, 158.30, 90, 90, 90	I222	deuterated	deuterated	apo	5% EtOH, 0.75M NaCl	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.09	96.3 (83.1)	5.6 (1.2)	71.87, 87.67, 158.14, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.2	95.2 (76.2)	7.9 (1.1)	71.61, 87.73, 158.12, 90, 90, 90	I222	deuterated	deuterated	apo	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID30B	
X-ray	cryo	1.67	96.8 (97.1)	7.4 (1.4)	71.80, 87.82, 158.50, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.2M Na/K tartrate	30% glycerol	ESRF ID30B	
X-ray	cryo	1.31	98 (98.2)	6.5 (0.99)	71.82, 87.60, 157.94, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.2M Na/K tartrate	30% glycerol	ESRF ID30B	
X-ray	cryo	1.28	98.3 (98.9)	8.2 (1.3)	71.94, 87.40, 157.75, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.2M Na/K tartrate	30% glycerol	ESRF ID30B	
X-ray	cryo	1.37	98.3 (93.9)	6.1 (1)	71.55, 88.08, 158.36, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.75M NaCl, 5% EtOH	30% glycerol	ESRF ID30B	
X-ray	cryo	1.15	92.8 (63.4)	8.8 (1.1)	71.89, 87.55, 157.98, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.1M Tris/HCl, pH 8.5, 20% EtOH	30% glycerol	ESRF ID30B	
X-ray	cryo	1.2	97.9 (97.3)	7.7 (1.1)	71.68, 87.60, 158.16, 90, 90, 90	I222	hydrogenated	hydrogenated	apo	0.1M Tris/HCl, pH 8.5, 10% EtOH	30% glycerol	ESRF ID30B	
X-ray	cryo	1.21	99.1 (97.8)	7 (1.4)	71.89, 87.88, 158.47, 90, 90, 90	I222	deuterated	hydrogenated	apo	0.3M Na/K tartrate	30% glycerol	ESRF ID23-1	
X-ray	cryo	1.44	97.6 (97.1)	10 (1.6)	71.65, 87.71, 158.23, 90, 90, 90	I222	deuterated	deuterated	hydrogenated fucose soaked	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID23-1	
X-ray	cryo	1.25	96.7 (93.3)	6.5 (0.92)	72.26, 88.00, 158.32, 90, 90, 90	I222	deuterated	deuterated	hydrogenated fucose co-crys	0.1M Na/K tartrate	30% d8-glycerol	ESRF ID23-1	7BB4
X-ray	cryo	1.51	97.9 (98.5)	4.8 (1.3)	72.84, 88.26, 158.26, 90, 90, 90	I222	deuterated	deuterated	hydrogenated fucose co-crys	0.2M Na/K tartrate	30% d8-glycerol	ESRF ID23-1	
X-ray	cryo	1.2	98.8 (99)	6.6 (1.3)	71.77, 87.70, 158.18, 90, 90, 90	I222	deuterated	deuterated	deuterated fucose soaked	0.2M Na/K tartrate	30% d8-glycerol	ESRF ID23-1	
X-ray	cryo	1.28	99.9 (99.3)	13 (1.1)	71.84, 87.64, 157.93, 90, 90, 90	I222	deuterated	deuterated	deuterated fucose soaked	0.1M Na/K tartrate	30% d8-glycerol	Diamond I03	

Figure 52: Data collection summary from PLL lectin

4.3 Article II

Visualisation of hydrogen atoms in a perdeuterated lectin-fucose complex
reveals key details of protein-carbohydrate interactions

L. Gajdos, M. P. Blakeley, A. Kumar, M. Wimmerová, M Haertlein, V. T. Forsyth,
A. Imberty and J. M. Devos. Structure, 2021, https://doi.org/10.1016/j.str.2021.03.003

Resumé

Il a été démontré que les protéines de liaison aux glucides des bactéries et champignons
pathogènes sont impliquées dans divers processus pathologiques, où elles interagissent
avec les glycanes présents à la surface des cellules hôtes. Ces interactions font partie
des processus initiaux d'infection de l'hôte et il est très important de les étudier au
niveau atomique. Nous présentons ici les structures neutroniques à température
ambiante de la lectine PLL de *Photorhabdus laumondii* sous sa forme apo et en
complexe avec le L-fucose deutéré, la première structure neutronique d'une protéine de
liaison aux glucides en complexe avec un ligand glucidique entièrement deutéré. Une
analyse structurelle détaillée des interactions lectine-glucide fournit de nouvelles
informations sur le réseau de liaison hydrogène, le rôle des molécules d'eau et l'étendue
des interactions d'empilement CH-π entre le fucose et les acides aminés aromatiques
dans le site de liaison.
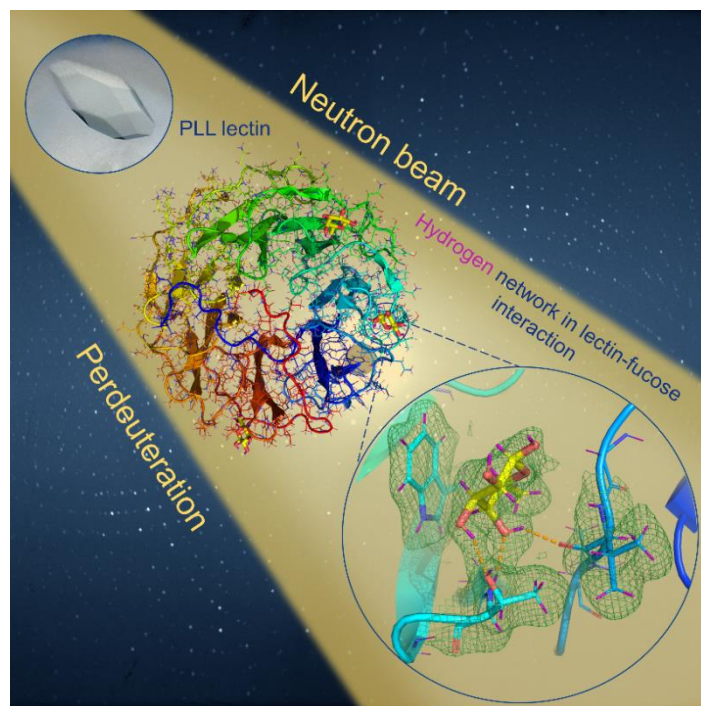
Visualisation of hydrogen atoms in a perdeuterated lectin-fucose complex reveals key details of protein-carbohydrate interactions

Lukas Gajdos[1,2,3], Matthew P. Blakeley[4], Atul Kumar[5,6,7], Michaela Wimmerová[5,6], Michael Haertlein[1,2], V. Trevor Forsyth[1,2,8], Anne Imberty[3]* and Juliette M. Devos[1,2]*

[1] Life Sciences Group, Institut Laue-Langevin, 38000 Grenoble, France
[2] Partnership for Structural Biology (PSB), 38000 Grenoble, France
[3] Université Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France
[4] Large Scale Structures Group, Institut Laue-Langevin, 38000 Grenoble, France
[5] CEITEC, Masaryk University, 625 00 Brno, Czech Republic
[6] NCBR, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic
[7] Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic
[8] Faculty of Natural Sciences, Keele University, ST5 5BG Staffordshire, UK

* To whom correspondence should be addressed. Anne Imberty (anne.imberty@cermav.cnrs.fr), Juliette Devos (devosj@ill.fr).

Summary

Carbohydrate-binding proteins from pathogenic bacteria and fungi have been shown to be implicated in various pathological processes, where they interact with glycans present on the surface of the host cells. These interactions are part of the initial processes of infection of the host and are very important to study at the atomic level. Here we report the room temperature neutron structures of PLL lectin from *Photorhabdus laumondii* in its apo form and in complex with deuterated L-fucose, that is to our knowledge, the first neutron structure of a carbohydrate-binding protein in complex with a fully deuterated carbohydrate ligand. A detailed structural analysis of the lectin-carbohydrate interactions provides new information on the hydrogen bond network, the role of water molecules, and the extent of the CH-π stacking interactions between fucose and the aromatic amino acids in the binding site.

Keywords: Lectin, ligand binding, stacking interaction, neutron macromolecular crystallography (NMX), fully deuterated L-fucose, carbohydrate-binding, perdeuteration, carbohydrates, neutron structure, Photorhabdus

Introduction

Neutron macromolecular crystallography (NMX) provides unique information on the location of hydrogen atoms in proteins and solvent molecules that is rarely obtainable by X-ray crystallography even at high resolution (Blakeley et al., 2015; Eriksson et al., 2013; Woińska et al., 2016). Neutrons are scattered from atomic nuclei with a scattering power that is isotope-dependent. Since the neutron scattering lengths of hydrogen and deuterium are of similar magnitudes to the other elements commonly found in proteins such as carbon, oxygen and nitrogen, they can be easily visualized in neutron crystallographic analyses (Blakeley and Podjarny, 2018). NMX is thus a powerful technique for structural glycobiology where the location of hydrogen atoms is of crucial importance for understanding the interactions between carbohydrates and amino acids that may involve direct and water-bridged hydrogen bonds. To date (December 2020), there are only 175 macromolecular structures deposited in the Protein Data Bank that have been determined using neutron diffraction data. Glycoside hydrolases (Niimura et al., 1997; Wan et al., 2015), lytic polysaccharide monooxygenase (Bacik et al., 2017; O'Dell et al., 2017) and xylose isomerase (Langan et al., 2014) are carbohydrate-processing enzymes that have been extensively studied by neutron diffraction. On the other hand, only two lectins (sugar-binding proteins) have been studied using neutron crystallography, namely Concanavalin A (Blakeley et al., 2004; Gerlits et al., 2017; Habash et al., 1997, 2000; Kalb et al., 2000) and Galectin-3C (Manzoni et al., 2016, 2018).

Lectins are carbohydrate-binding proteins having no enzymatic activity. They are involved in various physiological and pathological processes including molecular and cellular recognition and adhesion (Lis and Sharon, 1998). Many pathogenic bacteria use lectins as toxins or adhesins to promote bacterial infection (Beddoe et al., 2010; Moonens and Remaut, 2017) and the detailed knowledge of their binding site has been used for designing anti-adhesive compounds, some of them now in preclinical study (Mousavifar et al., 2018).

Photorhabdus laumondii (formerly *P. luminescens*) is a bioluminescent rod-shaped gram-negative bacterium that lives in a mutualistic relationship with entomopathogenic nematodes from the genus *Heterorhabditis* (Machado et al., 2018; Waterfield et al., 2009). This species can also adopt alternative lifestyle in the soil, with positive influence on plant rhizosphere (Regaiolo et al., 2020). *P. laumondii* produces several lectins; three of them, PLL (Kumar et al., 2016), PLL2 (Fujdiarová et al., 2020), and PLL3 (Faltinek et al., 2019) that are similar in sequences, have been recently characterized and are proposed to play important roles in the complex *Photorhabdus* life cycle. All three lectins bind L-fucose and their monomeric structure folds as a seven-bladed β-propeller characteristic of this family of lectins. PLL is a homotetramer with seven putative binding sites for the monosaccharide per monomer, of which three were occupied by L-fucose in the X-ray crystal structure (Kumar et al., 2016). The PLL lectin from *P. laumondii* is an excellent model system for the study of lectin/carbohydrate interactions since the binding of fucose involves several hydrogen bonds to the protein side chain and main chain atoms, bridging by water molecules, and, very interestingly, hydrophobic contact with two tryptophan residues, that involve CH/π dispersion interactions important in protein-carbohydrate complexes (Wimmerová et al., 2012).

In order to optimise the study of the lectin-carbohydrate interactions using neutron

diffraction experiments, it is highly advantageous to perdeuterate (*i.e.* fully deuterate) both the protein and the ligands, so that all hydrogen atoms are replaced by deuterium. This maximises the visibility of the coherent neutron diffraction signal (Haertlein et al., 2016) by eliminating the inherently large incoherent scattering from hydrogen. In addition, the negative sign of the coherent neutron scattering length of hydrogen gives rise to cancellation effects in neutron scattering length density maps determined at medium resolutions (~ 1.6-2.5 Å) that makes it difficult to visualise hydrogen and carbon atoms in -CH_2 groups (Fisher et al., 2014; Koruza et al., 2019). The use of perdeuterated molecules obviates this problem and provides better quality neutron scattering length density maps.

Perdeuterated recombinant proteins are now commonly produced using bacterial and yeast expressions systems that have been adapted to D_2O-based growth medium and a deuterated carbon source (Heartlein et al, 2016; Meilleur et al., 2009). Perdeuterated glycans, on the other hand, are not easily available and only a limited number have been produced via synthetic or enzymatic approaches (Cress et al., 2019; Kent et al., 2015; Sawama et al., 2012). Very few neutron structures of protein complexed with perdeuterated carbohydrate have been characterised (Kovalevsky et al., 2008, 2010; Langan et al., 2014). Recently, however, we reported the production of perdeuterated L-fucose-d_{12} (Fuc-d_{12}) obtained using a glyco-engineered *E. coli* strain, genetically modified to produce high titers of L-fucose (Gajdos et al., 2020).

Here, we report the neutron and X-ray structures of the apo and fucose-bound PLL lectin from *P. laumondii*. This is, to our knowledge, the first account of a neutron structure of a perdeuterated lectin bound to a perdeuterated monosaccharide, allowing

for a precise description of the hydrogen bond network and the π-stacking interaction between the hydrophobic face of fucose and the aromatic rings of amino acids in the binding site.

Results
Neutron diffraction data from PLL that was buffer-exchanged in D_2O (H/D-PLL) in the apo form and from perdeuterated PLL (D-PLL) complexed with perdeuterated fucose (Fuc-d_{12}) were collected from single crystals of 0.5-0.7 mm^3 to 2.2 Å resolution (Figure S1). For the joint X-ray/neutron structural refinement, room temperature (RT) X-ray data were collected on the same crystals to resolutions of 1.70 Å (H/D-PLL apo) and 1.84 Å (D-PLL/Fuc-d_{12}) respectively (Table S1). In addition, three RT X-ray structures of D-PLL apo, H/D-PLL/H-fucose (commercial L-fucose) and D-PLL/H-fucose complexes as well as one 100 K X-ray structure of D-PLL/H-fucose complex were collected for comparisons (Table S2).

Overall structure of D-PLL and comparisons of RT and 100 K X-ray structures
The overall X-ray structures of H/D-exchanged PLL and D-PLL consist of a homotetramer of seven-bladed β-propellers linked by disulphide bridges (Figure 1). There is only one PLL monomer in the asymmetric unit. It is similar to the previously reported H-PLL (Kumar et al., 2016) and will not be described in details here. The superposition of the RT jointly refined X-ray and neutron structures of D-PLL and H/D-PLL gives an overall root-mean-square deviation (RMSD) of 0.17 Å (backbone atoms) indicating that the deuteration did not lead to any modifications of the 3D structure and that both structures are comparable. Similar results were obtained after superposition of the 100 K X-ray crystal structures of the H-HLL and D-PLL giving an RMSD of 0.13 Å (backbone atoms).
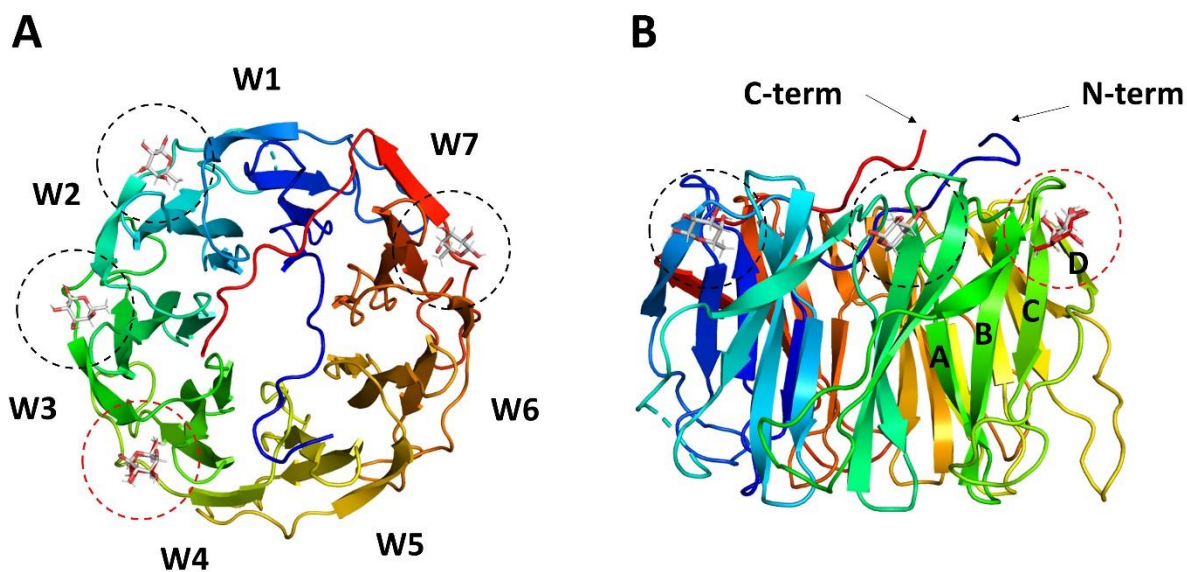
Figure 1: X-ray structure of D-PLL monomer with fucose binding sites. (A) Top view of the D-PLL monomer (PDB: 7BB4). Four molecules of fucose located between adjacent repeats (W1-W7) are shown in stick representation. Black-circled are sites that have been already described. The red-circled is an additional binding site described herein. (B) Side view of the monomer.

Interestingly, larger differences were observed when comparing the structures obtained at RT and 100 K, with mean RMSD of 0.64 Å (backbone atoms). After investigation, two clear differences were observed between the RT and 100 K structures of the protein, regardless of its H- or D-form. Differences were visibly pronounced in the short loops connecting β-strands C and D of the W1 and W5 repeats, respectively, caused by glycerol molecules that altered their conformations.

Specifically, in both short loops, a glycerol molecule is strongly bound to the amino acids via six hydrogen bonds (Figure 2) making the loops shift away from their normal positions observed in the room temperature structures. All 100 K structures of PLL lectin from the previous study (Kumar et al., 2016) and this work were collected from crystals soaked in glycerol solution, which served as a cryoprotectant during the crystal cryo-cooling. Glycerol readily binds to the protein with an average

number of 17 glycerol molecules per monomer for the apo protein and 11 for the fucose-bound complex. An unexpected difference between X-ray structures at RT and at 100 K is the number of occupied binding sites. In the crystal structure of the H-PLL/H-fucose complex (PDB: 5C9P), three fucose molecules were observed in three binding sites, site II (between W-motifs 1 and 2), III (between W-motifs 2 and 3) and VII (between W-motifs 6 and 7) respectively (Figure 1).

Here we report an additional binding site IV (between repeats W3 and W4) occupied by a fucose molecule in our RT and 100 K X-ray structure, in place of a glycerol molecule in the previously reported structure at 100 K (Figure 3). The fucose molecule could be modelled in a mixture of α/β-configurations with occupancies of 54/46. Generally, the fucose is bound in site IV in a similar manner to the other sites (Figure S2). The strongest hydrogen bonds are between the O4 oxygen of L-fucose and

115

the backbone ND of Ala190 and O of Thr168. Fucose is further stabilized via backbone hydrogen bonds of NH (Thr168)

to the sugar ring O or O4 oxygen atoms.
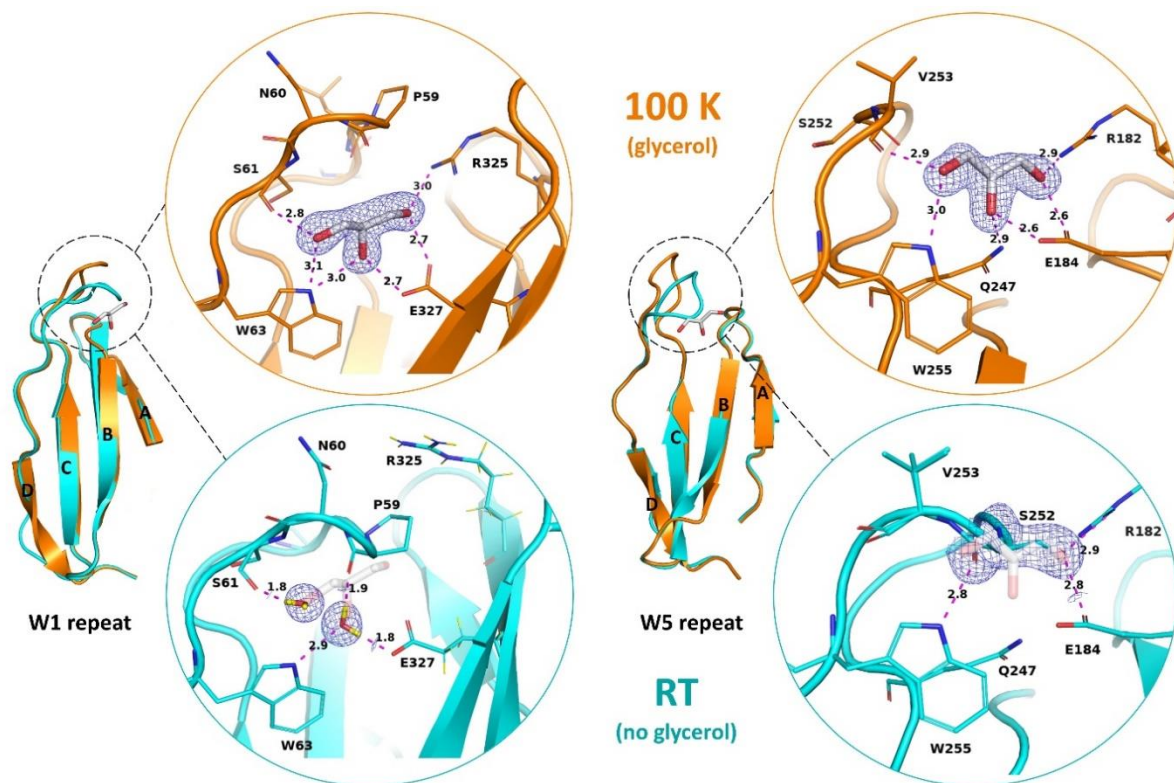
Figure 2: Superposition of 100 K (orange, PDB: 7BB4) and room temperature (cyan, PDB: 7BBC) X-ray structures of the repeats W1 and W5 of the D-PLL β-propeller. $2mF_o$-DF_c electron density is shown as blue mesh and contoured at 1.5 σ level. Hydrogen bonds are shown as magenta dashed lines and distances are given in Å. (Top) In the 100 K structure, a glycerol molecule (used as a cryoprotecting agent) binds strongly to the protein changing the native conformation of the short loop between strands C and D of the repeats W1 and W5 observed in the room temperature structure collected from crystals mounted in capillaries. (Bottom) In the RT structure, native conformation of loops is presented. Glycerol molecules (transparent grey sticks) from 100 K structure are superimposed with the RT structure.

Neutron structures and quality of the neutron density map

The use of perdeuterated protein yielded high quality neutron scattering length density maps (Figure 4A). Exchangeable protons could be located without ambiguity in the $2mF_o$-DF_c neutron map. The deuterium on the amine group of the main chain could be visualized as well as their involvement in hydrogen bonds (Figure 4B). However, the H/D-exchanged protein neutron structure exhibited clear

cancellation effects in the $2mF_o$-DF_c maps, mostly around aliphatic residues (Figure 4A).

The protonation states of the histidine residues could also be unambiguously determined from the perdeuterated structure analysis (D-PLL/Fuc-d_{12}). All histidines of the H/D-exchanged apo PLL were protonated on both nitrogen atoms ND1 and NE2 as a result of the used crystallization condition with a low pH of

4.6. In comparison, all but two of the histidine residues were equally protonated in the perdeuterated D-PLL/Fuc-d_{12} crystal (Table 1) grown at a physiological pH (0.1 M sodium potassium tartrate, pH ~ 7). Based on the examination of the $2mF_o$-DF_c omit neutron map, His155 and His176 were found to be in the neutral NE2-protonated and ND1-protonated tautomeric forms respectively (Figure 4C). It is noteworthy, that none of the histidine residues of the PLL lectin is involved or is in the vicinity of the sugar-binding sites. Further analysis of the neutron structures showed no protonation on any acidic side chains (Asp, Glu).
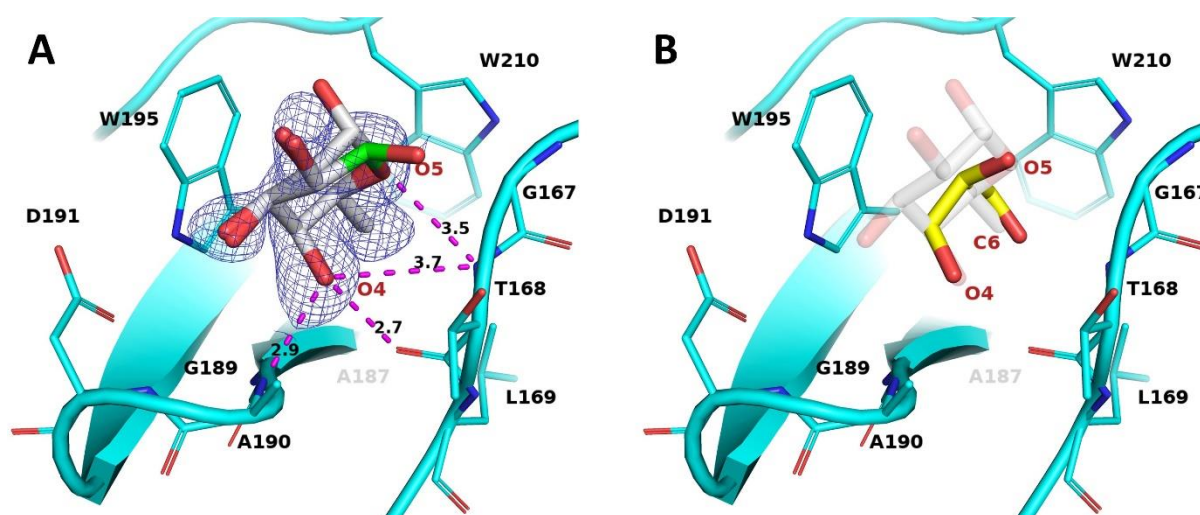
Figure 3. X-ray structure of fucose in the additional binding site IV of D-PLL located between repeats W3 and W4 of the seven-bladed β-propeller. (A) $2mF_o$-DF_c X-ray omit map (blue mesh) around L-fucose is contoured at 0.9σ. Position of fucose O1 atom is modelled in both alpha (grey) and beta (green) configurations. Hydrogen bonds are depicted as magenta dashed lines. All distances are given in Å. (B) Superposition of the fucose (transparent grey sticks) and a glycerol molecule (yellow sticks) modelled in apo PLL (PDB: 5C9O). The fucose atom labelling corresponds to the superimposed glycerol oxygen atoms.

Table 1: Protonation states of histidine residues observed in the neutron crystal structures.

State	H51	H80	H100	H148	H155	H176	H196	H203	H244	H298	H339
apo	Both	Both	Both	Both	Both	Both	Both	Both	Both	nd	Both
fucose	Both	Both	Both	Both	NE2	ND1	Both	Both	Both	nd	Both

The $2mF_o$-DF_c X-ray map for the D-PLL/Fuc-d_{12} complex at 1.85 Å resolution shows clear electron density for the ligand in binding sites II (Figure 5), III and VII. The fourth binding site described above for RT X-ray structure is not occupied by the ligand in the neutron structure of the complex. The site is occupied by solvent molecules (Figure S3B) observed also in the 100 K X-ray structure (PDB: 5C9P). Possible explanation is that this site has a lower affinity for fucose than the other sites as previously suggested (Kumar et al., 2016). In all sites, fucose could be modelled in both α/β configurations with occupancies of 30/70, which is in agreement with the

population of 30/70 observed by NMR in solution (Ryu et al., 2004). The neutron scattering density around the deuterated fucose molecule is clearly different from the electron density given that the deuterium atoms contribute significantly to the neutron scattering (Figure 5). As a result, neutron peaks of the methyl group together with deuteriums on carbon atoms are visible in the density map and directionality of the fucose hydroxyl groups could be determined. The map quality is better in the binding site II located between blades W1 and W2 of the β-propeller.
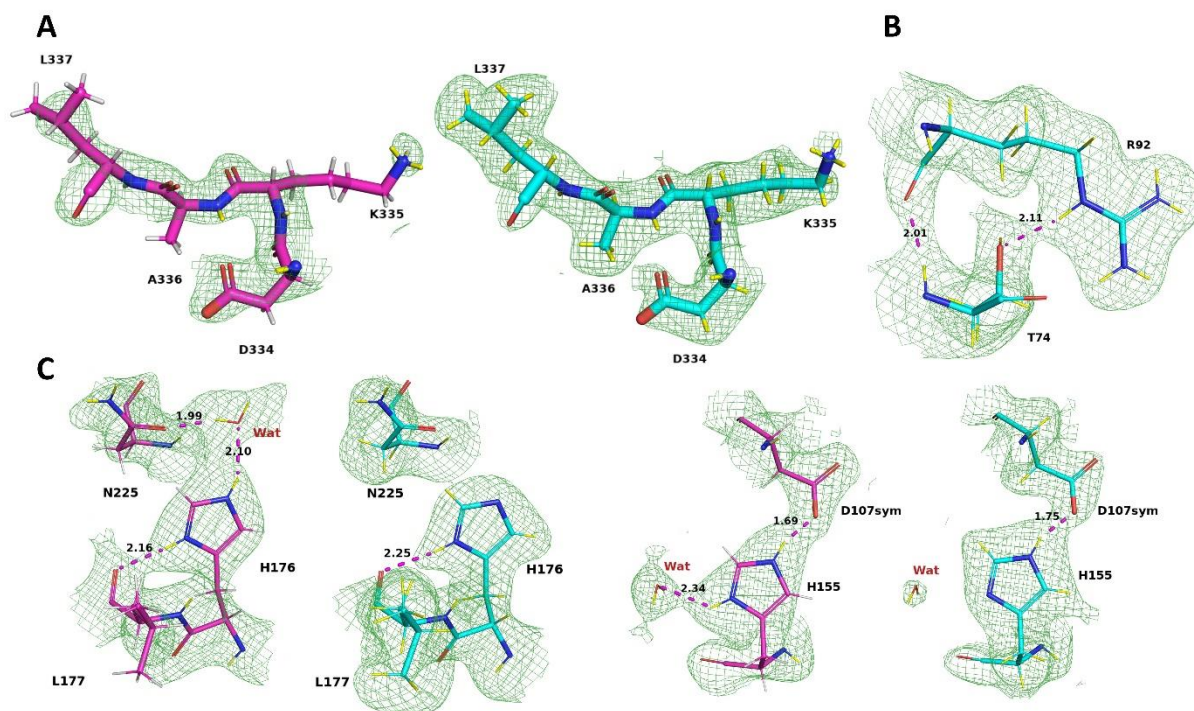
Figure 4: Quality of the neutron scattering length density map. (A) Stick representation and associated neutron density maps of selected residues of the H/D-PLL (magenta sticks) and D-PLL (cyan sticks). Hydrogens atoms are coloured white and deuterium atoms are coloured yellow. The $2mF_o$-DF_c neutron density map (green mesh) is contoured at 1σ. The use of perdeuterated protein provides better quality maps than those calculated for the H/D-exchanged protein where interpretation is limited by clear density cancellation effects around the CH-groups of the aliphatic amino acids. (B) Example of a backbone hydrogen bond with a corresponding neutron density. (C) Comparison of the protonation states of His176 and His155 for the H/D-exchanged apo PLL (magenta sticks) and D-PLL/Fuc-d$_{12}$ (cyan sticks) neutron structures. All distances are measured from a deuterium atom to the acceptor atom and are given in Å.
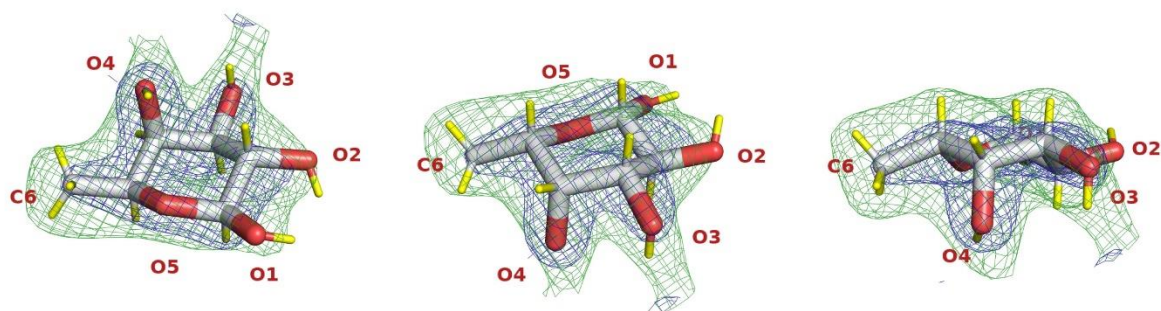
Figure 5: Electron and neutron density around perdeuterated fucose. Three different views of perdeuterated L-fucose (grey sticks) in the binding site II of the D-PLL/Fuc-d$_{12}$ complex (only β-configuration is displayed for the O1 atom). Deuterium atoms are coloured yellow. The 2mF_o-DF_c neutron scattering length density (green mesh) is contoured at 1σ and the 2mF_o-DF_c electron density (blue) is contoured at 2σ level.

Hydrogen bonding in the fucose binding sites

The sugar-binding mode is essentially identical in all binding sites and the orientation of the sugar is similar to the one previously observed in the X-ray structure (Kumar et al., 2016). Only novel features highlighted by the neutron maps are therefore described here. In site II, the two key fucose hydroxyl groups OD3 and OD4 are involved in direct hydrogen bonds with the protein (Figure S4). The neutron map displays perfect continuity of density between fucose O3 and the hydroxyl group of Thr94 (Figure 6). The neutron structure clarifies the direction of the hydrogen bond, with deuterium on Fuc-O3 pointing towards OG1 of Thr94 side chain (1.94 Å), while the deuterium on this oxygen goes to the solvent. The Fuc-O4 oxygen atom accepts an H-bond with the backbone amide of Thr94 (2.02 Å) clearly observed in the neutron density map while the deuterium of Fuc-O4 points towards the backbone oxygen of Val72 (2.06 Å). The same hydrogen bond network is observed in site III, with the addition of the ring oxygen O5 accepting one hydrogen bond from the backbone amide of Gly120. Fucose has the same orientation in the site VII but the map

is of lower quality. Measured distances and angles of hydrogen bonds are presented in Table 2.

The neutron structure also provides insight to the role of water molecules in the ligand binding. In the binding site II, one water molecule bridges between Fuc-O1 in β-configuration and the protein, with clear density around the two deuterium involved. Fuc-O1β atom accepts an H-bond (1.62 Å) from this water that in turn accepts another hydrogen bond from the backbone amide of Val72 (2.30 Å). Two additional waters seem to be bridging the interaction of Fuc-O3 atom and Asp95 although with weak neutron density suggesting higher mobility and weaker coordination.

119
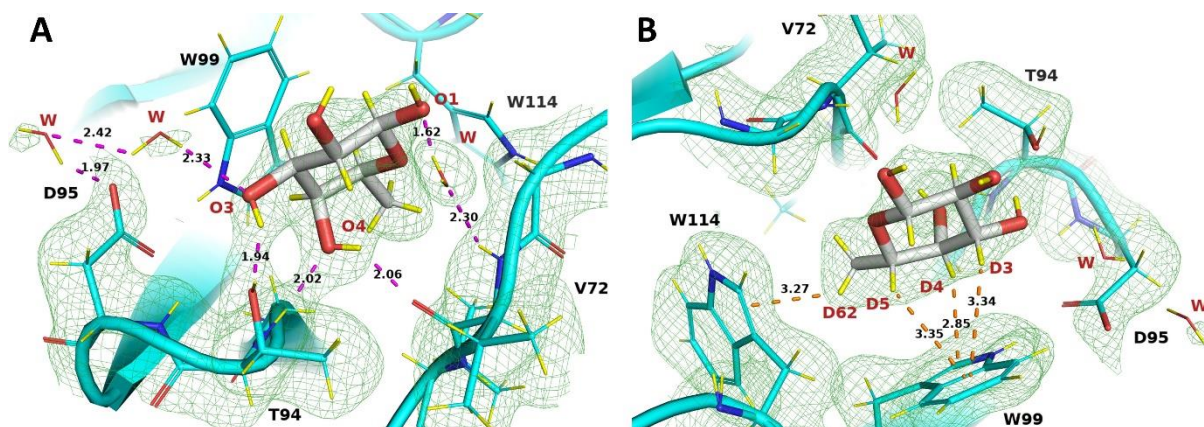
Figure 6: Two different views of binding site II in the complex of perdeuterated PLL lectin bound to the perdeuterated L-fucose. (A) Hydrogen-bonding network around the fucose-d$_{12}$ molecule. The 2mF_o-DF_c nuclear scattering length density (green mesh) is shown at a 1σ counter level. The L-fucose is shown as thick grey sticks and the protein is shown as thin cyan sticks. Deuterium atoms are shown as yellow sticks. Hydrogen bonds are shown as magenta dashed lines. (B) View rotated by 180 ° around the axis perpendicular to the panel to highlight the CD-π interactions between the nonpolar face of the fucose and the aromatic amino acids. CD-π dispersion interactions are shown as orange dashed lines with distances (Å) between C-D and benzene/pyrrole ring centroid of the relevant tryptophan residue.

Table 2: Measured distances and angles of hydrogen bonding interactions between perdeuterated L-fucose and PLL lectin based on the jointly refined X-ray/neutron crystal structure (D-PLL/Fuc-d$_{12}$).

Site	Donor	Acceptor	Distance (Å)	Angle (°)
II	Fuc-O4D	Val72-O	2.06	129.7
	Thr94-ND	Fuc-O4	2.02	158.3
	Fuc-O3D	Thr94-OG1	1.94	153.2
	W309-OD	Fuc-βO1	1.62	163.5
III	Fuc-O4D	Gly120-O	1.88	145.8
	Ser142-ND	Fuc-O4	2.05	166.8
	Fuc-O3D	Ser142-OG	2.13	129.1
	Gly120-ND	Fuc-O5	2.21	144.6
VII	Fuc-O4D	Ile311-O	1.95	165.7
	Leu333-ND	Fuc-O4	2.16	166.0

CD/π interactions observed in the neutron structure

All fucose binding sites of the PLL lectin are characterized by a conserved tryptophan-tryptophan cross-strand pair whose aromatic planes are directed at a near-right angle (varying from 87-115 °), creating a cleft to which the fucose is coordinated with interaction from both its hydrophobic face and its methyl group (Figure 6B). The neutron maps highlight the perfect complementarity of shape between the flat surface formed by aliphatic deuterium on C3, C4 and C5 of fucose and the aromatic ring of Trp99 (in site II) resulting in strong CD/π dispersion interactions. The methyl group of fucose is furthermore stabilized by interaction with the second tryptophan residue, Trp114 (in site II). The same interaction is observed in the two other binding sites (Figure S3) in the D-PLL/Fuc-d$_{12}$ structure.

Two selection criteria were chosen for defining CD-π dispersion interactions based on previous studies (Brandl et al., 2001). The first criterion was the closest distance between the fucose carbon atom and the ring centre of the π system of a tryptophan residue (d(Cn-π)) with the cut-off value of 4.5 Å. The second criterion was the angle at the deuterium (∠C-D-π) defined as the angle between the deuterium atom and the ring centre of the π system of a tryptophan residue with values being ≥ 120 ° (Figure S4).

Table 4: Experimentally determined geometrical parameters of CD-π interactions between perdeuterated L-fucose and aromatic amino acids in the binding sites of the perdeuterated PLL lectin based on the X-ray/neutron jointly refined crystal structure. Distances d(Cn-π) and d(Dn-π) are measured from the carbon and deuterium atom, respectively, to the geometrical centre of either the benzene or pyrrole aromatic rings of the tryptophan residue. Parameter ∠(Cn-Dn-π) is defined as an angle formed by carbon-deuterium-π centre of the aromatic ring.

Binding site	CDn	d(Cn-π)/Å		d(Dn-π)/Å		∠(Cn-Dn-π)/°	
		Benzene	Pyrrole	Benzene	Pyrrole	Benzene	Pyrrole
II Trp99 /Trp114*	3	4.09	4.42	3.34	4.07	135.3	104.5
	4	4.03	3.72	3.48	2.85	118.1	150.1
	5	4.10	4.00	3.35	3.57	136.3	109.8
	6	4.97	4.34	4.78	4.15	95.4	95.0
		4.93*	3.95*	4.15*	3.27*	139.0*	128.2*
III Trp147/Trp162*	3	3.70	4.25	2.95	3.93	134.6	102.4
	4	3.65	3.53	3.10	2.62	117.6	156.5
	5	3.83	3.97	3.15	3.69	128.4	99.5
	6	4.62	4.18	4.32	3.49	101.7	129.6
		3.85*	4.31*	2.97*	3.36*	151.3*	166.3*
VII Trp338/Trp354*	3	4.03	4.11	3.14	3.33	153.6	138.9
	4	4.46	4.01	4.27	3.37	95.3	125.2
	5	4.23	4.08	3.36	3.31	150.2	138.4
	6	5.32	4.81	5.13	4.58	96.0	98.0
		4.68*	3.71*	3.84*	3.00*	146.7*	130.8*

In order to better describe deuterium atoms involved in the CD-π interactions, the indole ring of tryptophan side chains was taken as two separate aromatic rings (benzene and pyrrole) with two geometrical centres. The measured interatomic distances and angles are listed in Table 4. Analysis of the measured distances and angles shows that up to four CD-π interactions occur in the fucose-binding by PLL lectin. Deuterium atoms on C3, C4, C5 and C6 are involved in CD-π stacking interaction in all binding sites with average distances of 3.14, 2.95, 3.29 and 3.15 Å respectively. In the binding site II, the strongest interaction is between D4 and the pyrrole ring of Trp99 (2.85 Å) whereas deuteriums on C3 and C5 interact weaker with the benzene part of Trp99 with distances of 3.34 and 3.36 Å respectively (Figure 6B). Furthermore, the deuterium on the fucose methyl group interacts with the pyrrole ring of the second tryptophan residue Trp114 with a distance of 3.27 Å. The binding interactions are the same in the other binding sites (Figure S3).

Water network in the apo (ligand-free) neutron structure

In the apo (ligand-free) neutron crystal structure extending to 2.2 Å resolution, the carbohydrate binding sites are occupied by three water molecules which are conserved in all three fucose-binding sites II, III and VII. The average crystallographic B-factors for the three water molecules are similar and rather high (50.7, 57.2, 53.5 Å^2 respectively) suggesting high mobility due to the limited occurrence of hydrogen bonds to the protein. The positions of water oxygen atoms were based on the $2mF_o\text{-}DF_c$ electron density map peaks. In the neutron map, the waters form a hydrogen bonding network creating a continuous nuclear density (Figure 7A). The most buried water molecule (Wat1) occupies the position of Fuc-O4. This is the most ordered water molecule establishing two hydrogen bonds with the protein. In the binding site II, Wat1

oxygen atom accepts an H-bond from the backbone amide of Thr94 (2.07 Å) while it donates one deuterium to the main chain oxygen of Val72 (2.05 Å). The second deuterium atom of Wat1 is coordinated towards the other water Wat2. The Wat1 water is coordinated in the same fashion in other binding sites (Figure S5).

The two other water molecules are more weakly coordinated and establish hydrogen bonds between them, rather than to the protein. Interestingly, Wat2 occupies positions between C5 and C6 (methyl group) of fucose, which is unconventional for sugar-binding proteins. Usually waters occupy positions of oxygen atoms of the sugar hydroxyl group mimicking the hydrogen bonding of the saccharide. PLL binding sites show hydrophobic character, suggesting the water molecules might be stabilized by nonconventional lone-pair···π or O-H···π interactions with the aromatic ring of the tryptophan (Durec et al., 2018).

Discussion

Protein-carbohydrate interactions are involved in many important physiological and pathological processes including cell trafficking, fertilization, immune response, cell-adhesion, or viral and bacterial infections. Lectins bind to carbohydrates mainly via hydrogen bonds, either directly or through water bridges, as well as via CH-π dispersion interactions. These latter can sometimes be the main driving force behind lectin-carbohydrate interactions (Asensio et al., 2013; Houser et al., 2020). The molecular mechanisms behind these interactions are predominantly studied at the atomic level by X-ray crystallography, which can locate individual hydrogens only at subatomic resolution (Takaba et al., 2019). The focus of this study was to describe lectin-carbohydrate interactions by neutron macromolecular crystallography as it provides unique and needed information on exact positions of

hydrogen atoms involved in the sugar recognition.

In this study, we report the first neutron structure, to our knowledge, of a bacterial lectin in its unbound and in a ligand-bound form. So far, two lectins have been studied by NMX, a glucose/mannose-specific lectin Concanavalin A from plants (Blakeley et al., 2004; Gerlits et al., 2017; Habash et al., 1997, 2000, Schirò et al., 2012) and a human galactose-specific Galectin-3C (Manzoni et al., 2016, 2018.
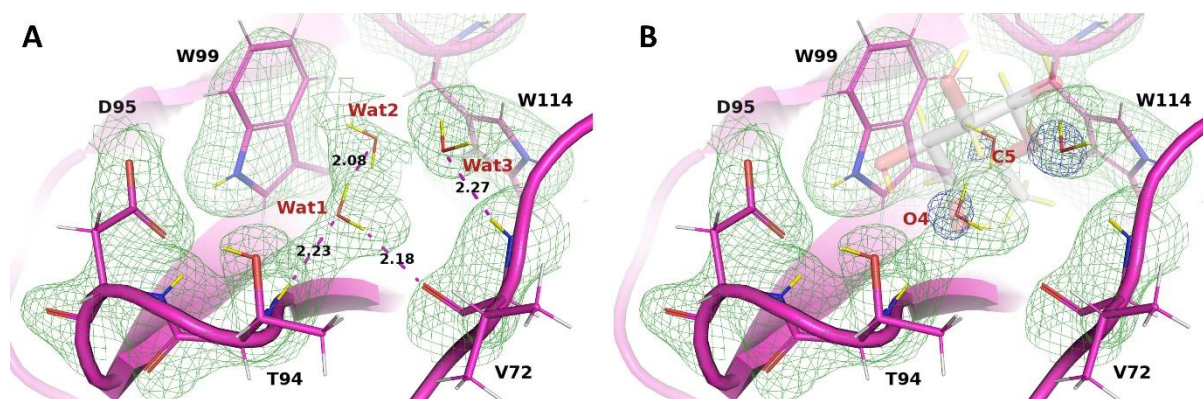
Figure 7: Water network in the binding site II of H/D-exchanged apo PLL. (A) Protein is depicted as magenta thin sticks. Hydrogen atoms are coloured grey and deuterium atoms are coloured yellow. Hydrogen bonds are shown as magenta dashed lines and distances are given in Å. The $2mF_o$-DF_c neutron density map is shown as green mesh and contoured at 1σ. (B) Superposition of the water molecules and fucose (transparent grey sticks) in the binding site. The $2mF_o$-DF_c electron density around water oxygen atoms is shown as blue mesh and contoured at 1σ. The fucose atoms at the positions of water molecules are labelled.

We compare RT jointly refined X-ray/neutron crystal structures of H/D-exchanged PLL lectin in its apo form and the perdeuterated D-PLL/Fuc-d$_{12}$ complex. No large differences were observed between H- and D-forms of PLL lectin concluding that perdeuteration did not affect the native state of the protein. Some differences were observed for the 100 K and RT X-ray structures of PLL lectin regardless of its H/D-state. We observed changes in conformation of short loops in repeats W1 and W5 that were caused by a bound glycerol used as a cryoprotectant in the 100 K structures. Moreover, a fucose molecule in our RT and 100 K structure of D-PLL lectin (PDB: 7B7C and 7BB4,

respectively) occupies an additional binding site IV between repeats W3 and W4 of the β-propeller. This site has not been described yet as it was occupied by water molecules (PDB: 5C9P) or a glycerol molecule (PDB: 5C9O) in 100 K X–ray crystal structures (Kumar et al., 2016). Even though most of the X-ray structures are solved from data collected at cryogenic conditions (100 K), room-temperature structures (293 K) can often provide more relevant information on the protein conformation state and ligand-binding since they are closer to physiological conditions and thus avoid artefacts caused during cryo-cooling of protein crystals as observed by others (Maeki et al., 2020).

The RT X-ray/neutron structure of the perdeuterated complex deciphered the hydrogen bond network involved in the fucose binding sites with clear directionalities of fucose hydroxyl groups observed in the neutron map. Fuc-OD4 is the key fucose hydroxyl group in the PLL sugar-recognition as it makes two strong and well-defined hydrogen bonds with the protein backbone atoms in all binding sites. Fucose is further stabilized by another hydrogen bond between Fuc-OD3 and a polar side chain in sites II and III, while in sites IV and VII amino acids in this position have non-polar character, which precludes formation of this hydrogen bond. This is consistent with the neutron map quality that is of lower quality in the binding site VII.

Moreover, deuterium atoms of the perdeuterated L-fucopyranose ring can be seen clearly in the neutron density. This is due to the contribution of the deuterium scatterers with a positive nuclear scattering length so that the positions of all deuteriums can be unambiguously determined. To date, only a couple of neutron structures used perdeuterated monosaccharides which were all chemically synthetized (Kovalevsky et al., 2008; Langan et al., 2014). In our study, we used a perdeuterated sugar produced via biosynthetic pathways in glyco-engineered *E. coli* bacteria that was recently reported (Gajdos et al., 2020). Using hydrogenated sugars can lead to cancellation effects in the nuclear maps, mainly around CH_2 groups, which might make interpretation of maps challenging. Here, no cancellation effects were observed and a strong neutron peak of the fucose methyl group could be clearly seen in the neutron scattering length density map. The fucose methine and methyl groups at C3-C5 and C6 respectively are involved in the CD/π interactions with the tryptophan side chains. The average distances between the deuterium atom and the ring centroid of the relevant tryptophan residue are in the range 2.62-3.49 Å which

is in agreement with the recently published analysis of CH-π interactions based on the 3D structures of protein-carbohydrate complexes deposited in the PDB (Houser et al., 2017, 2020). Experimentally-determined structures of CH-π dispersion interactions are necessary for a better understanding of the nature and physics behind the dispersion forces involved in the protein-carbohydrate interactions. These can be used in the quantum chemical calculations of interaction energies, which could eventually improve structure-based drug design of new inhibitors.

Interestingly, in the RT X-ray/neutron structure of the unbound PLL lectin, three water molecules were observed at the positions of fucose O4, C5 and C6 atoms. The strongest neutron peak was for the water located on the position of Fuc-OD4, which is the most ordered water molecule in all binding sites. The other two waters seem to be less coordinated with only one hydrogen bond to the protein. It is possible that these water molecules interact with tryptophan residues using their O-H bonds or lone-pairs of electrons as described by NMR, molecular dynamics simulations and quantum mechanics calculations (Durec et al., 2018; Špačková et al., 2018).

Conclusion
In this study, we reported two X-ray/neutron jointly refined structures of PLL lectin in apo form and in complex with perdeuterated L-fucose. The present results show that neutron macromolecular crystallography is a powerful technique for studying lectin-carbohydrate interactions since it can help resolve complete hydrogen network in the ligand-binding sites as well as directionality of water molecules involved in binding. We also described CD-π interactions that are crucial for the fucose specificity in the PLL lectin and that could be directly observed and their distances could be measured from the neutron structure. Finally, we also described the

additional fucose-binding site observed in our X-ray structures completing the number of PLL active sites to 4 per monomer and 16 per tetramer.

Author Contributions
A.I., M.H., J.M.D., M.P.B. and V.T.F. designed the experiment. M.W., A.K. provided expertise in the production and purification of the protein. L.G., J.M.D., M.H. and V.T.F. produced the deuterated biomolecules. L.G. purified the various components and carried out the experimental work. M.P.B. collected and processed the neutron data and provided expertise in structure determination. L.G. solved the crystal structures and refined them. L.G. prepared the figures and wrote the manuscript with A.I. and J.M.D., and with critical inputs from all authors.

Declaration of Interests
The authors declare no competing interests.

Accession numbers
The atomic coordinates and structure factors for the structures reported here have been deposited in the PDB with the accession numbers PDB: 7BBI, 7BBC, 7B7F, 7B7E, 7B7C, 7BB4.

References
Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 213–221.

Afonine, P. V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. *68*, 352–367.

Artero, J.B., Härtlein, M., McSweeney, S., and Timmins, P. (2005). A comparison of refined X-ray structures of hydrogenated and perdeuterated rat γE-crystallin in H2O and D2O. Acta Crystallogr. Sect. D Biol. Crystallogr. *61*, 1541–1549.

Arzt, S., Campbell, J.W., Harding, M.M., Hao, Q., and Helliwell, J.R. (1999). LSCALE - The new normalization, scaling and absorption correction program in the Daresbury Laue software suite. J. Appl. Crystallogr. *32*, 554–562.

Asensio, J.L., Ardá, A., Cañada, F.J., and Jiménez-Barbero, J. (2013). Carbohydrate-aromatic interactions. Acc. Chem. Res. *46*, 946–954.

Bacik, J.P., Mekasha, S., Forsberg, Z., Kovalevsky, A.Y., Vaaje-Kolstad, G., Eijsink, V.G.H., Nix, J.C., Coates, L., Cuneo, M.J., Unkefer, C.J., et al. (2017). Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. Biochemistry *56*, 2529–2532.

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 271–281.

Beddoe, T., Paton, A.W., Le Nours, J., Rossjohn, J., and Paton, J.C. (2010). Structure, biological functions and applications of the AB5 toxins. Trends Biochem. Sci. *35*, 411–418.

Blakeley, M.P., Kalb, A.J., Helliwell, J.R., and Myles, D.A.A. (2004). The 15-K neutron structure of saccharide-free concanavalin A. Proc. Natl. Acad. Sci.

U. S. A. *101*, 16405–16410.

Blakeley, M.P., Teixeira, S.C.M., Petit-Haertlein, I., Hazemann, I., Mitschler, A., Haertlein, M., Howard, E., and Podjarny, A.D. (2010). Neutron macromolecular crystallography with LADI-III. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 1198–1205.

Blakeley, M.P., Hasnain, S.S., and Antonyuk, S. V. (2015). Sub-atomic resolution X-ray crystallography and neutron crystallography: Promise, challenges and potential. IUCrJ *2*, 464–474.

Blakeley, M.P., and Podjarny, A.D. (2018). Neutron macromolecular crystallography. Emerg. Top. Life Sci. *2*, 39–55.

Brandl, M., Weiss, M.S., Jabs, A., Sühnel, J., and Hilgenfeld, R. (2001). C-H···π-interactions in proteins. J. Mol. Biol. *307*, 357–377.

Campbell, J.W., Hao, Q., Harding, M.M., Nguti, N.D., Wilkinson, C., and IUCr (1998). LAUEGEN version 6.0 and INTLDM. J. Appl. Crystallogr. *31*, 496–502.

Cress, B.F., Bhaskar, U., Vaidyanathan, D., Williams, A., Cai, C., Liu, X., Fu, L., M-Chari, V., Zhang, F., Mousa, S.A., et al. (2019). Heavy Heparin: A Stable Isotope-Enriched, Chemoenzymatically-Synthesized, Poly-Component Drug. Angew. Chemie Int. Ed. *58*, 5962–5966.

Dimasi, N., Flot, D., Dupeux, F., and Márquez, J.A. (2007). Expression, crystallization and X-ray data collection from microcrystals of the extracellular domain of the human inhibitory receptor expressed on myeloid cells IREM-1. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *63*, 204–208.

Durec, M., Marek, R., and Kozelka, J. (2018). Water–Tryptophan Interactions: Lone-pair···π or O−H···π? Molecular Dynamics Simulations of β-Galactosidase Suggest that Both Modes Can Co-exist. Chem. - A Eur. J. *24*, 5849–5859.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 486–501.

Eriksson, U.K., Fischer, G., Friemann, R., Enkavi, G., Tajkhorshid, E., and Neutze, R. (2013). Subangstrom resolution x-ray structure details aquaporin-water interactions. Science (80-.). *340*, 1346–1349.

Evans, P. (2006). Scaling and assessment of data quality. In Acta Crystallographica Section D: Biological Crystallography, (International Union of Crystallography), pp. 72–82.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. *69*, 1204–1214.

Faltinek, L., Fujdiarová, E., Melicher, F., Houser, J., Kašáková, M., Kondakov, N., Kononov, L., Parkan, K., Vidal, S., and Wimmerová, M. (2019). Lectin PLL3, a novel monomeric member of the seven-bladed β-propeller lectin family. Molecules *24*.

Fisher, S.J., Blakeley, M.P., Howard, E.I., Petit-Haertlein, I., Haertlein, M., Mitschler, A., Cousido-Siah, A., Salvay, A.G., Popov, A., Muller-Dieckmann, C., et al. (2014). Perdeuteration: Improved visualization of solvent structure in neutron macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. *70*, 3266–3272.

Fujdiarová, E., Houser, J., Dobeš, P., Paulíková, G., Kondakov, N., Kononov, L., Hyršl, P., and Wimmerová, M. (2020). Heptabladed β-propeller lectins PLL2 and PHL from *Photorhabdus* spp. recognize *O*-methylated sugars and influence the host immune system. FEBS J. febs.15457.

Gajdos, L., Forsyth, V.T., Blakeley, M.P., Haertlein, M., Imberty, A., Samain, E., and Devos, J.M. (2020). Production of perdeuterated fucose from glyco-engineered bacteria. Glycobiology. *31*, 151-158.

Gerlits, O.O., Coates, L., Woods, R.J., and Kovalevsky, A. (2017). Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A As Revealed by Neutron Crystallography. Biochemistry *56*, 4747–4750.

Habash, J., Raftery, J., Weisgerber, S., Cassetta, A., Lehmann, M.S., Høghøj, P., Wilkinson, C., Campbell, J.W., and Helliwell, J.R. (1997). Neutron Laue diffraction study of concanavalin A- The proton of Asp28. J. Chem. Soc. - Faraday Trans. *93*, 4313–4317.

Habash, J., Raftery, J., Nuttall, R., Price, H.J., Wilkinson, C., Kalb, A.J., and Helliwell, J.R. (2000). Direct determination of the positions of the deuterium atoms of the bound water in concanavalin A by neutron Laue crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. *56*, 541–550.

Haertlein, M., Moulin, M., Devos, J.M., Laux, V., Dunne, O., and Forsyth, V.T. (2016). Biomolecular Deuteration for Neutron Structural Biology and Dynamics. Methods Enzymol. *566*, 113–157.

Houser, J., Kozmon, S., Mishra, D., Mishra, S.K., Romano, P.R., Wimmerová, M., and Koca, J. (2017). Influence of trp flipping on carbohydrate binding in lectins. An example on Aleuria aurantia lectin AAL. PLoS One *12*.

Houser, J., Kozmon, S., Mishra, D., Hammerová, Z., Wimmerová, M., and Koča, J. (2020). The CH–π Interaction in Protein–Carbohydrate Binding: Bioinformatics and In Vitro Quantification. Chem. – A

Eur. J. *26*, 10769–10780.

Kabsch, W. (2010). XDS. Acta Crystallogr. D. Biol. Crystallogr. *66*, 125–132.

Kalb, A.J., Habash, J., Hunter, N.S., Price, H.J., Raftery, J., and Helliwell, J.R. (2000). Manganese(II) in concanavalin A and other lectin proteins. Met. Ions Biol. Syst. *37*, 279–304.

Kent, B., Hauß, T., Demé, B., Cristiglio, V., Darwish, T., Hunt, T., Bryant, G., and Garvey, C.J. (2015). Direct Comparison of Disaccharide Interaction with Lipid Membranes at Reduced Hydrations. Langmuir *31*, 9134–9141.

Koruza, K., Mahon, B.P., Blakeley, M.P., Ostermann, A., Schrader, T.E., McKenna, R., Knecht, W., and Fisher, S.Z. (2019). Using neutron crystallography to elucidate the basis of selective inhibition of carbonic anhydrase by saccharin and a derivative. J. Struct. Biol. *205*, 147–154.

Kovalevsky, A.Y., Katz, A.K., Carrell, H.L., Hanson, L., Mustyakimov, M., Zoe Fisher, S., Coates, L., Schoenborn, B.P., Bunick, G.J., Glusker, J.P., et al. (2008). Hydrogen location in stages of an enzyme-catalyzed reaction: Time-of-flight neutron structure of D-xylose isomerase with bound D-xylulose. Biochemistry *47*, 7595–7597.

Kovalevsky, A.Y., Hanson, L., Fisher, S.Z., Mustyakimov, M., Mason, S.A., Trevor Forsyth, V., Blakeley, M.P., Keen, D.A., Wagner, T., Carrell, H.L., et al. (2010). Metal Ion Roles and the Movement of Hydrogen during Reaction Catalyzed by D-Xylose Isomerase: A Joint X-Ray and Neutron Diffraction Study. Structure *18*, 688–699.

Kumar, A., Sýkorová, P., Demo, G., Dobeš, P., Hyršl, P., and Wimmerová, M. (2016). A Novel Fucose-binding Lectin from Photorhabdus luminescens (PLL) with an Unusual Heptabladed β-Propeller Tetrameric Structure. J. Biol. Chem. *291*, 25032–25049.

Langan, P., Sangha, A.K., Wymore, T., Parks, J.M., Yang, Z.K., Hanson, B.L., Fisher, Z., Mason, S.A., Blakeley, M.P., Forsyth, V.T., et al. (2014). L-Arabinose binding, isomerization, and epimerization by D-Xylose isomerase: X-Ray/Neutron crystallographic and molecular simulation study. Structure *22*, 1287–1300.

Lis, H., and Sharon, N. (1998). Lectins: Carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. *98*, 637–674.

Machado, R.A.R., Wüthrich, D., Kuhnert, P., Arce, C.C.M., Thönen, L., Ruiz, C., Zhang, X., Robert, C.A.M., Karimi, J., Kamali, S., et al. (2018). Whole-genome-based revisit of photorhabdus phylogeny:Proposal for the elevation of most

photorhabdus subspecies to the species level and description of one novel species photorhabdus bodei sp. nov., and one novel subspecies photorhabdus laumondii subsp. clarkei subsp. nov. Int. J. Syst. Evol. Microbiol. *68*, 2664–2681.

Maeki, M., Ito, S., Takeda, R., Ueno, G., Ishida, A., Tani, H., Yamamoto, M., and Tokeshi, M. (2020). Room-temperature crystallography using a microfluidic protein crystal array device and its application to protein-ligand complex structure analysis. Chem. Sci. *11*, 9072–9087.

Manzoni, F., Saraboji, K., Sprenger, J., Kumar, R., Noresson, A.L., Nilsson, U.J., Leffler, H., Fisher, S.Z., Schrader, T.E., Ostermann, A., et al. (2016). Perdeuteration, crystallization, data collection and comparison of five neutron diffraction data sets of complexes of human galectin-3C. Acta Crystallogr. Sect. D Struct. Biol. *72*, 1194–1202.

Manzoni, F., Wallerstein, J., Schrader, T.E., Ostermann, A., Coates, L., Akke, M., Blakeley, M.P., Oksanen, E., and Logan, D.T. (2018). Elucidation of Hydrogen Bonding Patterns in Ligand-Free, Lactose- and Glycerol-Bound Galectin-3C by Neutron Crystallography to Guide Drug Design. J. Med. Chem. *61*, 4412–4420.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Meilleur, F., Weiss, K.L., and Myles, D.A.A. (2009). Deuterium labeling for neutron structure-function-dynamics analysis. Methods Mol. Biol. *544*, 281–292.

Moonens, K., and Remaut, H. (2017). Evolution and structural dynamics of bacterial glycan binding adhesins. Curr. Opin. Struct. Biol. *44*, 48–58.

Moriarty, N.W., Grosse-Kunstleve, R.W., and Adams, P.D. (2009). Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. Acta Crystallogr. Sect. D Biol. Crystallogr. *65*, 1074–1080.

Mousavifar, L., Touaibia, M., and Roy, R. (2018). Development of Mannopyranoside Therapeutics against Adherent-Invasive Escherichia coli Infections. Acc. Chem. Res. *51*, 2937–2948.

Niimura, N., Minezaki, Y., Nonaka, T., Castagna, J.C., Cipriani, F., Høghøj, P., Lehmann, M.S., and Wilkinson, C. (1997). Neutron Laue diffractometry with an imaging plate provides an effective data collection regime for neutron protein crystallography. Nat. Struct. Biol. *4*, 909–914.

O'Dell, W.B., Agarwal, P.K., and Meilleur, F. (2017). Oxygen Activation at the Active Site of a Fungal Lytic

Polysaccharide Monooxygenase. Angew. Chemie - Int. Ed. *56*, 767–770.

Regaiolo, A. Dominelli, N., Andersen, K and Heermann, R. (2020). The biocontrol agent and insect pathogen *Photorhabdus luminescens* interacts with plant roots. Appl. Env. Microb. *86*, e00891-20.

Ryu, K.S., Kim, C., Park, C., and Choi, B.S. (2004). NMR analysis of enzyme-catalyzed and free-equilibrium mutarotation kinetics of monosaccharides. J. Am. Chem. Soc. *126*, 9180–9181.

Sawama, Y., Yabe, Y., Iwata, H., Fujiwara, Y., Monguchi, Y., and Sajiki, H. (2012). Stereo- and Regioselective Direct Multi-Deuterium-Labeling Methods for Sugars. Chem. - A Eur. J. *18*, 16436–16442.

Schirò, G., Vetri, V., Frick, B., Militello, V., Leone, M., and Cupane, A. (2012). Neutron scattering reveals enhanced protein dynamics in concanavalin a amyloid fibrils. J. Phys. Chem. Lett. *3*, 992–996.

Špačková, N., Trošanová, Z., Šebesta, F., Jansen, S., Burda, J. V., Srb, P., Zachrdla, M., Žídek, L., and Kozelka, J. (2018). Protein environment affects the water-tryptophan binding mode. MD, QM/MM, and NMR studies of engrailed homeodomain mutants. Phys. Chem. Chem. Phys. *20*, 12664–12677.

Takaba, K., Tai, Y., Eki, H., Dao, H.-A., Hanazono, Y., Hasegawa, K., Miki, K., and Takeda, K. (2019). Subatomic resolution X-ray structures of green fluorescent protein. IUCrJ *6*, 387–400.

Thaller, C., Weaver, L.H., Eichele, G., Wilson, E., Karlsson, R., and Jansonius, J.N. (1981). Repeated seeding technique for growing large single crystals of proteins. J. Mol. Biol. *147*, 465–469.

Wan, Q., Parks, J.M., Hanson, B.L., Fisher, S.Z., Ostermann, A., Schrader, T.E., Graham, D.E., Coates, L., Langan, P., and Kovalevsky, A. (2015). Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography. Proc. Natl. Acad. Sci. U. S. A. *112*, 12384–12389.

Waterfield, N.R., Ciche, T., and Clarke, D. (2009). Photorhabdus and a Host of Hosts . Annu. Rev. Microbiol. *63*, 557–574.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant,

M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. *27*, 293–315.

Wimmerová, M., Kozmon, S., Nečasová, I., Mishra, S.K., Komárek, J., and Koča, J. (2012). Stacking Interactions between Carbohydrate and Protein Quantified by Combination of Theoretical and Experimental Methods. PLoS One *7*, e46032.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. *67*, 235–242.

Woińska, M., Grabowsky, S., Dominiak, P.M., Woźniak, K., and Jayatilaka, D. (2016). Hydrogen atoms can be located accurately and precisely by x-ray crystallography. Sci. Adv. *2*, e1600192.
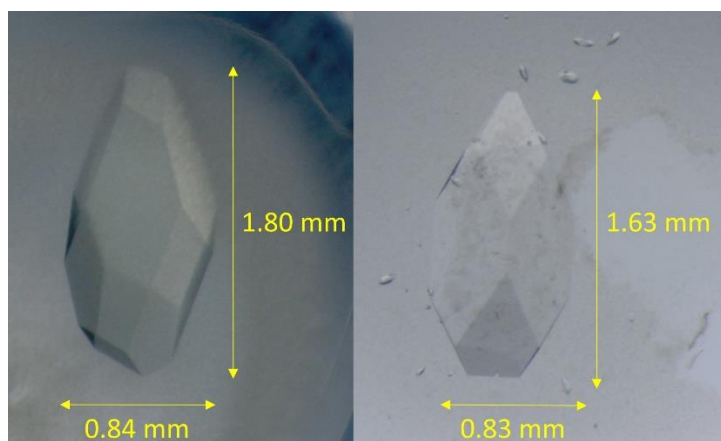
Figure S1: Crystals of the H/D-exchanged apo PLL lectin (H/D-PLL apo, left) and the perdeuterated PLL lectin soaked with perdeuterated L-fucose (D-PLL/Fuc-d$_{12}$, right) used for the neutron data collection at LADI-III. Both crystals diffracted to 2.2 Å resolution.
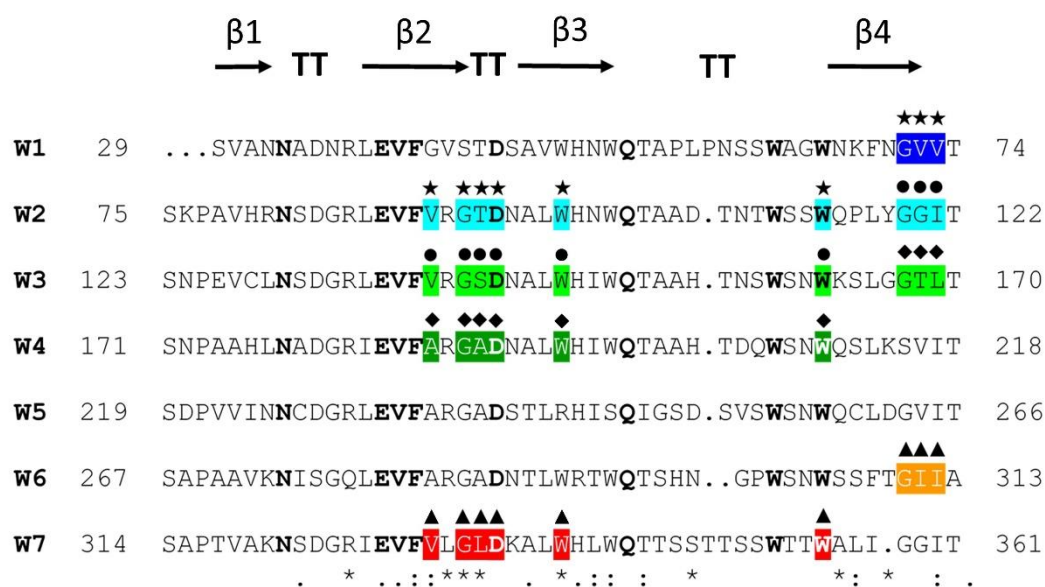
Figure S2, related to Figure 1: Alignment of the seven repeats (W-motifs) of the PLL β-propeller. Symbols indicate amino acids involved in the fucose binding sites II (stars), III (circles), IV (diamonds) and VII (triangles). The colours follow the colour code in Figure 1. In bold are the conserved amino acids.
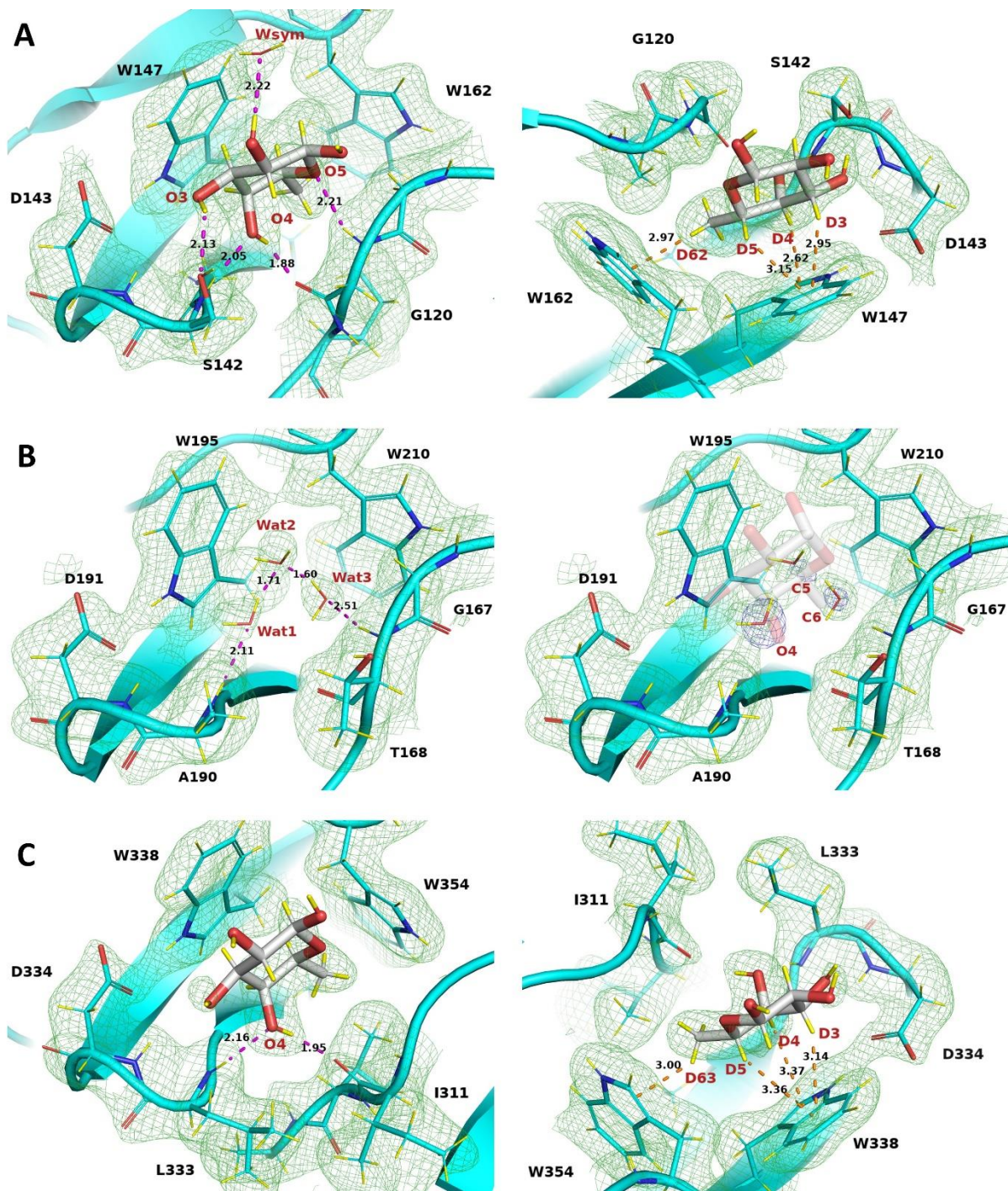
Figure S3, related to Figure 6: Neutron structure of the complex of perdeuterated PLL lectin bound to the perdeuterated L-fucose (D-PLL/Fuc-d12). (A) Binding site III. (B) Binding site IV. (C) Binding site VII. The $2m F_o$-$D F_c$ nuclear scattering length density (green mesh) is shown at a 1σ contour level. The L-fucose is shown as thick grey sticks and the protein is shown as thin cyan sticks. Deuterium atoms are shown as yellow sticks. Hydrogen bonds relevant for the discussion are shown as magenta dashed lines with distances (Å) between the

deuterium atom and H-bond acceptor atom. CD-π dispersion interactions are shown as orange dashed lines with distances (Å) between C-D and benzene/pyrrole ring centroid of the relevant tryptophan residue. (A) fucose-binding site III of PLL with hydrogen bond network (left) and a view rotated by 180° around the axis perpendicular to the panel to highlight the CD-π interactions between the nonpolar face of the fucose and the aromatic amino acids (right). (B) fucose-binding site IV occupied by solvent in the D-PLL/Fuc-d_{12} complex (left) and superposition of water molecules and fucose (transparent sticks) from the 100 K X-ray crystal structure (PDB: 7BB4) (right). (C) fucose-binding site VII (same views as panels A and B).
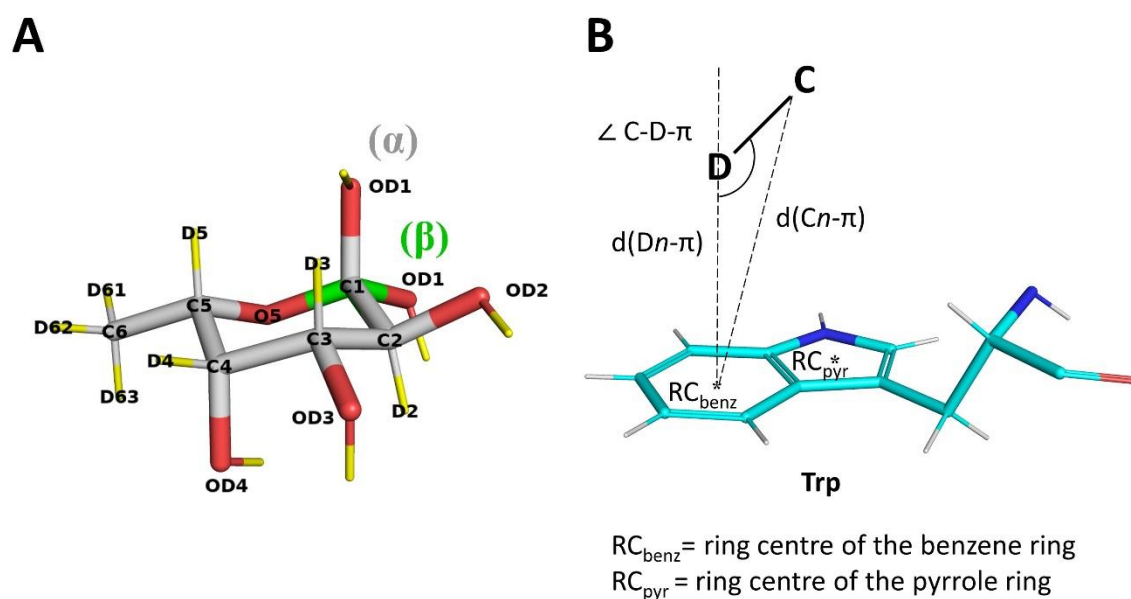
Figure S4, related to Figure 6 and Table 3: (A) Stick representation of L-fucose with atom labelling. Fucose is shown in both alpha (grey) and beta (green) configurations for O1 atom. (B) Schematic illustration of parameters defining CD-π interactions.
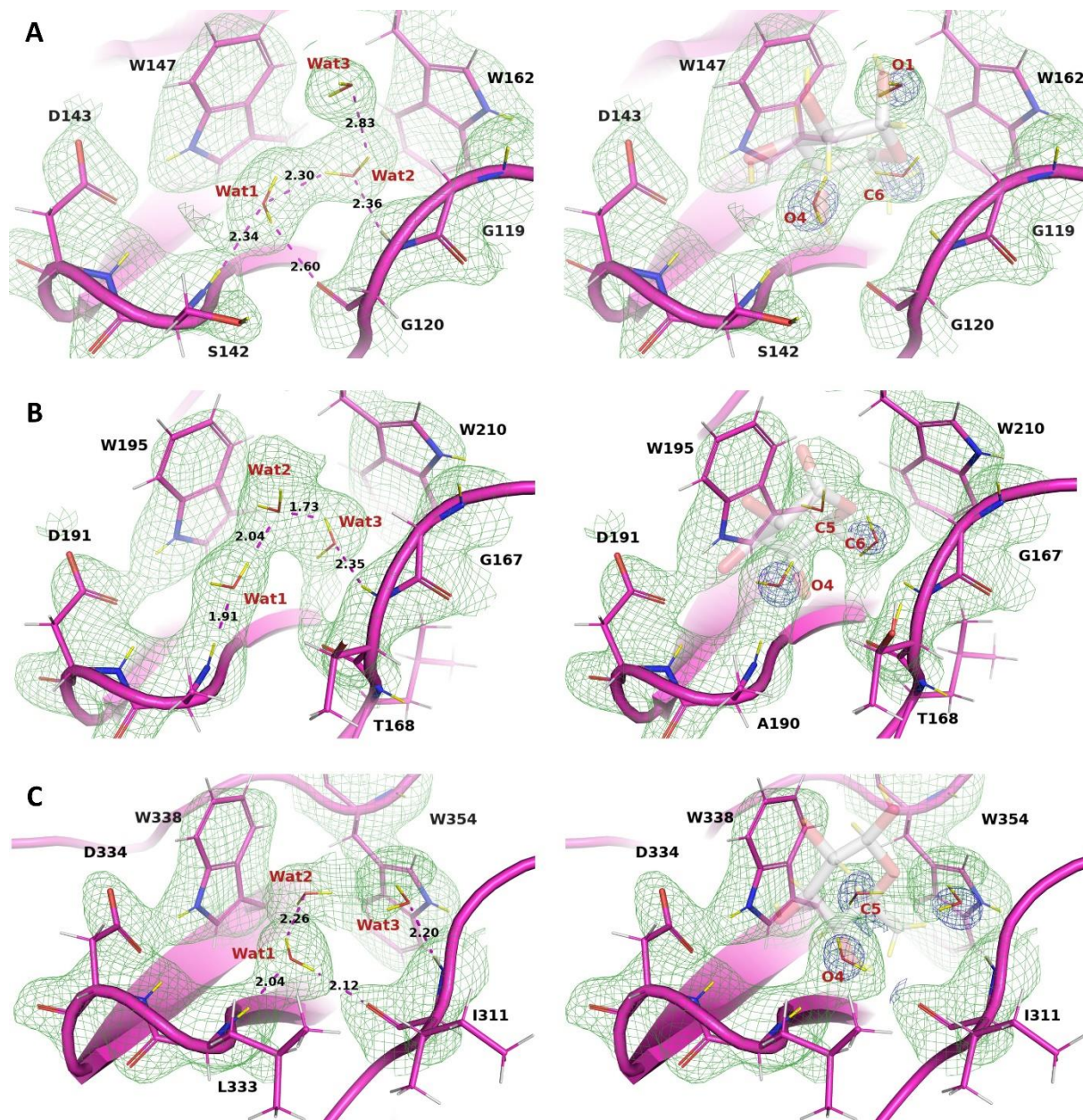
Figure S5, related to Figure 7: (Left) Water network in the binding sites III (A), IV (B) and VII (C) of the H/D-exchanged apo PLL lectin. Protein is depicted as magenta thin sticks. Hydrogen atoms are coloured grey and deuterium atoms are coloured yellow. Hydrogen bonds are shown as magenta dashed lines and distances are given in Å. The $2mF_o\text{-}DF_c$ neutron density map is shown as green mesh and contoured at 1σ. (Right) Superposition of the water molecules and fucose (transparent grey sticks) in the binding sites. The $2mF_o\text{-}DF_c$ electron density around water oxygen atoms is shown as blue mesh and contoured at 1σ. The fucose atoms at the positions of water molecules are labelled.

132

Table S1: Room temperature X-ray and neutron data-collection and structure refinement statistics for the H/D-exchanged apo PLL and D-PLL/Fucose-d₁₂ complex.

Values in parentheses are for the highest resolution shell.

	H/D-PLL Apo	D-PLL Fucose-d$_{12}$
Data collection		
Neutrons		
Instrument	LADI-III	LADI-III
Wavelengths (Å)	3.1-4.1	3.07-4.05
Detector	Image plate	Image plate
Resolution (Å)	39- 2.20 (2.32-2.20)	46-2.2 (2.32-2.20)
Spacegroup	I222	I222
Unit cell parameters		
a, b, c (Å)	72.7, 89.3, 159.4	72.7, 89.2, 159.2
α, β, γ (°)	90, 90, 90	90, 90, 90
R$_{merge}$ (*I*) (%)	18.1 (45.8)	16.3 (32.6)
R$_{pim}$ (*I*) (%)	13.2 (26.1)	8.3 (16.0)
Mean *I*/σ (*I*)	6.2 (2.0)	8.0 (3.4)
Completeness (%)	82.9 (73.5)	84.9 (74.7)
Multiplicity	2.8 (3.0)	4.1 (4.0)
No. of unique reflections	21 855 (2 797)	22348 (2822)
Crystal size (mm^3)	0.7	0.5
X-rays		
X-ray source	FIP-BM30A, ESRF	GeniX 3D Cu High Flux (Xenocs), IBS
Wavelength (Å)	0.9796	1.5418
Detector	ADSC Q315r	Mar 345 (marXperts)
Resolution (Å)	46-1.70 (2.02-1.70)	33-1.84 (1.88-1.84)
Unit cell parameters		
a, b, c (Å)	72.7, 89.3, 159.4	72.7, 89.2, 159.2
α, β, γ (°)	90, 90, 90	90, 90, 90
R$_{merge}$ (*I*) (%)	6.4 (77.5)	11.3 (58.2)
R$_{meas}$ (*I*) (%)	7.5 (90.8)	11.8 (64.0)
CC1/2 (%)	99.8 (63.3)	99.8 (82.7)
Mean *I*/σ (*I*)	12.4 (1.65)	14.7 (3.1)
Completeness (%)	98.6 (99.7)	100.0 (100.0)
Multiplicity	3.5 (3.5)	11.5 (5.8)
No. of unique reflections	56577 (4157)	45268 (2766)
Refinement		
Resolution range X-ray (Å)	39.84-1.70	33.05-1.84
Resolution range neutron (Å)	38.97-2.20	45.99-2.20
Reflections (used), X-ray	56568	45271
Reflections (test), X-ray	2828	2265
Reflections (used), neutron	21821	22234
Reflections (test), neutron	1092	1113
R$_{work}$ (%), X-ray	13.5	12.1
R$_{free}$ (%), X-ray	15.6	14.0
R$_{work}$ (%), neutron	21.5	19.1
R$_{free}$ (%), neutron	24.0	22.1
No. of atoms (protein)	2850	2939
No. of water molecules	340	334
Rmsd in bond lengths (Å)	0.010	0.010

Rmsd in bond angles (°)	1.312	1.323
Average B factors (Å^2)		
Protein	29.7	34.6
Ligand	-	51.9
Ramachandran statistics		
Favoured (%)	98.3	97.7
Allowed (%)	1.7	2.3
Outliers (%)	0.0	0.0
All-atom clashscore	2.6	2.7
PDB code	7BBI	7BBC

Table S2: X-ray data-collection and refinement statistics for H/D-exchanged and perdeuterated D-PLL lectin.

Values in parentheses are for the highest resolution shell.

	H/D-PLL H-fucose	D-PLL Apo	D-PLL H-fucose	D-PLL H-fucose
Data collection				
Temperature	RT	RT	RT	100 K
X-ray source	FIP-BM30A, ESRF	FIP-BM30A, ESRF	FIP-BM30A, ESRF	ID23-1, ESRF
Wavelength (Å)	0.9796	0.9796	0.9796	0.9717
Detector	ADSC Q315r	ADSC Q315r	ADSC Q315r	PILATUS 6M (Dectris)
Resolution (Å)	46-1.60 (1.64-1.60)	46-1.60 (1.64-1.60)	46-1.55 (1.59-1.55)	56-1.70 (1.73-1.70)
Spacegroup	I222	I222	I222	I222
Unit cell parameters				
a, b, c (Å)	72.6, 89.2, 159.3	72.6, 89.3, 159.2	72.6, 89.1, 159.2	72.8, 88.3, 158.3
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
R_{merge} (*I*) (%)	5.6 (91.3)	4.0 (108.1)	4.8 (71.8)	10.3 (36.4)
R_{meas} (*I*) (%)	6.2 (102.1)	4.4 (119.4)	5.3 (78.5)	11.9 (42.0)
CC1/2 (%)	99.8 (69.1)	100 (67.8)	99.9 (86.0)	98.9 (90.4)
Mean I/σ (*I*)	16.2 (2.1)	25.3 (1.8)	20.6 (3.2)	7.9 (2.9)
Completeness (%)	98.7 (98.9)	99.9 (100.0)	99.5 (100.0)	98.0 (99.3)
Multiplicity	5.2 (5.1)	5.6 (5.6)	5.9 (6.0)	4.2 (4.3)
No. of unique reflection	67576 (4953)	68418 (5004)	74700 (5479)	55209 (2949)
Model refinement				
Resolution (Å)	39.82-1.60	30.91-1.60	39.80-1.55	45.82-1.70
Reflections (used)	67565	68413	74690	55116
Reflections (test)	3378	3419	3734	2785
R_{work} (%)	14.9	15.0	14.6	17.2
R_{free} (%)	17.2	17.1	16.3	19.9
No. of atoms (protein)	2992	2987	3036	2897
No. of water molecules	406	396	410	593
Rmsd in bond lengths (Å)	0.005	0.005	0.005	0.006
Rmsd in bond angles (°)	0.819	0.840	0.820	0.896
Average B factors (Å^2)				
Protein	19.8	22.4	19.8	16.3
Ligand	34.8	-	33.3	25.9
Ramachandran statistics				
Favoured (%)	98.0	97.7	98.0	98.0
Allowed (%)	2.0	2.3	2.0	2.0
Outliers (%)	0.0	0.0	0.0	0.0
All-atom clashscore	2.6	1.7	2.5	6.0
PDB code	7B7F	7B7E	7B7C	7BB4

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli BL21(DE3)	Invitrogen / Life technologies	Cat#C600003
Chemicals, Peptides, and Recombinant Proteins		
L-fucose	Carbosynth	Cat#MF06710 CAS: 2438-80-4
Perdeuterated L-Fucose	(Gajdos et al., 2020)	N/A
Recombinant PLL	(Kumar et al., 2016)	N/A
Deuterated recombinant PLL	This paper	N/A
Deposited Data		
X-ray/neutron RT H/D-exchanged apo PLL	This paper	PDB: 7BBI
X-ray/neutron RT D-PLL/Fuc-d_{12}	This paper	PDB: 7BBC
X-ray RT D-PLL apo	This paper	PDB: 7B7E
X-ray RT H/D-exchanged PLL/H-fucose	This paper	PDB: 7B7F
X-ray RT D-PLL/H-fucose	This paper	PDB: 7B7C
X-ray 100 K D-PLL/H-fucose	This paper	PDB: 7BB4
Coordinates of rPLL	(Kumar et al., 2016)	PDB: 5C9O
Coordinates of rPLL in the presence of L-fucose	(Kumar et al., 2016)	PDB: 5C9P
Recombinant DNA		
Plasmid pET29a	Novagen	Cat#69871
Plasmid pET29a-pll	(Kumar et al., 2016)	N/A
Software and Algorithms		
XDS	(Kabsch, 2020)	http://www.xds.mpimf-heidelberg.mpg.de/
Phenix 1.16.3	(Adams et al., 2010)	https://www.phenix-online.org/
CCP4	(Winn et al., 2011)	http://www.ccp4.ac.uk
Coot	(Emsley et al., 2010)	https://bernhardcl.github.io/coot
Molprobity	(Williams et al., 2018)	http://molprobity.biochem.duke.edu/
Pymol 2.3.2	Schrodinger, LLC	https://pymol.org/2/
iMosflm	(Battye et al., 2011)	http://www.ccp4.ac.uk

LAUEGEN	(Campbell et al., 1998)	https://web.archive.org/web/20001024010254/http://www.dl.ac.uk/SRS/PX/jwc_laue/laue_top.html
LSCALE	(Arzt et al., 1999)	https://web.archive.org/web/20001024010254/http://www.dl.ac.uk/SRS/PX/jwc_laue/laue_top.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Anne Imberty (anne.imberty@cermav.cnrs.fr).

Materials availability

This study did not generate new unique reagent.

Data and code availability

Model coordinates and experimental data have been deposited in the Protein Data Bank under PDB accession codes for jointly refined x-ray and neutron structures 7BBI for H/D-exchanged apo PLL structure, 7BBC for the D-PLL/Fuc-d12 complex. In addition, the following X-ray structures have also been deposited: 7B7F for H/D-PLL complexed with H-fucose at room temperature, 7B7E for D-PLL apo at room temperature, 7B7C for D-PLL complexed with H-fucose at room temperature, and 7BB4 for D-PLL complexed with H-fucose at 100K.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial cell culture

Both hydrogenated and perdeuterated PLL lectin were expressed in *Escherichia coli* BL21(DE3) cells using pET-29a-*pll* (Kumar et al., 2016) under kanamycin selection. All

culture media were supplemented with 35 $\mu g\ mL^{-1}$ kanamycin and grown at 37 °C with shaking at 180 rpm.

METHOD DETAILS

Protein expression

The hydrogenated protein was produced in LB medium as described previously (Kumar et al., 2016). Briefly, *E. coli* BL21(DE3) cells containing pET29a-*pll* plasmid were cultured in LB medium with 35 $\mu g\ mL^{-1}$ kanamycin at 37 °C until OD_{600} of 0.6-0.8. The protein expression was induced with IPTG (0.2 mM) and cells were grown at 25 °C for 12 h. The cells were harvested (8000 rpm at 4 °C for 1 h) and the cell paste was frozen at -80 °C.

Adaptation to D_2O and deuterated glycerol-d_8

The *E. coli* cells containing pET29a-*pll* plasmid (Kumar et al., 2016) were adapted to deuterated Enfors minimal medium with a define composition as previously described (Artero et al., 2005; Heartlein et al., 2016). The cells were adapted to fully deuterated Enfors minimal medium with the following composition: 6.86 g L^{-1} $(NH_4)_2SO_4$, 1.56 g L^{-1} KH_2PO_4, 6.48 g L^{-1} $Na_2HPO_4{\cdot}2H_2O$, 0.49 g L^{-1} $(NH_4)_2HC_6H_5O_7$ (diammonium hydrogen citrate), 0.25 g L^{-1} $MgSO_4{\cdot}7H_2O$, with 1.0 mL L^{-1} of trace metal stock solution (0.5 g L^{-1} $CaCl_2{\cdot}2H_2O$, 16.7 g L^{-1} $FeCl_3{\cdot}6H_2O$, 0.18 g L^{-1} $ZnSO_4{\cdot}7H_2O$, 0.16 g L^{-1} $CuSO_4{\cdot}5H_2O$, 0.15 g L^{-1} $MnSO_4{\cdot}4H_2O$, 0.18 g L^{-1} $CoCl_2{\cdot}6H_2O$, 20.1 g L^{-1} EDTA), 5 g L^{-1} glycerol-d_8. A single colony of *E. coli* containing pET29a-*pll* plasmid grown overnight on an LB agar plate supplemented with kanamycin was used to inoculate 15 mL of hydrogenated Enfors minimal medium and was grown overnight at 37 °C. The culture was then used to inoculate 15 mL of 100% D_2O Enfors minimal medium (with deuterated glycerol-d_8) at OD_{600} of 0.1 and was grown overnight at 37 °C. This passaging step was repeated five times until the doubling time for *E. coli* reached values similar to those for hydrogenated cultures.

Deuterated fed-batch fermentation

A deuterium-adapted preculture of 150 mL was used to inoculate 1.2 L of deuterated minimal medium in a 3 L bioreactor (Infors, France) used for fed-batch fermentation. The pD of the culture medium was maintained at 7.2 by addition of 4% NaOD and the temperature at 30 °C. After exhaustion of glycerol-d_8 (Euriso-top, France) from the culture medium, the fed-batch

phase was initiated by continuous feeding with additional 30 g of glycerol-d_8. PLL expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at OD_{600} of 16 and harvested after 19 h induction by centrifugation (8000 g for 1 h at 4 °C). The cell paste was frozen at -80 °C for long-term storage. The cell paste yield was 66 g wet weight from a 1.8-liter culture.

Protein purification

Both hydrogenated and perdeuterated PLL were purified the same way using Ni^{2+} affinity chromatography (AKTA Prime) and hydrogenated buffers. Cell paste was resuspended in lysis buffer (20 mM potassium phosphate pH 7.5, 2 mM trehalose) in the presence of EDTA-free protease inhibitor cocktail (Roche, Mannheim, Germany). The cells were lysed by cell disruption at a pressure of 1.8 kBar (Constant Systems Ltd, UK). After centrifugation (24 000 g for 40 min at 4 °C), the supernatant was filtered (0.45 μm) and the cleared cell lysate was loaded onto a His Trap FF column (GE Healthcare Life Sciences (now Cytiva), USA) pre-equilibrated with the lysis buffer. Protein was eluted using phosphate buffer containing 100 mM imidazole. The purity of the protein was verified by 12% Tris-Tricine SDS-PAGE stained with Coomassie Blue. The protein MW estimated from SDS-PAGE was about 40 kDa for both proteins, in agreement with the theoretical MW of 41.9 kDa. Fractions containing pure protein were pooled together and dialyzed against the lysis buffer and concentrated using 30-kDa cut-off centrifugal filter units (Amicon, Merck Millipore). The concentrated protein was flash-frozen in small aliquots of 200 μL in liquid nitrogen for long-term storage. The typical yield of perdeuterated PLL lectin was about 4 mg of protein per 1 g of wet cell paste.

Production and purification of Fuc-d_{12}

The fully-deuterated fucose (Fuc-d_{12}) was produced using glyco-engineered bacteria in a bioreactor as described previously (Gajdos et al. 2020). The secreted fully-deuterated fucose was further purified and characterised before being used for the crystallisation experiments described in this work.

Crystallization

Crystallization experiments were carried out at the High Throughput Crystallisation Laboratory (HTX Lab) of the EMBL Grenoble (Dimasi et al., 2007). Both hydrogenated and deuterated PLL lectin were crystallized using vapour-diffusion sitting drop method in 24-well crystallization plates (Hampton Research, USA) in the following three conditions: 0.1 M sodium acetate, pD 4.6, 8% (w/v) polyethylene glycol (PEG) 4000; 0.1 M Tris-DCl, pD 8.5, 6-8% (w/v) PEG 8 000; 0.05-0.4 M sodium-potassium tartrate. The final volume of the drops was 8-16 µL, the protein : reservoir ratio was 3:1 and the volume of the reservoir solution was 0.5 mL. Crystallization plates were incubated at 18 °C. Hydrogenated crystals used for neutron diffraction experiments were H/D-exchanged by vapour diffusion. Reservoir solution was replaced by deuterated solution in a stepwise manner starting from 25% to 100% deuterated solution over a period of two weeks.

Crystals of deuterated PLL lectin in complex with deuterated fucose-d_{12} were obtained in the following conditions: 0.05-0.1 M sodium-potassium tartrate. Crystals were either co-crystallized or soaked with Fuc-d_{12} to a final concentration of 50 mM. All solutions were dissolved in D_2O. A macro-seeding technique was applied to grow large crystals for neutron diffraction experiments (Thaller et al., 1981). Crystals of 0.05-0.1 mm^3 were harvested and introduced into pre-equilibrated drops containing the protein solution. Crystals were growing for several weeks. In order to avoid damaging the crystals by repeated seeding, crystals were regularly fed with fresh protein solution. The lid was open once per week and 3-4 µL of a pre-equilibrated protein solution was added onto the drop. Final volumes of crystals obtained using this method were 0.5-0.7 mm^3. To exchange all labile hydrogens that might have been back-exchanged when the sitting-drop lid was open, the reservoir solution was replaced for freshly-prepared deuterated solution three times prior to crystal mounting in quartz capillaries.

Neutron data collection and processing

Crystals of PLL lectin were mounted either in thin (0.01 mm thickness, Hampton Research, US) or thick-walled quartz capillaries (0.2 mm thickness, Vitrocom from CM Scientific, UK) with inner diameter 1.5-2.0 mm and sealed with wax in preparation for data collection at room-temperature using the LADI-III diffractometer at the Institut Laue-Langevin (Blakeley et al., 2010). Using a neutron wavelength range from 3.1-4.1 Å, neutron diffraction data extending to 2.2 Å resolution were collected for D-PLL complexed with Fuc-d_{12} and for H/D-exchanged apo PLL. The neutron diffraction data for the D-PLL/Fuc-d_{12} complex were collected from a

single crystal with volume of 0.5 mm^3 using 18h exposures. Nineteen images were collected from two different crystal orientations. The neutron diffraction data for H/D-exchanged apo PLL were collected from two crystals with volumes of ~0.7 mm^3 that had been grown in the same crystallization condition. Ten images (each of 18h) were collected in total from the two crystals. Neutron data were indexed and integrated using *LAUEGEN* (Campbell et al., 1998), wavelength normalized using *LSCALE* (Arzt et al., 1999) and merged using *SCALA* (Evans, 2006).

X-ray data collection and processing

Room temperature (RT) X-ray datasets were collected for both crystals from which the neutron data were collected. For the H/D-exchanged apo PLL crystal, the data images were recorded on the BM30A (newly BM07) beamline at the ESRF (Grenoble, France). The same beamline was used to collect RT datasets of H/D-PLL-H-fucose, D-PLL apo and D-PLL-H-fucose crystals (hydrogenated commercial L-fucose was used for the crystals). The collected data were processed using *XDS* (Kabsch, 2010), scaled and merged in *AIMLESS* (Evans and Murshudov, 2013) and converted to structure factors using *TRUNCATE* from the *CCP4* program suite (Winn et al., 2011). For the D-PLL-Fuc-d$_{12}$ crystal, the RT X-ray datasets were recorded on the GeniX 3D Cu High Flux diffractometer (Xenocs) at the IBS, Grenoble, France. Data were integrated using *iMOSFLM* (Battye et al., 2011),(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011)(Battye et al., 2011) scaled and merged using *AIMLESS* and converted to structure factors using *TRUNCATE* from the *CCP4* program suite.

To collect the 100 K X-ray data, crystals were cryo-protected by soaking in 30% (v/v) glycerol in a crystallization solution for a time as short as possible and were subsequently cryo-cooled at 100 K in liquid nitrogen. Datasets from crystals of D-PLL/H-fucose complex were collected on the ID23-1 beamline at the ESRF equipped with the PILATUS 6M detector. The data were processed via XIA2 pipeline. Complete data collection and structure refinement statistics are presented in Table S2.

Phasing and Structure refinement

Monomeric structure of the recombinant PLL lectin (PDB: 5C9O) stripped of water molecules was used as the initial model for initial phasing using molecular replacement (McCoy et al., 2007). Crystallographic refinement was carried out with *phenix.refine* (Afonine et al., 2012)

from the *PHENIX* package (Adams et al., 2010) altered with a manual model building and model adjustments using *Coot* (Emsley et al., 2010). Water molecules were introduced automatically using *phenix.refine* and inspected manually. Ligand molecules were introduced manually in *Coot*.

The joint x-ray/neutron refinement was carried out after the satisfactory R factors were achieved. The final X-ray and neutron R_{work}/R_{free} values for the H/D-exchanged apo PLL were 13.5%/15.6% and 21.5%/24.0% respectively and for the D-PLL/Fuc-d_{12} complex were 12.1%/14.0% and 19.2%/22.3% respectively. The *ReadySet* utility in *PHENIX* was used to introduce hydrogen and deuterium atoms at appropriate positions in the protein and ligand molecules. The *eLBOW* (Moriarty et al., 2009) tool in *PHENIX* was used to generate restraint files for the hydrogenated and deuterated fucose molecules. Molecular figures were prepared in *PyMOL* (Schrödinger, Inc.).

Complete data collection and structure refinement statistics are presented in Table S1. Model coordinates and experimental data have been deposited in the Protein Data Bank under PDB accession codes: 7BBI for H/D-exchanged apo PLL structure and 7BBC for the D-PLL/Fuc-d_{12} complex.

Neutron structure of the perdeuterated LecB/fucose complex

5.1 Motivation

Pseudomonas aeruginosa is a human opportunistic pathogen that produces two soluble lectins LecA (PA-IL) and LecB (PA-IIL) specific for galactose and fucose, respectively, that are involved in the attachment of bacteria to host epithelium and biofilm formation (Imberty et al., 2004). Both lectins are considered as potential drug targets for anti-adhesive therapy (Grishin et al., 2015). LecB binding to fucose is established via direct and water-bridged hydrogen bonds but also involves two calcium ions bridging the sugar and the amino acids of the protein. Three of the fucose hydroxyl groups participate in the coordination of these calcium ions which is proposed to be the basis for the unusual high affinity observed for LecB fucose recognition (Mitchell et al., 2004). More than 20 X-ray crystal structures of LecB lectin have been solved but some questions remain to be answered such as the particular role of water molecules involved in the binding of fucose, as well as determination of protonation state of acidic residues in the calcium and fucose-binding site. Neutron protein crystallography is an ideal technique to provide these pieces of information. Production of perdeuterated LecB lectin for neutron diffraction studies had already been carried out in the D-Lab at ILL more than 10 years ago and an X-ray structure of the complex with H-fucose has been solved (PDB: 4CE8). Unfortunately, crystals large enough for neutron diffraction experiments could not be grown at that time. With the improvements of neutron instrument and availability of perdeuterated protein and ligand we believed it was now possible to collect high-quality neutron diffraction data.

5.2 Methods development

The *E. coli* cells for production of LecB lectin were adapted to the deuterated Enfors minimal medium in a similar way as described previously for PLL lectin, through several steps of passaging from LB medium to hydrogenated and finally deuterated minimal medium. After expression tests and adaptation, we carried out a high cell-density cultivation using deuterated medium and deuterated glycerol as carbon source. The perdeuterated LecB lectin was produced in high amount (4 mg of pure protein per g of cell paste) using a batch/fed-batch fermentation (Figure 53).
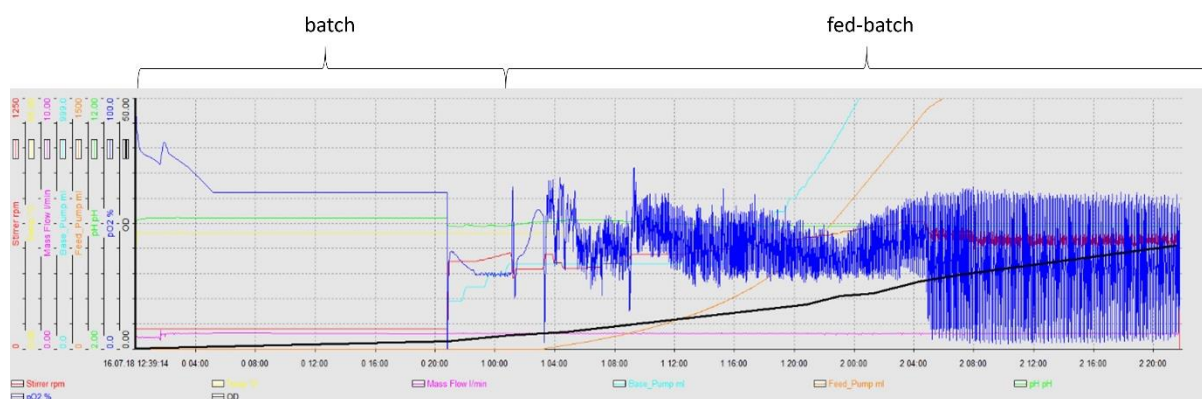
Figure 53: High cell-density culture profile of the production of perdeuterated LecB lectin using *E. coli* batch/fed-batch fermentation in a bioreactor. A volume of 1.2-L of deuterated Enfors minimal medium was inoculated with a 150 mL preculture of adapted cells. Fed-batch phase was initiated after consumption of the initial glycerol supply (5 g) in the culture medium indicated by a sudden increase in pO_2 (i.e. dissolved oxygen in the medium, blue line). The culture was then continuously fed with a 12% glycerol-d_8 solution (orange line) in a limiting way to ensure that the cell specific growth did not reach its maximum value. The mass flow of sterile air (pink line) was maintained at ≥ 0.5 L min^{-1}. The temperature (yellow line) was maintained at 30 °C and the pD (green line) of the culture medium was maintained at 7.2 by addition of a base solution (cyan line) of 4 % NaOD. Stirring (red lined) was adjusted to keep the dissolved oxygen tension of 30%. When the set-point was reached, the stirrer speed (rpm) was increased to compensate for the oxygen demand of growing cells. The bacterial growth (black line) was monitored by measuring OD_{600} off-line.

LecB has a weak affinity for mannose (Garber et al., 1987) and therefore perdeuterated LecB could be purified in a single-step affinity chromatography on a mannose column using the same protocol described for H-LecB (Figure 54 and 55) (Gillon et al., 2020). After purification, we used the High Throughput Crystallization Facility (HTX) at the EMBL institute (Grenoble, France) to screen the crystallization conditions for D-LecB at a concentration of 10 mg mL^{-1}. Several hits with single crystals were obtained (Figure 56). Lithium sulphate and polyethylene glycol (PEG) were confirmed to be the best crystallization reagents, as previously described for other LecB structures deposited in the PDB.
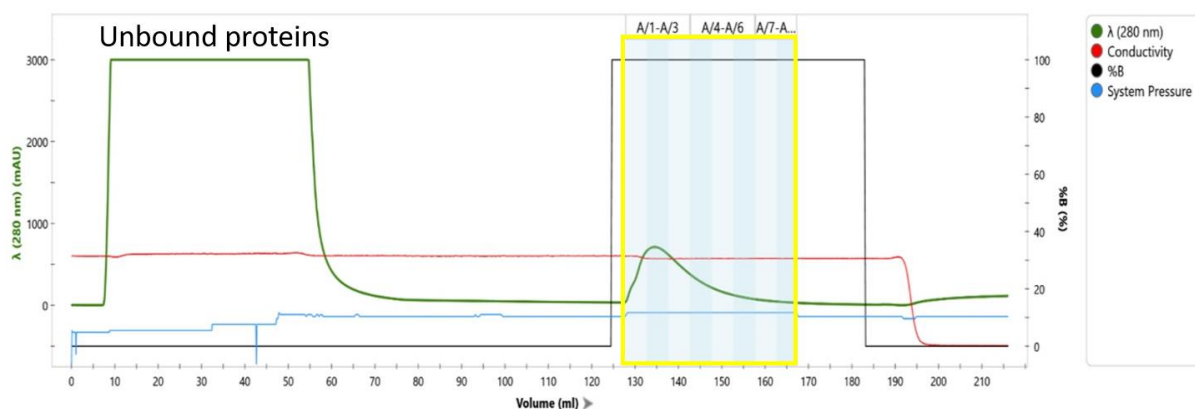
Figure 54: Affinity chromatography purification of perdeuterated LecB (D-LecB) on a 10 mL column filled with D-mannose-agarose resin (Merck). The protein was eluted with a buffer (20 mM Tris/HCl, pH 7.5, 100 mM NaCl, 100 μM) containing 100 mM free D-mannose (yellow box).
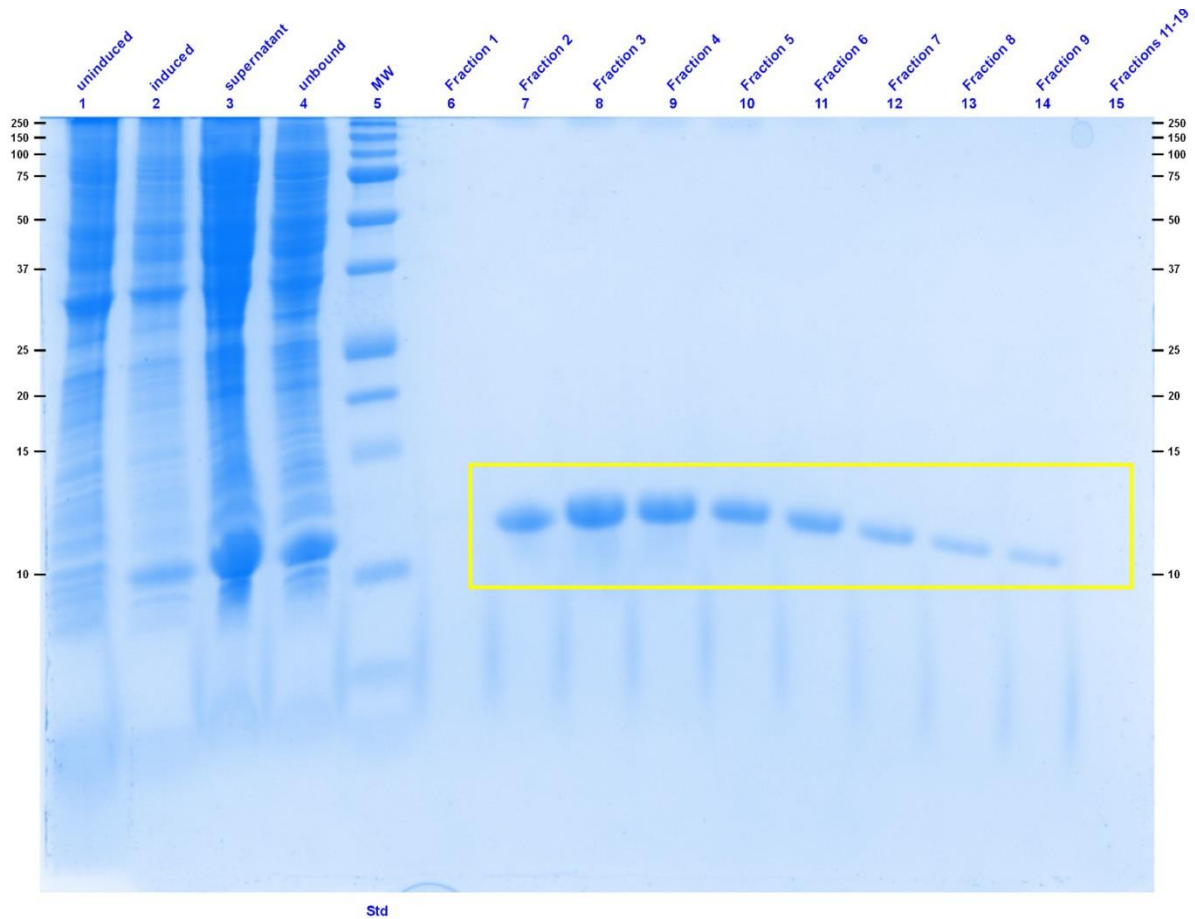
Figure 55: SDS-PAGE gel (16% Tris-glycine) of expression and purification of D-LecB lectin. Lane 1 (uninduced *E. coli* cells), lane 2 (induced cells), lane 3 (supernatant, soluble protein), lane 4 (unbound proteins), lane 5 (molecular weight marker), lane 6-15 (elution with 100 mM D-mannose). The yellow box indicates the bands that correspond to the D-LecB protein.

0.2 M NaAc
0.1 M Na cacodylate, pH 6.5
30% PEG 8K (w/v)

0.2 M Li$_2$SO$_4$
0.1 M Tris/HCl, pH 8.5
30% PEG 4K (w/v)

0.2 M Li$_2$SO$_4$
0.1 M Bis Tris, pH 5.5
25% PEG 3350 (w/v)

0.2 M Li$_2$SO$_4$
0.1 M Tris/HCl, pH 8.5
30% PEG 4K (w/v)
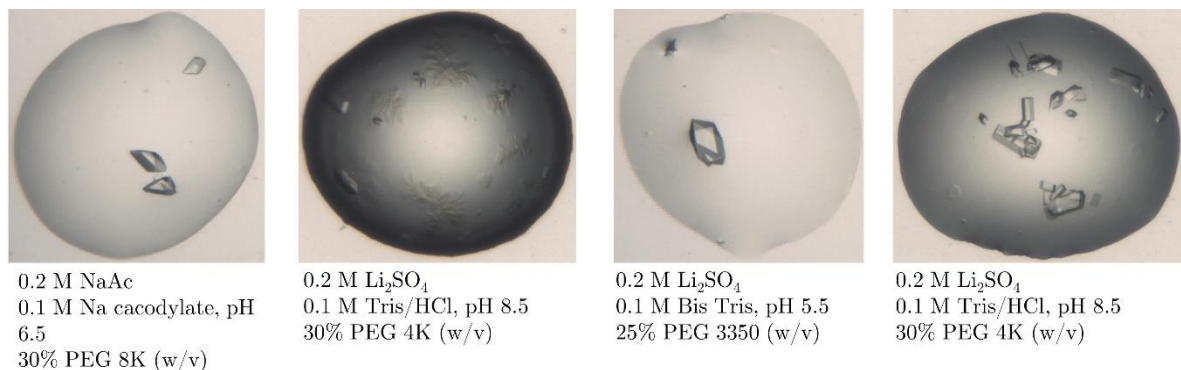
Figure 56: Crystals of perdeuterated LecB/Fuc-d$_{12}$ complex obtained by vapour diffusion hanging drop method. Crystallization experiments were set up at the High Throughput Crystallization Facility at the EMBL (Grenoble, France).

These conditions were optimized for growing large crystals of the D-LecB/Fuc-d_{12} complex suitable for neutron diffraction studies. All crystals were grown using a sitting-drop vapour diffusion method. Macro-seeding and feeding techniques were applied to increase the size of the crystals. Small crystals with edges of about 100 μm were first macro-seeded into pre-equilibrated drops. In order to reduce potential damages caused by repetitive transfer and manipulation of crystals during macro-seeding, they were regularly fed with a fresh protein solution over a period of five months. Two crystals with volumes of ~ 0.3 mm³ could be grown using this approach (Figure 57). A smaller crystal with volume ~ 0.1 mm³ grew from a drop containing a liquid-liquid phase separation without macroseeding or feeding (Figure 58).
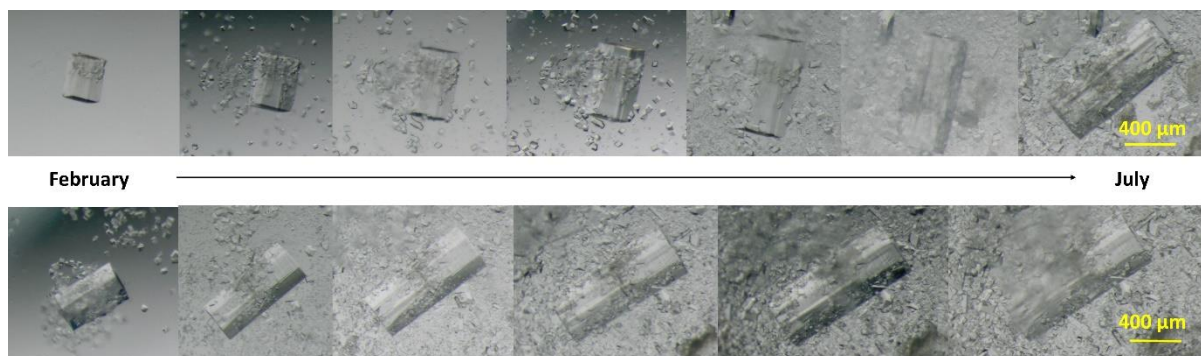
Figure 57: Evolution over time of two largest D-LecB/Fuc-d_{12} crystals from the following crystallization condition: 0.2 M Li_2SO_4, 0.1 M Tris/DCl, pD 8.5, 20% (w/v) PEG 4000. Crystals were regularly fed with a fresh protein solution ($\sim$ 2 μL) that was added onto the drop roughly every two weeks.
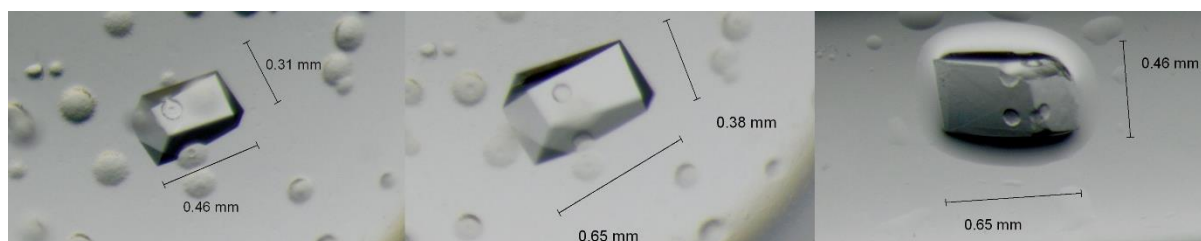
Figure 58: Crystal of the perdeuterated LecB/Fuc-d_{12} complex grown in the following crystallization condition: 0.2 M Li_2SO_4, 0.1 M Tris/DCl, pD 7.5, 20% (w/v) PEG 4000. The crystal mounted in a quartz capillary (right) was used for the neutron data collection on the LADI-III at ILL (Grenoble, France).

147

Crystals were mounted into quartz capillaries and tested on LADI-III diffractometer at ILL. The first two crystals grown by a feeding technique diffracted neutrons but diffraction images displayed streaked spots indicating global disorder and these crystals were not suitable for a high-quality neutron data collection (Figure 59). The smaller crystal diffracted neutrons to 1.9 Å resolution and diffraction images exhibited nice round spots allowing for a full neutron data collection to be carried out. Macroscopically, this crystal looked much better than the previous two crystals that were grown by macro-seeding/feeding techniques. The crystal grew in less than two months from setting up the crystallization plate and it is possible it would have continued to grow had it not been mounted for the experiment. The room temperature X-ray data were collected from the same crystal to carry out a joint X-ray/neutron refinement. Several 100K X-ray data sets were collected for comparisons (Figure 60).
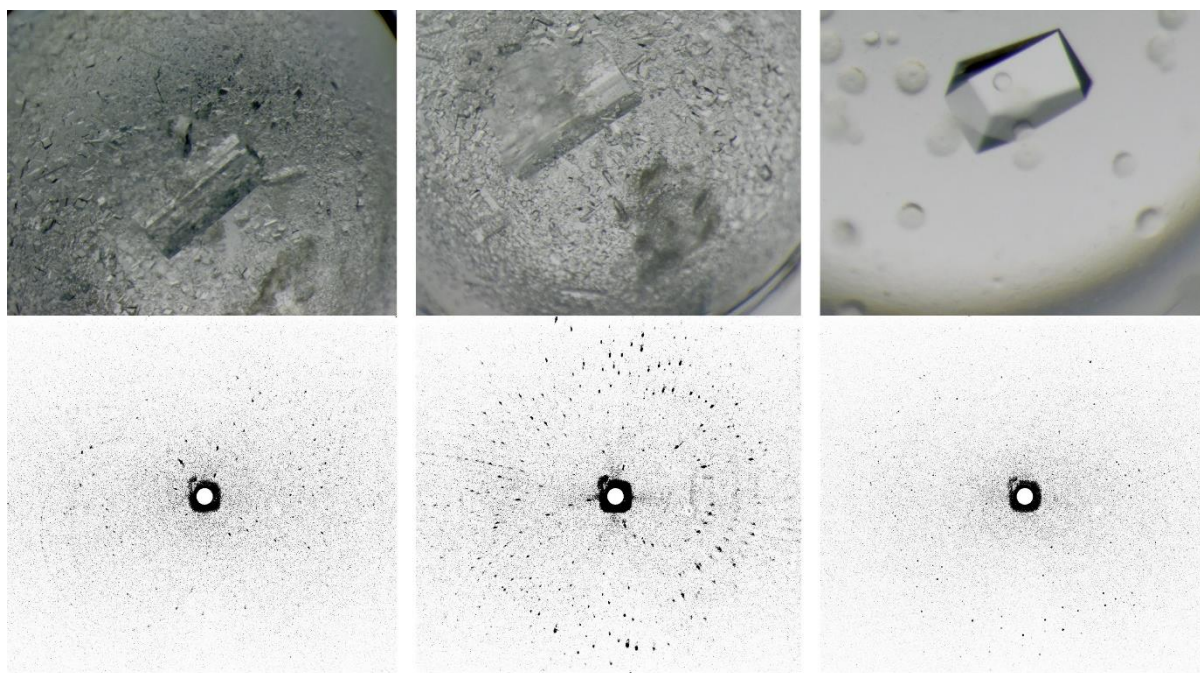
Figure 59: quasi-Laue neutron diffraction images of three crystals of the D-LecB/Fuc-d_{12} complex tested on LADI-III at the ILL.

X-ray/neutron	T	Resolution [Å]	Completeness [%]	$\langle I/\sigma(I) \rangle$	Unit cell [a, b, c, α, β, γ]	Spacegroup	Protein form	Crystallization form	Ligand	Crystallization condition	Cryoprotectannt	Beamline
Neutron	RT	1.9	73.9 (62.1)	5.6 (2.3)	52.90 73.87 55.00 90 94.58 90	P21	deuterated	deuterated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 M Tris/DCl, pD 7.5, 20% (w/v) PEG 4K	-	LADI-III, ILL
X-ray	RT	1.85	99.5 (84.7)	11.3 (4.2)	52.64 73.49 54.84 90 94.51 90	P21	deuterated	deuterated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 M Tris/DCl, pD 7.5, 20% (w/v) PEG 4K	-	IBS
X-ray	cryo	1.52	99.3 (91.4)	16.2 (2.3)	52.48 72.78 54.85 90 94.48 90	P21	deuterated	deuterated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 M Tris/DCl, pD 7.5, 30% (w/v) PEG 4K	-	Proxima-1, SOLEIL
X-ray	cryo	1.28	99.3 (94.8)	12.3 (1.1)	52.54 72.96 55.16 90 94.62 90	P21	deuterated	deuterated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 M Tris/DCl, pD 7.5, 30% (w/v) PEG 4K	-	Proxima-1, SOLEIL
X-ray	cryo	0.9	87.6 (25.6)	16.3 (0.6)	52.60 72.99 55.19 90 94.58 90	P21	deuterated	deuterated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 30% (w/v) PEG 3350	-	Proxima-1, SOLEIL
X-ray	cryo	1.11	84.0 (23.0)	6.8 (1.0)	52.59 72.94 54.91 90 94 90	P21	deuterated	hydrogenated	-	2 M (NH$_4$)$_2$SO$_4$, 0.2 M Na/K tartrate, 0.1 M Na$_3$-citrate, pH 5.6	-	I03 beamline, Diamond
X-ray	cryo	2.07	98.5 (99.9)	5.4 (1.1)	47.66 179.96 95.00 90 97.62 90	P21	deuterated	hydrogenated	-	2 M (NH$_4$)$_2$HPO$_4$, 20% (w/v) PEG 3350	-	I03 beamline, Diamond
X-ray	cryo	1.21	96.4 (73.7)	7.7 (1.0)	52.51 72.73 54.99 90 94.41 90	P21	deuterated	hydrogenated	Fucose	0.2 M Li$_2$SO$_4$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350	-	I03 beamline, Diamond
X-ray	cryo	1.33	96.5 (93.9)	4.4 (1.1)	52.40 72.42 54.25 90 94.16 90	P21	deuterated	hydrogenated	Fucose	0.2 M Li$_2$SO$_4$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350	-	I03 beamline, Diamond
X-ray	cryo	1.08	85.4 (14.6)	10.8 (0.9)	49.03 72.73 100.01 90 90 90	P 21 21 21	deuterated	hydrogenated	-	0.2 M Li$_2$SO$_4$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350	-	I03 beamline, Diamond
X-ray	cryo	1.14	93.8 (58.2)	8.7 (1.1)	49.12 72.84 99.90 90 90 90	P 21 21 21	deuterated	hydrogenated	-	0.2 M Li$_2$SO$_4$, 0.1 M Bis-Tris, pH 5.5, 25% (w/v) PEG 3350	-	I03 beamline, Diamond
X-ray	cryo	1.44	96.7 (96.8)	3.2 (1.0)	52.48 145.43 109.90 90 94.58 90	P21	deuterated	hydrogenated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 Tris/HCl, pH 8.5, 30% (w/v) PEG 4K	-	I03 beamline, Diamond
X-ray	cryo	1.28	98.0 (92.9)	8.1 (1.2)	52.29 72.38 54.14 90 94.56 90	P21	deuterated	hydrogenated	Fucose-d$_{12}$	0.2 M Li$_2$SO$_4$, 0.1 Tris/HCl, pH 8.5, 30% (w/v) PEG 4K	-	I03 beamline, Diamond
X-ray	cryo	1.15	88.3 (39.1)	6.7 (1.1)	52.46 73.13 55.05 90 94.57 90	P21	deuterated	hydrogenated	Fucose	0.2 Li$_2$SO$_4$, 20% (w/v) PEG 3350	-	I03 beamline, Diamond

Figure 60: Data collection summary of LecB lectin.

5.3 Article III

Hydrogen-bonding network involved in fucose binding of *Pseudomonas aeruginosa* lectin LecB revealed by neutron crystallography
L. Gajdos, M. P. Blakeley, M Haertlein, V. T. Forsyth, J. M. Devos, A. Imberty. To be submitted.

Resumé

La lectine LecB de *Pseudomonas aeruginosa* est spécifique des glycannes fucosylés présents dans les tissus de l'hôte. La lectine contient un site de reconnaissance des sucres très inhabituel avec deux ions calcium proches l'un de l'autre et coordonnant trois des groupes hydroxyle du fucose, ce qui lui confère à la fois spécificité et haute affinité. Nous décrivons ici la structure neutronique à température ambiante de LecB deutéré en complexe avec du L-fucose deutéré. Les cartes neutroniques démontrent l'absence d'atome de deutérium attaché aux chaînes latérales carboxylate impliquées dans la liaison au calcium. Le deutérium des groupes hydroxyle du fucose s'écarte des ions métaux, établissant des liaisons hydrogène avec les résidus d'acides aminés dans le site de liaison. L'orientation de la molécule d'eau conservée (sous forme de D2O) dans le site de liaison du fucose est également définie, confirmant son rôle dans la stabilisation de l'une des boucles impliquées dans la liaison des sucres. Nous avons également déterminé la structure cryo aux rayons-X du complexe perdeutéré à 0,9 Å de résolution, permettant une comparaison avec le complexe hydrogéné à haute résolution précédemment déterminé. Cette comparaison, ainsi que les données thermodynamiques des expériences ITC, démontrent que la deutération n'affecte pas l'interaction de la lectine avec son ligand glucidique.

Hydrogen-bonding network involved in fucose binding of *Pseudomonas aeruginosa* lectin LecB revealed by neutron crystallography

Lukas Gajdos[a,b,c], Matthew P. Blakeley[d], Michael Haertlein[a,b], V. Trevor Forsyth[a,b,e], Juliette M. Devos[a,b*], Anne Imberty[c*]

[a] Life Sciences Group, Institut Laue-Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France

[b] Partnership for Structural Biology (PSB), 71 Avenue des Martyrs, 38000 Grenoble, France

[c] Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France

[d] Large Scale Structures Group, Institut Laue-Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France.

[e] Faculty of Natural Sciences, Keele University, ST5 5BG Staffordshire, UK

* To whom correspondence should be addressed. Anne Imberty (anne.imberty@cermav.cnrs.fr, Tel: +33 4 7603 7640, Twitter: @AnneImberty)

Summary
The LecB lectin from *Pseudomonas aeruginosa* is specific for fucosylated glycans present in host tissue. The lectin displays a highly unusual carbohydrate-binding site with two calcium ions in close proximity coordinating three of the fucose hydroxyl groups, providing both the specificity and the high affinity. Here we report the room temperature neutron structure of deuterated LecB in complex with deuterated L-fucose. The neutron maps demonstrate the absence of any deuterium atom attached to the carboxylate side chains involved in calcium binding. The deuteriums of the fucose hydroxyl groups point away from the metal ions, establishing hydrogen bonds with the amino acid residues in the binding site. The orientation of the conserved water molecule (as D_2O) in the fucose-binding site is also defined, confirming its role in stabilizing one of the loops involved in sugar binding. We have also determined the 0.9 Å resolution cryo X-ray structure of the perdeuterated complex, allowing comparison with the previously determined high-resolution complex of the hydrogenated complex. This comparison, along with thermodynamic data from ITC experiments, demonstrate that deuteration does not affect the interaction of the lectin with its carbohydrate ligand.

Introduction

The biological effects that glycans elicit depend on the recognition of their specific features by protein receptors, such as enzymes, antibodies, carbohydrate-binding modules (CBMs) or lectins[1-4]. Lectins bind non-covalently and specifically to complex oligosaccharides, and are therefore able to decipher the so-called glycocode, i.e. the information contained in the structure and dynamics of complex branched glycans[5]. Lectins are involved in many biological processes, and are also actors in several diseases. In infection, lectins present at the surface of pathogenic microorganisms bind to glycans on the host cell surface, triggering adhesion, but also early response from the human system such as inflammation[6,7]. Competing with the attachment of pathogen lectins to the host glycans is one of the promising tools to fight the on-going crisis in antibiotic resistance[8,9]. Such strategy uses a structure-based knowledge for designing high-affinity glycomimetics with anti-infectious properties, as exemplified by mannose-derivatives against *E. coli* urinary infection[10]. The interactions between carbohydrates and amino acids have been characterized from the many crystal structures of lectins complexed with ligands (1447 3D X-ray structures in UniLectin3D curated database)[11](Bonnardel et al., 2019)(Bonnardel et al., 2019). The forces involved include hydrogen bonds, ionic bonds, van der Waals contacts, and CH-π interactions. Water molecules are often observed to bridge the carbohydrate hydroxyl groups with amino acids. Interestingly, a significant number of enzymes and/or lectins use divalent ions that directly coordinate with the hydroxyl groups of carbohydrates and with side chains of amino acids. Among the 350 different lectin families crystallized to date, more than 40 involve calcium ions in their binding sites (from UniLectin3D). Most of them belong to the C-type lectin family.

The opportunistic pathogen *Pseudomonas aeruginosa* has been identified as the number one priority in the list of antibiotic-resistant bacteria by the WHO in 2017 in order to prioritize the efforts for the research and development of new antibiotics[12]. *P. aeruginosa* produces two soluble lectins LecA (PA-IL)[13] and LecB (PA-IIL)[14] specific for galactose and fucose, respectively that are involved in the bacterial pathogenicity, adhesion and biofilm formation[15]. LecB is of special interest, since it is present only in a few pathogenic microorganisms. LecB has a unique binding site with two calcium ions at a very close distance (3.76 Å) that are involved in the sugar binding by coordinating three hydroxyl groups of the carbohydrate ligand. These two calcium ions directly contribute to the receptor specificity since they can only coordinate monosaccharides bearing the specific sequence of two equatorial and one axial hydroxyl group, as present in "fuco" and "manno" configurations. The ions also play a role in the enhanced affinity through delocalization of charge as evaluated by quantum chemical calculations, and through compensation for binding entropy losses by releasing strongly coordinated water molecules.

Results and discussion

Perdeuterated LecB (D-LecB) was produced in an *Escherichia coli* high cell-density culture with a high yield (4 mg D-LecB per g of cell paste). A quantity of 51 g of wet cell paste was obtained in a bioreactor from 1.6-L of deuterated medium and 45 g of deuterated glycerol-d$_8$ added in a batch/fed-batch mode. The protein was co-crystallized with perdeuterated fucose (Fuc-d$_{12}$) using the vapour diffusion method. To grow large crystals for neutron diffraction experiments, a feeding technique in which the protein crystal was regularly fed with a fresh protein solution, was originally used. Using this technique, crystals with volumes

of ~ 0.3 mm^3 over a period of 4 months could be grown. The crystals were tested on LADI-III instrument, but unfortunately showed split spots indicating high mosaicity. Finally, the neutron data were collected from a much smaller crystal with volume ~ 0.1 mm^3 that self-nucleated from a boundary of the two phases in the crystallization drop (Figure 1).
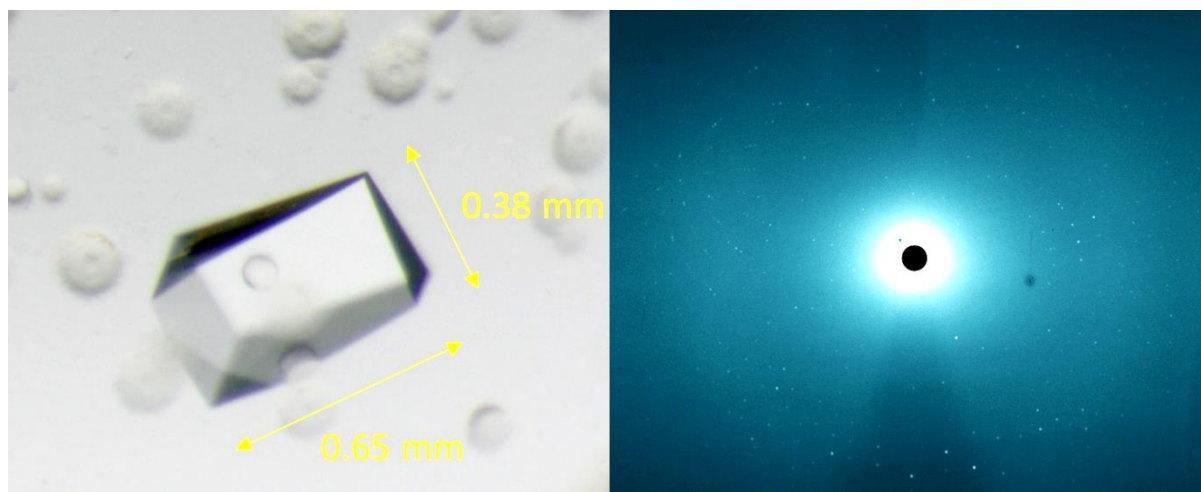
Figure 1: (Left) Crystal of the D-LecB/ Fuc-d$_{12}$ complex used for the neutron data collection on the LADI-III diffractometer at the Institut Laue-Langevin (Grenoble, France). (Right) quasi-Laue neutron diffraction pattern recorded from the crystal shown on the left.

We have determined a room temperature neutron structure of D-LecB complexed with Fuc-d$_{12}$ to 1.90 Å resolution that was jointly refined against both neutron and X-ray data (1.85 Å) collected from the same crystal. Additionally, we also report a 100K X-ray structure of the same perdeuterated complex determined at 0.9 Å resolution collected from a different crystal. The data collection and refinement statistics are presented in Table 2.

The crystal structures displayed similar features to those previously reported[16-18], belonging to the same space group $P2_1$ with four chains per asymmetric unit. Both the neutron and X-ray structures of the perdeuterated LecB complex are similar to the X-ray crystal structure previously described (PDB: 1GZT,[14]) with the root-mean-square deviation (RMSD) for the backbone atoms being 0.2 Å.

Room temperature jointly-refined X-ray/neutron structure of LecB

The overall fold of the LecB monomer is a β-sandwich composed of two curved sheets consisting of five and four antiparallel β-strands, respectively[16] (Figure 2a).

The excellent quality of the neutron scattering length density map allows clear observation of aliphatic amino acids and hydrogen bonds between adjacent β-sheets in the monomer. The tetramerization that occurs by the antiparallel association of β-strands from each dimer with their counterparts in the other dimer involves mostly hydrogen bonds that could be directly visualised from the neutron maps. The orientation of water molecules and polar amino acid side chains could also be unambiguously determined from the neutron analysis (Figure 2c). Furthermore, the use of perdeuterated protein allowed observation of the monomer interface between chains that participate in the tetramerization (i.e. chains A, C and D). Hydrophobic interactions include contacts between methyl groups of aliphatic amino

acid side chains of Ala1 (chain A), Val77 (chain C) and Thr84 (chain D). Hydrogen and ionic bonds are also involved in the stabilization of the tetramer (Figure 2b).
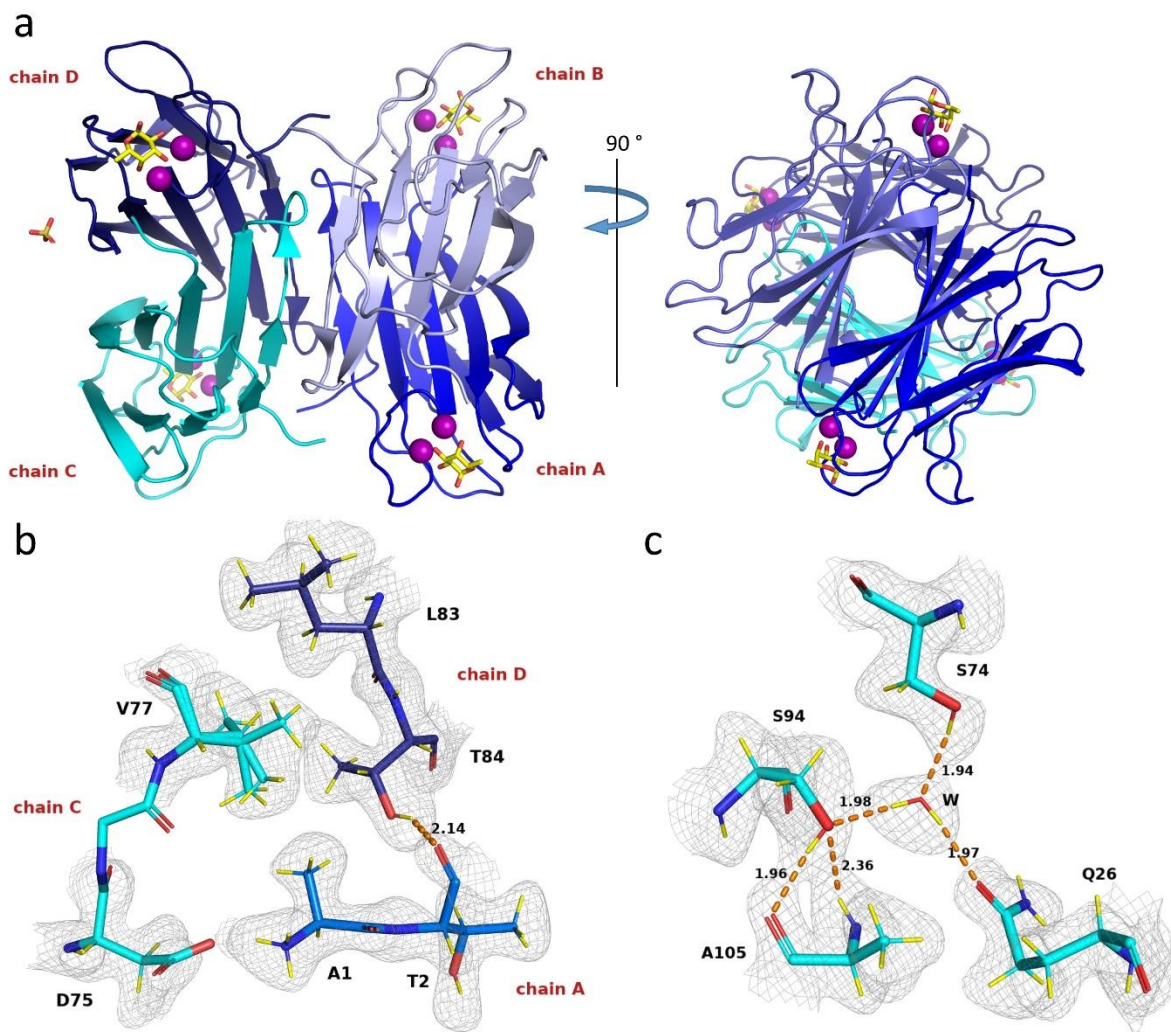
Figure 2: Structure of D-LecB in complex with Fuc-d$_{12}$. Calcium ions are shown as purple spheres, fucose molecules are shown as yellow sticks, a sulphate ion is shown as orange sticks. Hydrogen bonds (Å) are shown as orange dashed lines. Deuterium atoms are coloured yellow. The 2mF_o-DF_c neutron density (grey mesh) is contoured at 0.8σ. (a) Overall structure of D-LecB tetramer in complex with Fuc-d$_{12}$. (b) Protein interface between chains A, C and D. (c) Hydrogen bonding network involving a water molecule.

Coordination of the two calcium ions in the binding site

In addition to four fucose molecules and eight calcium ions, one sulphate anion in chain B and 364 water molecules have been modelled. Each monomer contains two adjacent calcium ions with a mean distance of 3.76 Å +/- 0.02 Å, in agreement with the high-resolution X-ray structure[19]. Each calcium is hepta-coordinated, with the involvement of three oxygen atoms of the ligand, the carboxylate atoms of six acidic amino acids (three aspartates, one glutamate and the C-terminal group of Gly from the neighbouring monomer), and oxygen atoms from two asparagine residues (Figure 3a). Analysis of the neutron maps confirmed the absence of any deuterium atom on the carboxylate, in agreement with previous experimental and theoretical

studies that concluded that none of the six acidic amino acid is protonated[19] (Figure 3b, 3c).
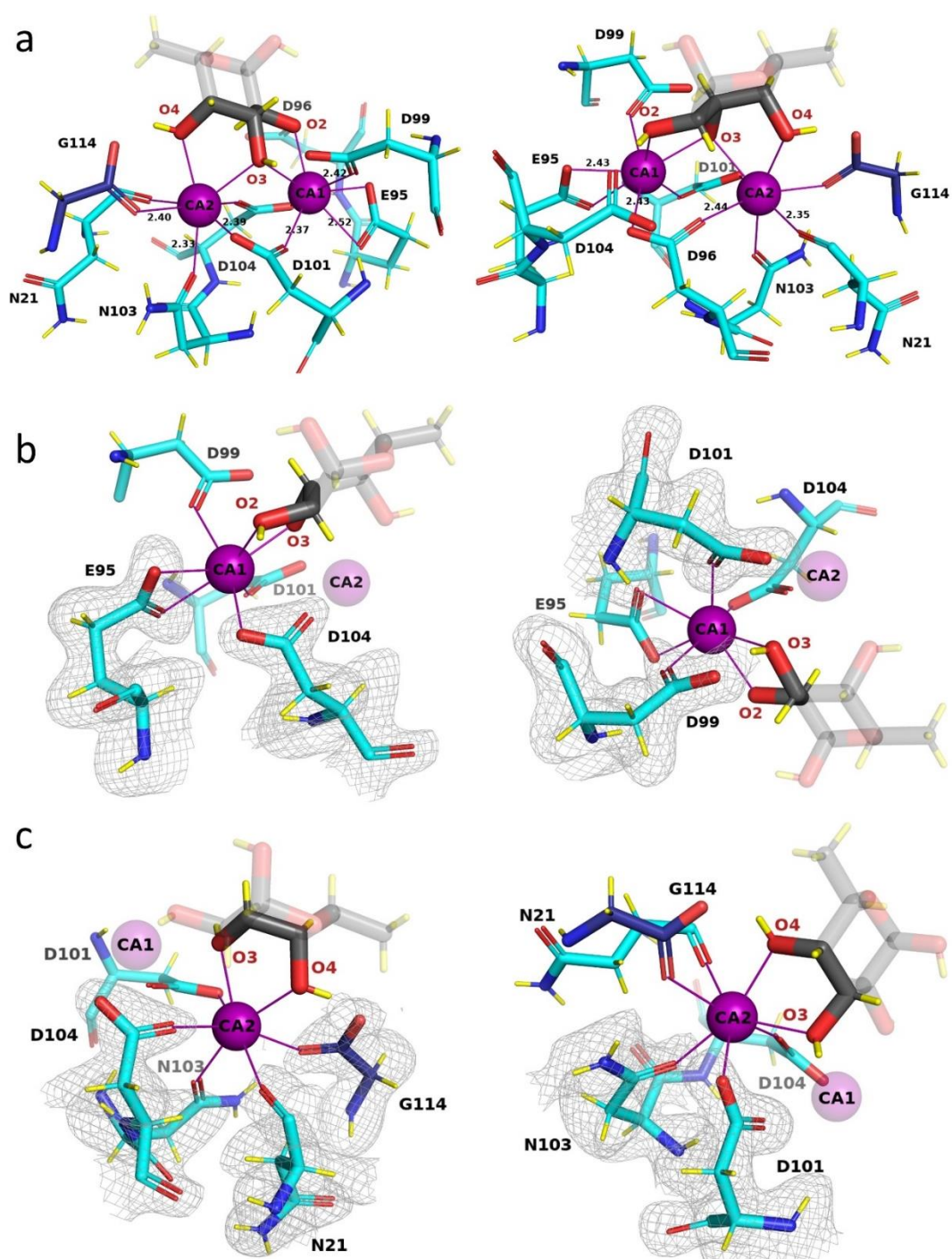
Figure 3: Calcium ions environment. 2mF_o-DF_c neutron scattering length density (grey mesh) is contoured at 0.8σ. (a) General view. (b) Environment of Ca 1. (c) Environment of Ca 2. Calcium ions are shown as purple spheres and metal coordination (Å) as solid purple lines.

Hydrogen network in the fucose-binding site

Four fucose residues could be modelled into the $2mF_o\text{-}DF_c$ neutron scattering length density map and the absence of cancellation effects allows for very fine visualisation of all hydrogen (D) atoms. In all four chains, the fucose ring adopts the stable $^1C_4$ chair conformation, with the alpha configuration at the anomeric position. Three fucose oxygen atoms coordinate two calcium ions with O2 coordinating the first calcium, O3 both calcium ions and O4 the second calcium (Figure 4). Deuterium atoms on these hydroxyl groups can be clearly visualized pointing away from the metal ions and establishing hydrogen bonds with the amino acid residues in the binding site. The use of perdeuterated fucose provided high-quality neutron density maps with positive peaks for methine deuterium atoms on the ring carbon atoms, as well as the fucose methyl group (Figure 4).
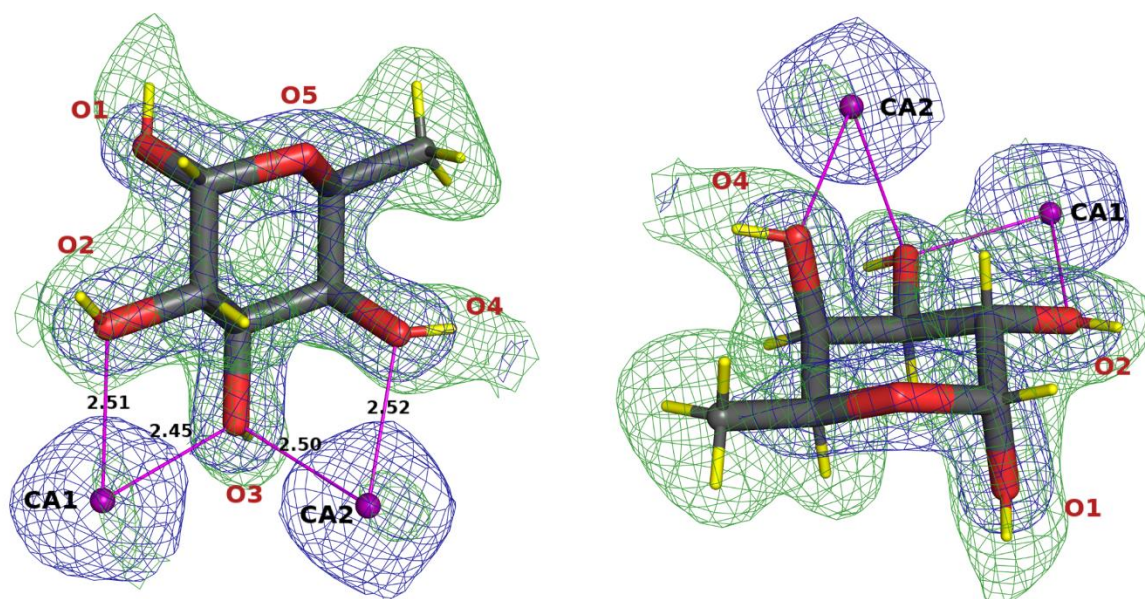
Figure 4: Two different orientations of perdeuterated ʟ-fucose coordinating two calcium atoms in the binding site of LecB. Calcium ions are shown as purple spheres, metal coordination as solid purple lines (Å) and deuterium atoms as yellow sticks. $2mF_o\text{-}DF_c$ neutron scattering length density (green mesh) and $2mF_o\text{-}DF_c$ electron density (blue mesh) are contoured at 1.4σ and 2.2σ, respectively.

Neutron scattering length density peaks can be observed for several of the hydrogen bonds established between the fucose and the protein (Figure 5a). Fuc-OD2 hydroxyl group donates its D to OD1 of Asp96 (1.80 Å +/- 0.05 Å), Fuc-OD3 establishes a hydrogen bond with the side chain OD2 oxygen of Asp99 (1.78 Å +/- 0.10 Å) and Fuc-OD4 donates a deuterium atom to the OXT oxygen of the carboxyl group of Gly114 of the neighbouring monomer (1.78 Å +/- 0.04 Å). The ring oxygen of fucose accepts a hydrogen bond from the amide backbone of Ser23 (2.13 Å +/- 0.03 Å). The binding of fucose to LecB puts Fuc-OD1, -OD3 and –OD4 hydrogen atoms in a gauche conformation (dihedral angles 62°, 35° and 46° respectively) with respect to the aliphatic hydrogens on C2, C3 and C4 respectively. The distances in parentheses represent a mean value from four independent molecules in the asymmetric

156

unit and are measured from the acceptor to the donor atom.
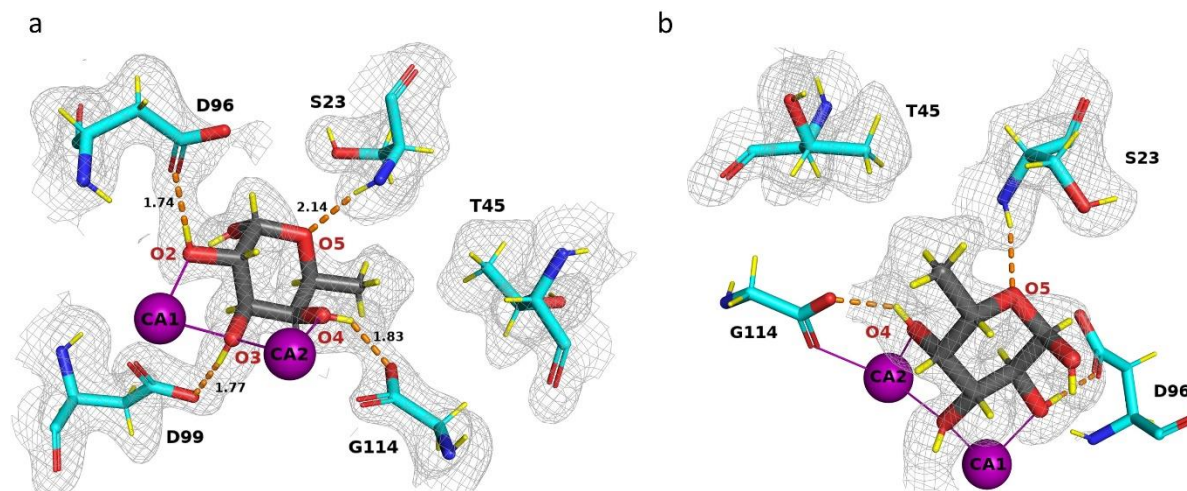
Figure 5: Two orientations of the fucose-binding site in the neutron structure of the D-LecB/Fuc-d_{12} complex (chain B). Fucose is represented as dark grey sticks and protein as cyan sticks. Calcium ions are shown as purple spheres and deuterium atoms are shown as yellow sticks. Hydrogen bonds and metal coordination are shown as orange dashed lines and as solid purple lines, respectively. The distances are given in Å. 2mF_o-DF_c neutron scattering length density map (grey mesh) is contoured at 0.8σ. Amino acids involved in coordination of the calcium ions were omitted for clarity (N21, N103, D101, D104, E95). (a) This view highlights the hydrogen-bonding network between fucose and the protein. (b) This view underlines the hydrophobic contact between methyl group of fucose and methyl and methine groups of Thr45 and Ser23 side chains, respectively.

The neutron maps clearly shows the hydrophobic contacts between the methyl group on the C6 position of fucose and the -CH$_2$ and methyl groups of Ser23 and Thr45, respectively (Figure 5b). The deuterium atoms in the -CH$_3$ group create a triangular-shaped neutron density and are in the most stable staggered conformation with respect to the aliphatic deuterium atom on C5.

The complete geometrical parameters including distances and angles of the hydrogen-bonding network in the fucose-binding site are presented in Table 1.

Water network in the ligand-binding site
In the carbohydrate-binding site, there are several conserved water molecules observed in all structures (PDB: 1GZT, 1OXC, 1UZV) that establish a hydrogen-bonding networks between the protein, the ligand and the other ordered waters[16,18,19]. A conserved water molecule, Wat1, present in all monomers, mediates an additional hydrogen bond between fucose and the protein. Even though the four chains of the LecB tetramer are identical and the fucose molecule is bound in the same manner in each of the four chains, the water networks in each of the binding sites display mild differences.

Table 1: Geometrical parameters of the hydrogen-bonding network in the four binding sites of the LecB/Fuc-d_{12} complex in the neutron structure

Donor	Acceptor	Hydrogen bond distance (Å)/ angle (°)			
		Chain A	Chain B	Chain C	Chain D
Direct interactions between the fucose residue and the protein					
Fuc-OD2	Asp96.OD1	1.81/165.1	1.74/179.1	1.77/151.5	1.86/162.6
Fuc-OD3	Asp99.OD2	1.74/160.0	1.77/161.8	1.92/133.3	1.69/153.7
Fuc-OD4	Gly114.OXT[a]	1.75/148.2	1.83/137.3	1.76/153.4	1.77/165.4
Ser23.ND	Fuc-O5	2.10/169.8	2.14/169.1	2.12/176.4	2.16/171.2
Water-bridged hydrogen bonds between the fucose residue and the protein					
Thr98.ND	O (Wat1)	2.12/162.1	2.05/155.8	2.15/164.1	2.01/164.1
Fuc-OD1	O (Wat1)	-	2.48/170.4	-	2.40/162.6
D1 (Wat1)	Fuc-O1	2.46/144.4	-	-	-
D1 (Wat1)	Fuc-O2	2.43/142.2	-	2.54/143.0	-
D2 (Wat1)	Fuc-O1	-	-	2.56/125.6	-
Fuc-OD1	O (Wat2)	2.24/160.8	-	-	-
D2 (Wat2)	Ser23.OG	1.87/149.5	-	-	-
D1 (Wat2)	Fuc-O1	-	-	-	2.50/147.9
D2 (Wat2)	Ser23.OG	-	-	-	2.40/160.7

[a] C-terminal residue from the neighbouring monomer

In the neutron structure, Wat1 accepts a hydrogen bond from the amide backbone of Thr98 (2.07 Å +/- 0.05 Å) in all monomers. In chain A, Wat1 donates one hydrogen to FucO1/FucO2 and the second hydrogen to the adjacent water molecule (see Figure 6a). In chains B and D, Wat1 accepts a hydrogen bond from Fuc-OD1 while it donates one hydrogen to the adjacent water molecule (see Figure 6b, 6d). In chain C, Wat1 donates both its hydrogens to Fuc-O1 (2.56 Å) and Fuc-O2 (2.54 Å) atoms, respectively, while Fuc-OD1 points its deuterium towards the solvent. It is interesting that this is the most conserved water molecule in the carbohydrate-binding site present in all LecB/fucose complexes (PDB: 1GZT, 1OXC, 1UZV) as well as in the native structure (PDB: 1OUX) and that

it might stabilize the calcium-binding loop composed of 14 amino acids in a β-hairpin motif as previously suggested[19]. Mitchell *et al*. tried to deduce the orientation of this water molecule assuming that it accepts hydrogen bonds from NH of Thr98 and Asp99. They suggested that the water donates both its hydrogens to O2 and O1 hydroxyls of fucose, respectively. Here, we show that this water can adopt different bonding orientations in the four chains according to the presence of other water molecules in the vicinity of O1 and Wat1 that are not conserved in all binding sites.
The second water molecule Wat2 bridges the anomeric oxygen atom to the side chain of Ser23 and is not conserved in all LecB crystal structures. In the neutron structure, Wat2 is present in two out of the four

monomers. The neutron density is less clear for Wat2 but it could still be oriented based on the 2mF_o-DF_c neutron map and the orientation of both OD1 fucose hydroxyl group and the side chain of Ser23.
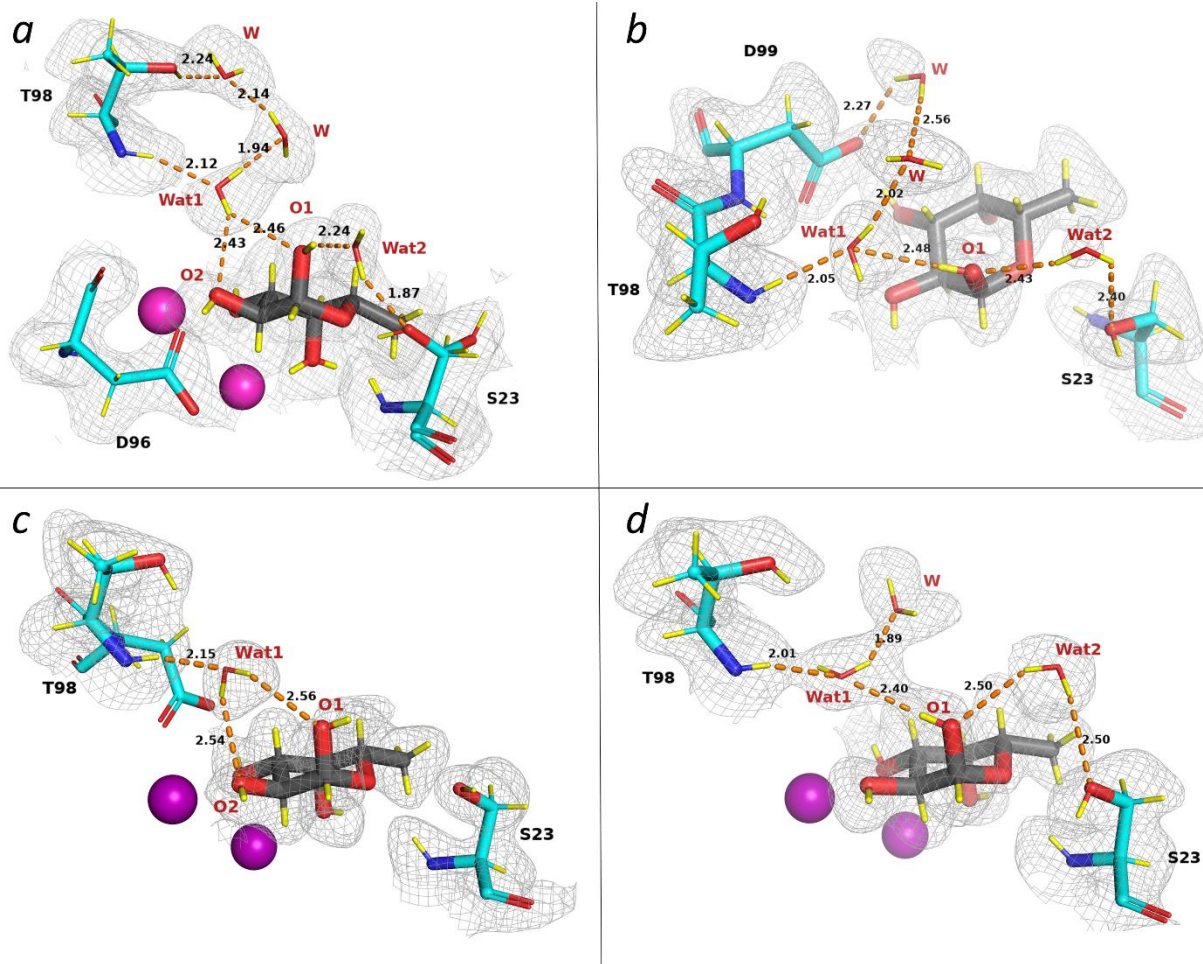
Figure 6: Stick representation of the fucose-binding sites in chains A, B, C and D of the perdeuterated LecB/fucose tetrameric complex. 2mF_o-DF_c neutron density map (grey mesh) is contoured at 0.7σ. Hydrogen bonds are shown as orange dashed lines with distances in Å. Fucose is depicted as grey sticks. Deuterium atoms are coloured yellow. Calcium ions are shown as purple spheres.

Comparison between 100K and room temperature structure

The 100K and the room temperature structures are both highly similar with RMSD for the backbone atoms of the tetramer of 0.3 Å. The 100K structure displays lower crystallographic B-factors for the protein and ligands with an average of 11.1 Å^2 and 12.8 Å^2, respectively, compared to the room temperature X-ray/neutron structure with values of 19.3 Å^2 and 17.6 Å^2, respectively. Despite the differences in the temperature and resolution, the static and dynamic disorder of both crystals is comparable, with 39 and 32 amino acids refined with two or more conformations in the 100K and room temperature structure, respectively. A total number of 364 water molecules were modelled in the room temperature neutron

structure, which is less than 50% of the water molecules observed in the 0.9 Å 100K X-ray crystal structure (Figure 7).
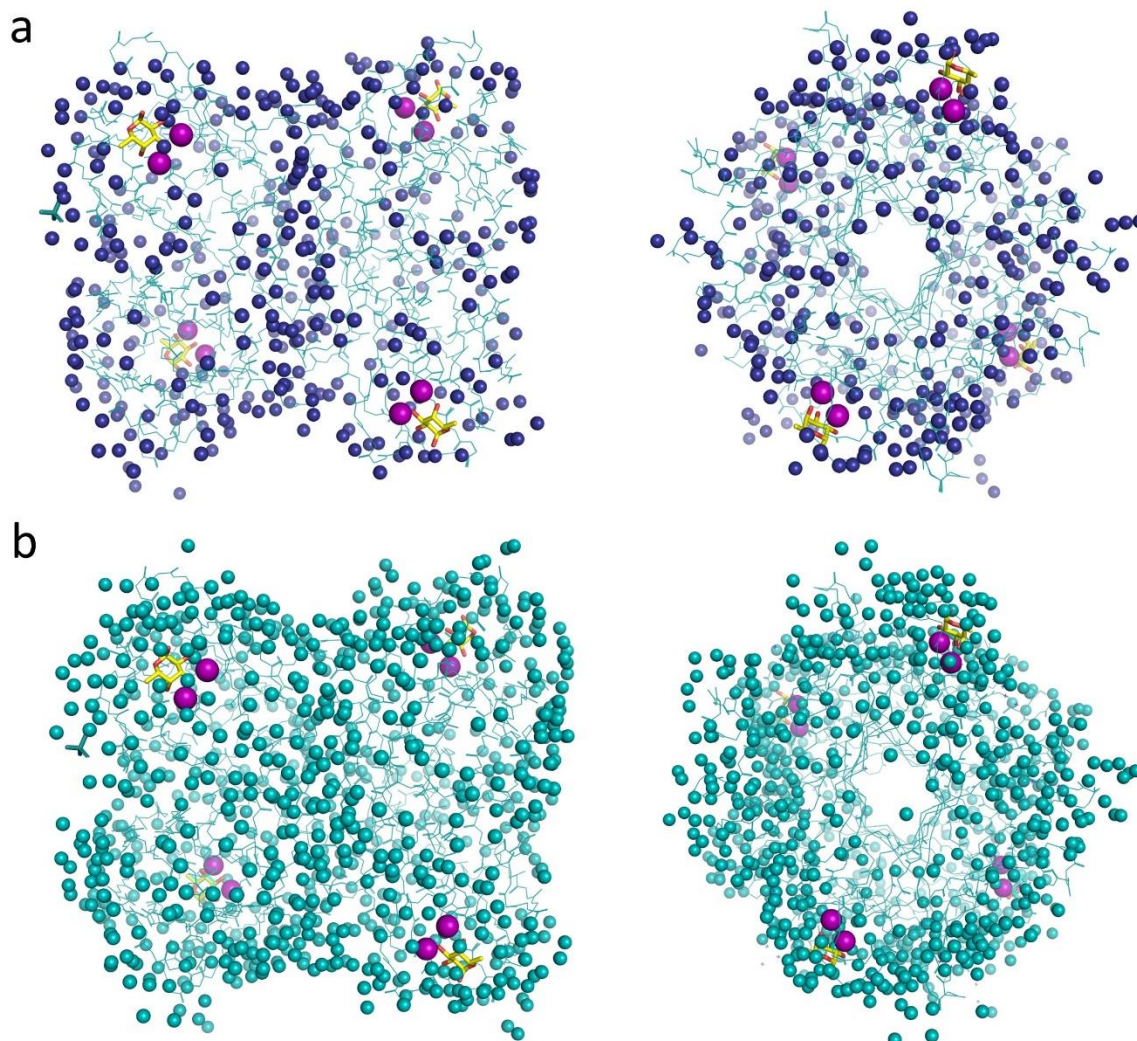
Figure 7: Front and side orientation of the perdeuterated LecB tetramer in complex with perdeuterated fucose. Fucose is shown as yellow sticks and calcium atoms as purple spheres. (Top) Schematic illustration of water molecules bound in the room temperature structure (dark blue spheres) and (bottom) in 100K (cyan spheres) and of the perdeuterated LecB/fucose complex.

In the fucose-binding site, the 100K X-ray structure allowed the positions of all aliphatic deuterium atoms in the fucose molecule to be determined based on the mF_o-DF_c omit electron density map, including the fucose methyl group with a strong triangular-shaped positive peaks in the omit electron density contoured at 2.5 σ

(Figure 8). The quality of the map allowed the orientation of two hydroxyl groups OH2 and OH4 pointing towards OD1 oxygen of Asp96 side chain and OXT oxygen atom of Gly114 of the neighbouring monomer, respectively. This is in agreement with the previously determined high-resolution structure of the LecB/fucose complex

determined at 1 Å, although only one hydrogen on OH2 hydroxyl group could be located in that structure then[19]. It is also in agreement with the neutron structure, which allows for an exact positioning and orientation of all hydrogen atoms as described above.
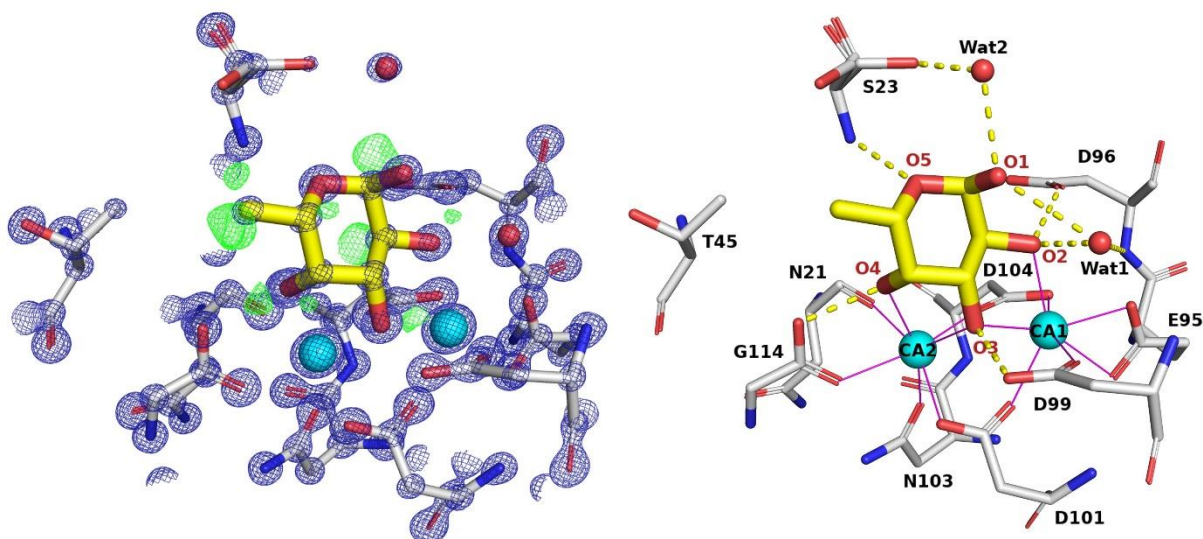
Figure 8: Stick representation of the 0.9 Å 100K X-ray structure of the fucose-binding site of the perdeuterated LecB/Fuc-d_{12} complex. Fucose is shown as thick yellow sticks and the protein as thin grey sticks. $2mF_o$-DF_c electron density (blue mesh) is contoured at 4σ. The mF_o-DF_c omit electron density (green mesh) is contoured at 2.5σ. Hydrogen bonds are shown as yellow dashed lines and distances are in Å. The metal coordination is represented by purple solid lines. The calcium ions are shown as cyan spheres.

Two conserved water molecules Wat1 and Wat2 are present in all monomers in the 0.9 Å X-ray crystal structure, which is different from the room temperature neutron structure where Wat2 bridging the anomeric oxygen O1 to the side chain of Ser23 is present in only two monomers out of the four. This is possibly due to the lower resolution and the temperature used for the neutron data collection. Comparably, in the 1 Å X-ray structure of the LecB/fucose complex (PDB: 1UZV), Wat2 is absent in one monomer and in one it splits to two close sites (refined at 0.5 occupancy each).

Comparison between deuterated and hydrogenated structures in cryo conditions

The perdeuterated structures of LecB complexed with fucose can be directly compared to the previously reported hydrogenated LecB/Fuc complexes[17] since the crystals were obtained using similar crystallization conditions. The X-ray data were collected to very high resolution in both cases. The structures are very similar with an RMSD of 0.13 Å over the whole tetramer main chain. Both structures contain two sulphate ions, originating from the crystallization conditions, bound on the same positions near chains A and B. Differences were limited to the loop connecting β-strands 5 and 6 of the monomer, not far from the binding site. The calcium and fucose-binding sites were also very similar.

161
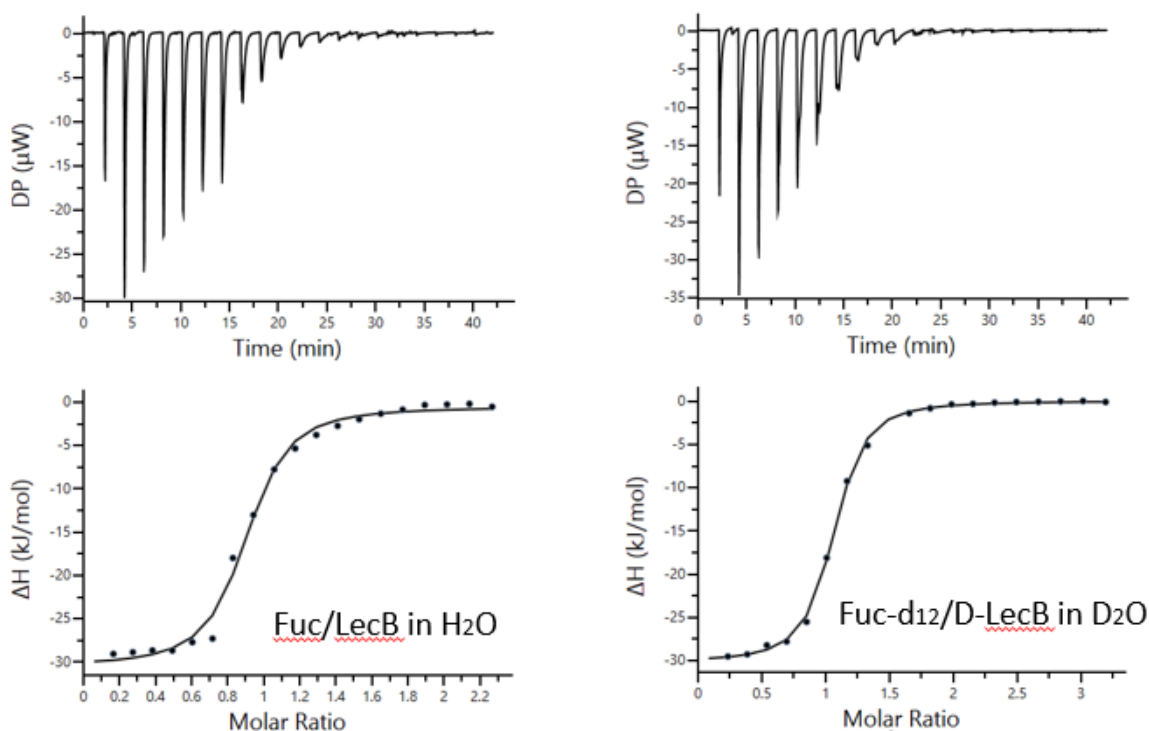
Figure 9: Isothermal titration calorimetry results of L-Fuc/LecB and L-Fuc-d$_{12}$/D-LecB binding in 20 mM Tris buffer, pH 7.5 with 100 mM NaCl and 100 µM CaCl$_2$ prepared in H$_2$O and D$_2$O, respectively. Top: data obtained from 20 automatic injections, 2 µL each, of L-fuc (4 mM)/L-Fuc-d$_{12}$ (7.6 mM) into the cell containing LecB (0.46 mM)/D-LecB (0.50 mM). Lower: plot of the total heat released as a function of total ligand concentration for the titration shown above. The solid line represents the best least-squares fit for the obtained data.

The thermodynamics of fucose binding to LecB was also compared between the fully hydrogenated and fully deuterated systems, in H$_2$O and D$_2$O, respectively (Figure 9). No significant differences could be observed with an enthalpy of binding ΔH of -30.45 (+/-0.05) kJ/mol for the hydrogenated system and -30.95 (+/-0.15) kJ/mol for the deuterated one. The dissociation constants were also almost identical with Kd values of 7.64 µM (+/-0.15) and 7.96 µM (+/- 0.42), respectively. Assays with mixing molecules (deuterated protein with hydrogenated ligand, or vice versa – data not shown) were also carried out and yielded similar results. This confirms that, for this particular protein/carbohydrate complex, the deuterated system does not diverge from the natural one, and that the results obtained

for the visualization and the positioning of the deuterium atoms can be safely extrapolated to hydrogen atoms.

Conclusion

This is the first record of an experimental determination of the directionality of the fucose hydroxyl groups and the protonation state of acidic residues in the carbohydrate-binding site of LecB lectin from human pathogen *Pseudomonas aeruginosa*. Previously, from the high-resolution 100K X-ray crystal structure (PDB: 1UZV) to 1 Å resolution only one hydrogen atom of O2-hydroxyl group could be oriented, pointing towards oxygen OD1 of the Asp96 side chain[19]. The protonation state of the acidic residues could not be unambiguously determined until now. The neutron analysis provided evidence that all of the six acidic

162

residues coordinating the calcium atoms are non-protonated. This is in line with the prediction based on the *ab initio* calculations of bond distances in acidic residues with different protonation states.

We show here that high-resolution X-ray crystallography can provide the locations of some hydrogen atoms involved in the sugar recognition, but that accurate determination of hydrogen positions and directionality of water molecules could only be unambiguously located by neutron crystallography. This information can be important in the rational structure-based drug design of novel potent inhibitors of *Pseudomonas aeruginosa*.

Materials and methods

Protein expression

The LecB lectin was expressed in *Escherichia coli* BL21(DE3) bacteria harbouring a pET29b(+)-pa2l plasmid with a kanamycin-resistance gene. All cultures were grown at 37 °C with shaking at 180 rpm and were supplemented with 50 μg mL^{-1} kanamycin.

Adaptation to D_2O and deuterated glycerol-d_8

Precultures of the LecB-producing strain were first grown in LB medium. Cells were then adapted to the deuterated Enfors minimal medium with the following composition: 6.86 g L^{-1} (NH$_4$)$_2$SO$_4$, 1.56 g L^{-1} KH$_2$PO$_4$, 6.48 g L^{-1} Na$_2$HPO$_4$·2H$_2$O, 0.49 g L^{-1} (NH$_4$)$_2$HC$_6$H$_5$O$_7$ (diammonium hydrogen citrate), 0.25 g L^{-1} MgSO$_4$·7H$_2$O, with 1.0 mL L^{-1} of trace metal stock solution (0.5 g L^{-1} CaCl$_2$·2H$_2$O, 16.7 g L^{-1} FeCl$_3$·6H$_2$O, 0.18 g L^{-1} ZnSO$_4$·7H$_2$O, 0.16 g L^{-1} CuSO$_4$·5H$_2$O, 0.15 g L^{-1} MnSO$_4$·4H$_2$O, 0.18 g L^{-1} CoCl$_2$·6H$_2$O, 20.1 g L^{-1} EDTA), 5 g L^{-1} glycerol-d_8 (Eurisotop). A single colony of *E. coli* containing the pET29b(+)-pa2l plasmid grown on an LB agar plate supplemented with 50 μg mL^{-1} kanamycin was used to inoculate 15 mL of LB medium with kanamycin. The preculture was then used to inoculate 15 mL of the hydrogenated Enfors minimal medium to OD$_{600}$ of 0.1 and grown overnight. This preculture was used to inoculate 15 mL of the deuterated Enfors minimal medium to OD$_{600}$ of 0.1 and grown overnight. This step was repeated 5 times prior to the fed-batch fermentation.

Deuterated fed-batch fermentation of LecB

A final preculture of 150 mL was used to inoculate 1.2 L of deuterated Enfors minimal medium in a 3 L bioreactor (Infors, Switzerland). The pD of the culture medium was regulated at 7.2 by addition of 4% NaOD (Eurisotop, France). The temperature was maintained at 30 °C. After consumption of the deuterated glycerol-d_8 from the culture medium, the fed-batch phase was initiated by continuous feeding with the additional 30 g of the deuterated glycerol-d_8. The protein expression was induced overnight (17 h) with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at the OD$_{600}$ of 13 at 30 °C. After the fermentation, the cells were recovered by centrifugation (10 500 g for 1 h at 6 °C) and the wet cell paste was frozen at -80 °C for long-term storage. The final yield was 51 g of the deuterated cell paste from a final volume of 1.6 L culture medium.

Protein purification

Deuterated LecB was purified by affinity chromatography using the NGC system (Bio-Rad, Marnes-la-Coquette, France) on a 10 mL mannose-Sepharose resin packed in the C10/10 column. All buffers used for purification were prepared in H$_2$O. Cells were resuspended in buffer A (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 100 μM CaCl$_2$) in the presence of an EDTA-free protease inhibitor cocktail (cOmplete™, Roche) and treated with DENARASE® endonuclease (c-LEcta GMBH, Leipzig, Germany). Cells were lyzed using a cell disruptor with pressure of 1.8 kbar (Constant Systems Ltd., Northants UK). After centrifugation (24 000 g for 30 min at 4 °C), the

supernatant was filtered (0.45 μm) and the clear cell lysate was loaded onto 10 mL mannose-agarose column pre-equilibrated with the buffer A. After extensive washing the unbound proteins, the deuterated LecB was eluted with buffer A containing 100 mM free D-mannose. The purity of the protein was examined by 16% Tris-tricine SDS-PAGE stained with Coomassie Blue. The fractions containing the pure protein were pooled together and dialyzed against 5 L of buffer A at 4 °C for a week with changing the buffer once per day. The protein was concentrated using a 5 kDa cut-off concentrator (Corning® Spin-X® UF, England) and flash-frozen in liquid nitrogen for long term storage. The typical yield of the deuterated LecB was about 4 mg of protein per gram of wet cell paste.

Isothermal titration calorimetry
All measurements were carried out at 25 °C using a MicroCal iTC200 isothermal titration calorimeter (Microcal-Malvern Panalytical, Orsay, France). A total of 20 injections of 2 μL of fucose solution each were injected at intervals of 120 s with constant stirring of 750 rev min^{-1} in the 200 μL sample cell containing the LecB protein. Experimental data were fitted to the theoretical titration curve with the one set of sites model in the software supplied by MicroCal with ΔH (enthalpy change), Ka (association constant) and n (number of binding sites per monomer) as adjustable variables. The free energy change (ΔG) and entropy contributions (TΔS) were calculated from the equation ΔG= ΔH – TΔS = –RT ln K_a (T is the absolute temperature and R = 8.314 J K^{-1} mol^{-1}). Two independent titrations were performed for each experiment.

Crystallization
The perdeuterated LecB lectin was crystallized using a vapour-diffusion sitting drop method. The protein solution of 10 mg mL^{-1} was incubated with 250 μg mL^{-1} of perdeuterated fucose in the presence of 100 μM CaCl$_2$ during 1 h at room temperature prior to crystallization. Perdeuterated fucose-d$_{12}$ was produced by glyco-enginereed *E. coli* in a high cell-density culture as reported previously[20]. Co-crystals of deuterated LecB with L-fucose-d$_{12}$ were obtained in the following conditions: 0.1 M Tris/DCl, pD 7.1, 20% (w/v) PEG 4000. The protein was mixed with the reservoir solution in 1:1 ratio and incubated at 19 °C.

Neutron data collection and processing
Neutron quasi-Laue diffraction data from the crystal of perdeuterated LecB/fucose complex were collected at room temperature using the LADI-III diffractometer[21] at the Institut Laue-Langevin (ILL) in Grenoble using a crystal of perdeuterated LecB in complex with perdeuterated fucose-d$_{12}$ with volume of ∼ 0.1 mm^3. A neutron wavelength range (Δλ/λ=30 %) of 2.8-3.8 Å was used for data collection with diffraction data extending to 1.9 Å resolution. The crystal was held stationary at different φ (vertical rotation axis) for each exposure. A total of 24 images were recorded (18 h per exposure) from 2 different crystal orientations. The neutron diffraction images were processed using the *LAUEGEN* software[22]. The *LSCALE* program[23] was used to determine the wavelength normalization curve using intensities of symmetry-equivalent reflections measured at different wavelengths. The data were merged and scaled using *SCALA*[24].

X-ray data collection and processing
Room temperature X-ray data were collected from the same crystal used for the neutron data collection. The X-ray data were recorded on the GeniX 3D Cu High Flux diffractometer (Xenocs) at the Institut de Biologie Structurale (IBS) in Grenoble, France. The data were processed using the *iMosflm* software[25], scaled and merged with *AIMLESS*[26] and converted to structure

factors using *TRUNCATE* in the *CCP4* suite[27].

Other crystals of deuterated LecB co-crystallized with deuterated fucose were harvested manually and cryocooled at 100K by flash-freezing in liquid nitrogen. X-ray datasets were collected at the SOLEIL synchrotron (Saint Aubin, France) on the PROXIMA-1beamline. Images were recorded on the EIGER-X 16M detector (Dectris Ltd., Baden, Switzerland) and processed by *XDS*[28].

Joint neutron and X-ray refinement

The initial model (PDB: 1GZT) with water molecules, metal ions and ligands removed was used for the molecular replacement using *PHASER* in *PHENIX* suite. All further data analysis was done using the *PHENIX* package[29]. Refinement was performed using a restrained maximum-likelihood refinement in *phenix.refine*[30]. Restraint files for deuterated fucose were generated by *eLBOW* utility[31] in *PHENIX*. Ligands were placed manually in *COOT*[32]. Water molecules were introduced automatically in *phenix.refine* based on the positive peaks in the mF_o-DF_c neutron scattering length density and inspected manually. Deuterium atoms were introduced using the *Readyset* utility in *Phenix* and refined individually. All water molecules were modelled as D_2O. Water oxygen atoms were initially placed according to the electron density peaks. The orientation of water molecules were refined and modified based on the potential hydrogen donor and acceptor orientation.

The quasi-Laue neutron R_{work} and R_{free} values for the final model were 19.2% and 24.6 %, respectively, while the X-ray R_{work} and R_{free} values were 10.4% and 14.2%, respectively. *Molprobity* software was used for structure validation. Refinement statistics are presented in Table 2.

Table 2: Room temperature neutron data collection and refinement statistics for the perdeuterated LecB/fucose complex

	D-LecB Fucose-d_{12}	D-LecB Fucose-d_{12}
Data collection		
Temperature	RT	100K
Neutrons		
Instrument	LADI-III, ILL	
Wavelengths (Å)	2.8-3.8	
Detector	Image plate	
Resolution (Å)	43- 1.90 (2.08-1.90)	
Spacegroup	$P2_1$	
Unit cell parameters		
a, b, c (Å)	52.9, 73.9, 55.0	
α, β, γ (°)	90, 94.6, 90	
R_{merge} (I) (%)	21.1 (38.5)	
R_{pim} (I) (%)	10.9 (17.0)	
Mean I/σ (I)	5.5 (2.1)	
Completeness (%)	73.8 (61.7)	
Multiplicity	3.7 (3.9)	
No. of unique reflections	24 478 (4860)	
Crystal size (mm³)	0.1	
X-rays		
X-ray source	GeniX 3D Cu High Flux (Xenocs), IBS	Proxima-1, SOLEIL
Wavelength (Å)	1.5418	0.97856
Detector	Mar 345 (marXperts)	EIGER-X 16M (Dectris Ltd.)
Resolution (Å)	33-1.85 (1.89-1.85)	40-0.90 (0.92-0.90)
Unit cell parameters		Ffff
a, b, c (Å)	52.9, 73.9, 55.0	52.6, 73.0, 55.2
α, β, γ (°)	90, 94.6, 90	90, 94.6, 90
R_{merge} (I) (%)	9.7 (29.8)	4.3 (141.6)
R_{meas} (I) (%)	10.4 (36.2)	4.7 (171.6)
CC1/2 (%)	99.5 (84.7)	100 (63.2)
Mean I/σ (I)	11.3 (4.2)	16.3 (0.6)
Completeness (%)	96.6 (70.4)	87.6 (25.6)
Multiplicity	7.0 (2.5)	6.1 (2.4)
No. of unique reflections	34855 (1586)	270061 (5821)
Refinement		
Resolution range X-ray (Å)	27.41-1.85	36.55-0.90
Resolution range neutron (Å)	42.92-1.90	
Reflections (used), X-ray	34734	526849
Reflections (test), X-ray	1736	25914
Reflections (used), neutron	24283	
Reflections (test), neutron	1195	
R_{work} (%), X-ray	10.4	11.3
R_{free} (%), X-ray	14.2	12.8
R_{work} (%), neutron	19.2	
R_{free} (%), neutron	24.6	
No. of atoms (protein)	3569	3609
Rmsd in bond lengths (Å)	0.010	0.08
Rmsd in bond angles (°)	1.305	1.072

Average B factors (Å^2)		
Protein	19.3	11.1
Ligand	17.6	12.7
Ramachandran statistics		
Favoured (%)	97.3	96.2
Allowed (%)	2.7	3.8
Outliers (%)	0.0	0.0
All-atom clashscore	2.0	5.9
PDB code		

References

1. Noach, I. et al. Recognition of protein-linked glycans as a determinant of peptidase activity. Proc. Natl. Acad. Sci. U. S. A. 114, E679–E688 (2017).

2. Tommasone, S. et al. The challenges of glycan recognition with natural and artificial receptors. Chemical Society Reviews 48, 5488–5505 (2019).

3. Lis, H. & Sharon, N. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. 98, 637–674 (1998).

4. Gregg, K. J., Finn, R., Abbott, D. W. & Boraston, A. B. Divergent modes of glycan recognition by a new family of carbohydrate-binding modules. J. Biol. Chem. 283, 12604–12613 (2008).

5. Ambrosi, M., Cameron, N. R. & Davis, B. G. Lectins: Tools for the molecular understanding of the glycocode. Organic and Biomolecular Chemistry 3, 1593–1608 (2005).

6. Kingeter, L. M. & Lin, X. C-type lectin receptor-induced NF-κB activation in innate immune and inflammatory responses. Cellular and Molecular Immunology 9, 105–112 (2012).

7. Imberty, A. & Varrot, A. Microbial recognition of human cell surface glycoconjugates. Curr. Opin. Struct. Biol. 18, 567–576 (2008).

8. Cecioni, S., Imberty, A. & Vidal, S. Glycomimetics versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands. Chem. Rev. 115, 525–561 (2015).

9. Sharon, N. Carbohydrates as future anti-adhesion drugs for infectious diseases. Biochimica et Biophysica Acta - General Subjects 1760, 527–537 (2006).

10. Mousavifar, L., Touaibia, M. & Roy, R. Development of Mannopyranoside Therapeutics against Adherent-Invasive Escherichia coli Infections. Acc. Chem. Res. 51, 2937–2948 (2018).

11. Bonnardel, F. et al. Unilectin3d, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. Nucleic Acids Res. 47, D1236–D1244 (2019).

12. Tacconelli, E. & Magrini, N. WHO | Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. WHO (2017).

13. Cioci, G. et al. Structural basis of calcium and galactose recognition by the lectin PA-IL of *Pseudomonas aeruginosa*. FEBS Lett. 555, 297–301 (2003).

14. Mitchell, E. et al. Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients. Nat. Struct. Biol. 9, 918–921 (2002).

15. Chemani, C. et al. Role of LecA and LecB Lectins in *Pseudomonas aeruginosa-*Induced Lung Injury and Effect of Carbohydrate Ligands. Infect. Immun. 77, 2065–2075 (2009).

16. Mitchell, E. et al. Structural basis for oligosaccharide-mediated adhesion of Pseudomonas aeruginosa in the lungs of cystic

fibrosis patients. Nat. Struct. Biol. 9, 918–921 (2002).

17.   Mitchell, E. P. et al. High affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins Struct. Funct. Bioinforma. 58, 735–746 (2004).

18.   Loris, R., Tielker, D., Jaeger, K. E. & Wyns, L. Structural basis of carbohydrate recognition by the lectin LecB from *Pseudomonas aeruginosa*. J. Mol. Biol. 331, 861–870 (2003).

19.   Mitchell, E. P. et al. High affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins Struct. Funct. Bioinforma. 58, 735–746 (2004).

20.   Gajdos, L. et al. Production of perdeuterated fucose from glyco-engineered bacteria.   Glycobiology   (2020). doi:10.1093/glycob/cwaa059

21.   Blakeley, M. P. et al. Neutron macromolecular crystallography with LADI-III. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 1198–1205 (2010).

22.   Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. LAUEGEN version 6.0 and INTLDM. J. Appl. Crystallogr. 31, 496–502 (1998).

23.   Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. LSCALE - The new normalization, scaling and absorption correction program in the Daresbury Laue software suite. J. Appl. Crystallogr. 32, 554–562 (1999).

24.   Evans, P. Scaling and assessment of data quality. in Acta Crystallographica Section

D: Biological Crystallography 62, 72–82 (International Union of Crystallography, 2006).

25.   Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 271–281 (2011).

26.   Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1204–1214 (2013).

27.   Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallographica Section D: Biological Crystallography 67, 235–242 (2011).

28.   Kabsch, W. et al. XDS. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 125–132 (2010).

29.   Adams, P. D. et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 213–221 (2010).

30.   Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 352–367 (2012).

31.   Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. Acta Crystallogr. Sect. D Biol. Crystallogr. 65, 1074–1080 (2009).

32.   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501 (2010).

6
Conclusion and perspectives

Carbohydrates present on the cell surface in the form of glycoproteins and glycolipids act as binding epitopes and carriers of biological information. Sugar-binding proteins such as lectins are widely expressed in nature and can specifically recognize various mono- and oligosaccharides. Lectin-carbohydrate interactions have been shown to be implicated in various important processes in health involving cell-trafficking, immune response and also in diseases including cell-adhesion of infectious agents such as bacteria and viruses to host tissues. The interactions have been extensively studied at the molecular level using X-ray crystallography, which helped to describe the molecular basis of sugar recognition. These interactions are predominantly mediated by hydrogen bonds and Van der Waals interactions, they sometimes involve water-bridging interactions, stacking interactions between the apolar face of a sugar molecule and aromatic amino acid residues, and metal ion coordination.

In this thesis, neutron protein crystallography (NMX) has been used to describe the interactions between lectins from pathogenic bacteria and their carbohydrate ligand. NMX is a powerful technique that is complementary to X-ray crystallography and provides further details on hydrogen atom positions. Direct visualization of hydrogen atoms is of crucial importance to better describe positions of sugar hydroxyl groups, define the protonation state of charged amino acids and determine the directionality of bonding to water molecules involved in the carbohydrate recognition. In this thesis, a series of perdeuteration experiments has been carried out together with single-crystal neutron diffraction studies of selected bacterial lectins. The first part of the thesis was dedicated to the production of perdeuterated fucose using a unique *in vivo* approach. Perdeuterated fucose has been then used as ligand for crystallization of

bacterial lectins, PLL from *Photorhabdus laumondii* and LecB from *Pseudomonas aeruginosa.*

In Chapter 3, the production of perdeuterated L-fucose (Fuc-d_{12}) using glyco-engineered *Escherichia coli* is described. The genetically modified strain of *E. coli* bacteria was designed by Dr. Eric Samain at the CERMAV institute in Grenoble. The strain was modified with the insertion and deletion of genes required for the biosynthesis of fucose. At ILL, the strain was adapted to D_2O-based minimal medium with perdeuterated glycerol-d_8 as carbon source. A series of fermentation experiments was carried out in the hydrogenated and deuterated media to optimize the production of fucose.

A simple batch fermentation was performed using 1.5-L of deuterated medium with 45 g of glycerol-d_8. A comparison between the hydrogenated and perdeuterated cultures showed that fucose is produced in the exponential phase of the bacterial growth in both cultures. The production was about 10 times more efficient in the hydrogenated medium than in the deuterated one. The final yield of the product was estimated to be about 2.1 mg L^{-1} in hydrogenated fermentation and 0.2 mg L^{-1} in the deuterated one. Lower yields of recombinant proteins and other biomolecules such as cholesterol produced by engineered microorganisms have been observed in perdeuterated cultivations (Moulin et al., 2018). Nevertheless, results have shown that *in vivo* production of fucose using glyco-engineered bacteria can yield high amounts of product and can be suitable for large-scale production of carbohydrates.

The product was characterized by mass spectrometry (MS) and nuclear magnetic resonance (NMR) experiments. The MS analysis showed that the molecular weight of perdeuterated fucose corresponds to its theoretical molecular weight indicated by the molecular ion adduct peaks observed in the mass spectra, which are in agreement with the calculated mass. In order to examine the level of perdeuteration,

both the proton and deuterium NMR experiments were carried out, both confirming full deuteration on all carbon atoms of the sugar ring.

Metabolic engineering of bacterial and yeast strains has been widely used in biotechnology and the production of metabolites with commercial interest. Results have shown that the production of perdeuterated monosaccharides can be efficiently done *in vivo* using engineered *E. coli*. In collaboration with CERMAV, several other carbohydrates of interest have been selected that could be produced in a deuterated form using a similar approach. Two *E. coli* strains for production of galactose and of a human blood group H type II trisaccharide have been designed by Dr. Samain and are ready for adaptation to D_2O and expression tests. Perdeuterated galactose could be used as a ligand for neutron studies of galactose-specific lectins such as LecA from *Pseudomonas aeruginosa* and the human fucosylated oligosaccharide could be used as a natural ligand for fucose-specific lectins including LecB from *P. aeruginosa* and BambL lectin from human pathogen *Burkholderia ambifaria*.

The results obtained in this study have been described in the published article "Production of perdeuterated fucose from glyco-engineered bacteria" published in the journal Glycobiology in 2021.

In Chapter 4, a series of crystallographic studies including X-ray and neutron single-crystal diffraction experiments was carried out on PLL lectin from bacteria *Photorhabdus laumondii*. *P. laumondii* is an insect pathogen that lives in a symbiotic association with nematodes. Recently, several lectins from this bacterium have been structurally and biophysically characterized. PLL is a fucose-specific lectin that binds fucose via direct and water-bridged hydrogen bonds as well as via CH-π interactions with tryptophan side chains. Perdeuterated PLL was produced using an *E. coli* high cell-density culture in a deuterated minimal medium with glycerol-d$_8$ as a carbon source.

Both the hydrogenated and the deuterated proteins were crystallized in the apo form and also co-crystallized with hydrogenated and perdeuterated fucose. Two room-

temperature jointly refined structures against X-ray and neutron data have been solved. The neutron structure of H/D-exchanged apo PLL allowed the observation of continuous neutron density for three water molecules bound in the binding site at the positions of fucose hydroxyl and carbon atoms. These waters displayed high crystallographic B-factors and a limited hydrogen-bonding network to the protein.

The neutron structure of the perdeuterated PLL/fucose complex provided a neutron map of excellent quality due to the use of perdeuterated protein and ligand. The positions of all hydrogen atoms involved in sugar recognition, consisting of direct hydrogen bonds and stacking interactions with aromatic amino acids, could be described. A direct observation and description of the CD-π interactions between fucose aliphatic deuterium atoms and aromatic rings of tryptophan residues provided further details into the sugar-recognition pattern.

Furthermore, an additional fourth fucose-binding site in PLL lectin was occupied by a fucose molecule in the 100 K and room temperature structures (RT). When comparing the 100 K and RT X-ray structures, it has been observed that glycerol readily binds to protein and thus altering conformations of several loops. Room temperature structures, thus, could allow confident description of the protein that was not affected by artefacts caused by cryo-protecting and cryo-cooling.

A possible perspective of this study could be the use of a higher concentration of fucose during co-crystallization experiments in order to improve the neutron scattering length density for fucose in all four binding sites. Another perspective could be to carry out a cryo neutron diffraction experiment on apo D-PLL. As described previously for Concanavalin A lectin, a 15 K neutron data collection reduced B-factors for the water molecules in the binding site resulting in a better quality of the neutron density map. This allowed a confident placement of D atoms and the orientation of water molecules (Blakeley et al., 2004).

The work carried out in this thesis resulted in an accepted article "Visualization of hydrogen atoms in a perdeuterated lectin-fucose complex reveals key details of protein-carbohydrate interactions" in the journal Structure, 2021.

In Chapter 5, a neutron diffraction study has been carried out on LecB lectin from bacteria *Pseudomonas aeruginosa*. *P. aeruginosa* is a human opportunistic pathogen that causes fatal lung infections in cystic fibrosis patients. LecB is a fucose-specific lectin that acts as a virulence factor involved in bacterial adhesion and biofilm formation and that is viewed as a potential target for anti-adhesive therapy of *P. aeruginosa* infection. LecB displays a high affinity towards fucose and fucosylated oligosaccharides. Two calcium ions are involved in the carbohydrate-binding site that is characterized by six acidic residues coordinating both the calcium ions and fucose. NMX was used to provide insights into the hydrogen-bonding network and protonation state of the charged amino acid residues involved in the sugar and calcium-binding site.

In this thesis, perdeuterated LecB was successfully produced using a high cell-density culture. The protein was co-crystallized with perdeuterated fucose and large crystals were grown using macro-seeding and feeding techniques applied in the vapour-diffusion sitting-drop set up. A room temperature neutron diffraction data set was collected from a ~ 0.1 mm^3 crystal of the D-LecB/Fuc-d$_{12}$ complex and the structure was jointly refined against both neutron and X-ray data collected from the same crystal to 1.90 and 1.85 Å resolution respectively. The analysis of the structure allowed clear observation of four direct hydrogen bonds established between the fucose and the protein. Perdeuterated fucose further allowed a detailed analysis of the hydrophobic contacts between fucose methyl group and aliphatic side chains. No neutron density for hydrogen atoms was observed in the vicinity of the acidic residues in the binding site indicating non-protonated state. This was in agreement with the previously reported chemical calculations based on the high-resolution X-ray structure of hydrogenated LecB complexed with commercial fucose (PDB: 1UZV). The neutron analysis also revealed the positions of the hydrogen atoms in water molecules that participate in the binding. The complete description of the hydrogen-bonding network involved in the fucose binding could be unambiguously described.

Since LecB is a drug target for the treatment of *P aeruginosa* infection, structural information is needed to guide a structure-based drug design of novel potent inhibitors with high potencies. The experimentally obtained hydrogen atom positions in the fucose molecule as well as in conserved water molecules revealed by neutron protein crystallography in this study may help in the geometry optimization calculations. A possible continuation of this work could be to collect neutron diffraction data from the apo protein to compare the water network involved in the sugar-binding. Given that high-quality neutron data can be collected from rather small crystals of LecB, it would also be interesting to crystallize LecB with high-affinity glycomimetic compounds that showed therapeutical potential (Sommer et al., 2018).

In this thesis, the feasibility of deciphering lectin-glycan interactions using neutrons has been demonstrated. The biggest challenge in neutron protein crystallography has been for a long time the sample size. It has been shown that high-quality neutron data can be collected from submillimeter crystals typically used for X-ray crystallography. The progress in sample preparation with facilities dedicated to deuterium labelling of molecules opens the field of neutron crystallography to a broad range of structural biologists. Perdeuteration of proteins and small molecules is now almost routinely performed which leads to reduced crystal volumes needed for neutron data collections and enhanced signal-to-noise ratios in the neutron scattering experiments. The current developments in neutron instrumentation, construction of new facilities and commissioning of new instruments dedicated to NMX, will certainly have an impact on the increasing number of deposited neutron protein structures including lectin-glycan complexes in the PDB.

References

Adam, E.C., Mitchell, B.S., Schumacher, D.U., Grant, G., Schumacher, U., 1997. Pseudomonas aeruginosa II lectin stops human ciliary beating: Therapeutic implications of fucose. Am. J. Respir. Crit. Care Med. 155, 2102–2104. https://doi.org/10.1164/ajrccm.155.6.9196121

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 213–221. https://doi.org/10.1107/S0907444909052925

Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R.W., Moriarty, N.W., Langan, P., Adams, P.D., 2010. Joint X-ray and neutron refinement with *phenix.refine*. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 1153–1163. https://doi.org/10.1107/S0907444910026582

Ahmed, H.U., Blakeley, M.P., Cianci, M., Cruickshank, D.W.J., Hubbard, J.A., Helliwell, J.R., 2007. The determination of protonation states in proteins. Acta Crystallogr. Sect. D Biol. Crystallogr. 63, 906–922. https://doi.org/10.1107/S0907444907029976

Aldred, K.J., Kerns, R.J., Osheroff, N., 2014. Mechanism of quinolone action and resistance. Biochemistry. https://doi.org/10.1021/bi5000564

Ambrosi, M., Cameron, N.R., Davis, B.G., 2005. Lectins: Tools for the molecular understanding of the glycocode. Org. Biomol. Chem. https://doi.org/10.1039/b414350g

Arzt, S., Campbell, J.W., Harding, M.M., Hao, Q., Helliwell, J.R., 1999. LSCALE - The new normalization, scaling and absorption correction program in the Daresbury Laue software suite. J. Appl. Crystallogr. 32, 554–562. https://doi.org/10.1107/S0021889898015350

Asensio, J.L., Ardá, A., Cañada, F.J., Jiménez-Barbero, J., 2013. Carbohydrate-aromatic interactions. Acc. Chem. Res. 46, 946–954. https://doi.org/10.1021/ar300024d

Audfray, A., Beldjoudi, M., Breiman, A., Hurbin, A., Boos, I., Unverzagt, C., Bouras, M., Lantuejoul, S., Coll, J.L., Varrot, A., Le Pendu, J., Busser, B., Imberty, A., 2015. A recombinant fungal lectin for labeling truncated glycans on human cancer cells. PLoS One 10. https://doi.org/10.1371/journal.pone.0128190

Bajolet-Laudinat, O., Girod-de Bentzmann, S., Tournier, J.M., Madoulet, C., Plotkowski, M.C., Chippaux, C., Puchelle, E., 1994. Cytotoxicity of Pseudomonas

aeruginosa internal lectin PA-I to respiratory epithelial cells in primary culture. Infect. Immun. 62, 4481–4487. https://doi.org/10.1128/iai.62.10.4481-4487.1994

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G.W., 2011. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 271–281. https://doi.org/10.1107/S0907444910048675

Becker, D.J., Lowe, J.B., 2003. Fucose: biosynthesis and biological function in mammals. Glycobiology 13, 41R-53R. https://doi.org/10.1093/glycob/cwg054

Beshr, G., Sikandar, A., Jemiller, E.M., Klymiuk, N., Hauck, D., Wagner, S., Wolf, E., Koehnke, J., Titz, A., 2017. Photorhabdus luminescens lectin A (PllA): A new probe for detecting -galactoside–terminating glycoconjugates. J. Biol. Chem. 292, 19935–19951. https://doi.org/10.1074/jbc.M117.812792

Bird, J.M., Kimber, S.J., 1984. Oligosaccharides containing fucose linked α(1-3) and α(1-4) to N-acetylglucosamine cause decompaction of mouse morulae. Dev. Biol. 104, 449–460. https://doi.org/10.1016/0012-1606(84)90101-5

Blakeley, M.P., Kalb, A.J., Helliwell, J.R., Myles, D.A.A., 2004. The 15-K neutron structure of saccharide-free concanavalin A. Proc. Natl. Acad. Sci. U. S. A. 101, 16405–10. https://doi.org/10.1073/pnas.0405109101

Blakeley, M.P., Teixeira, S.C.M., Petit-Haertlein, I., Hazemann, I., Mitschler, A., Haertlein, M., Howard, E., Podjarny, A.D., 2010. Neutron macromolecular crystallography with LADI-III. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 1198–1205. https://doi.org/10.1107/S0907444910019797

Blanas, A., Sahasrabudhe, N.M., Rodríguez, E., van Kooyk, Y., van Vliet, S.J., 2018. Fucosylated antigens in cancer: An alliance toward tumor progression, metastasis, and resistance to chemotherapy. Front. Oncol. https://doi.org/10.3389/fonc.2018.00039

Blanchard, B., Imberty, A., Varrot, A., 2014. Secondary sugar binding site identified for LecA lectin from Pseudomonas aeruginosa. Proteins Struct. Funct. Bioinforma. 82, 1060–1065. https://doi.org/10.1002/prot.24430

Bleves, S., Viarre, V., Salacha, R., Michel, G.P.F., Filloux, A., Voulhoux, R., 2010. Protein secretion systems in Pseudomonas aeruginosa: A wealth of pathogenic weapons. Int. J. Med. Microbiol. https://doi.org/10.1016/j.ijmm.2010.08.005

Bohlool, B.B., Schmidt, E.L., 1974. Lectins: A possible basis for specificity in the Rhizobium-legume root nodule symbiosis. Science (80-.). 185, 269–271. https://doi.org/10.1126/science.185.4147.269

Bonnardel, F., Mariethoz, J., Pérez, S., Imberty, A., Lisacek, F., 2021. LectomeXplore, an update of UniLectin for the discovery of carbohydrate-binding proteins based on a new lectin classification. Nucleic Acids Res. 49, D1548–D1554. https://doi.org/10.1093/nar/gkaa1019

Bonnardel, F., Mariethoz, J., Salentin, S., Robin, X., Schroeder, M., Perez, S., Lisacek, F.D.S., Imberty, A., 2019. Unilectin3d, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. Nucleic Acids Res. 47, D1236–D1244. https://doi.org/10.1093/nar/gky832

Boteva, R.N., Bogoeva, V.P., Stoitsova, S.R., 2005. PA-I lectin from Pseudomonas aeruginosa binds acyl homoserine lactones. Biochim. Biophys. Acta - Proteins Proteomics 1747, 143–149. https://doi.org/10.1016/j.bbapap.2004.10.011

Bowen, D.J., Ensign, J.C., 2001. Isolation and Characterization of Intracellular Protein Inclusions Produced by the Entomopathogenic Bacterium Photorhabdus luminescens. Appl. Environ. Microbiol. 67, 4834–4841. https://doi.org/10.1128/AEM.67.10.4834-4841.2001

Boyd, W.C., Shapleigh, E., 1954. Specific precipitating activity of plant agglutinins (lectins). Science (80-.). 119, 419. https://doi.org/10.1126/science.119.3091.419

Bruckner, J., 1955. Estimation of monosaccharides by the orcinol-sulphuric acid reaction. Biochem. J. 60, 200–205. https://doi.org/10.1042/bj0600200

Bucior, I., Abbott, J., Song, Y., Matthay, M.A., Engel, J.N., 2013. Sugar administration is an effective adjunctive therapy in the treatment of Pseudomonas aeruginosa pneumonia. Am. J. Physiol. - Lung Cell. Mol. Physiol. 305, L352. https://doi.org/10.1152/ajplung.00387.2012

Bucior, I., Pielage, J.F., Engel, J.N., 2012. Pseudomonas aeruginosa Pili and Flagella mediate distinct binding and signaling events at the apical and basolateral surface of airway epithelium. PLoS Pathog. 8. https://doi.org/10.1371/journal.ppat.1002616

Burton, R.A., Gidley, M.J., Fincher, G.B., 2010. Heterogeneity in the chemistry, structure and function of plant cell walls. Nat. Chem. Biol. https://doi.org/10.1038/nchembio.439

Cabanettes, A., Perkams, L., Spies, C., Unverzagt, C., Varrot, A., 2018. Recognition of Complex Core-Fucosylated N-Glycans by a Mini Lectin. Angew. Chemie - Int. Ed. 57, 10178–10181. https://doi.org/10.1002/anie.201805165

Caldararu, O., Manzoni, F., Oksanen, E., Logan, D.T., Ryde, U., 2019. Refinement of protein structures using a combination of quantum-mechanical calculations with neutron and X-ray crystallographic data. Acta Crystallogr. Sect. D Struct. Biol. 75, 368–380. https://doi.org/10.1107/S205979831900175X

Campbell, J.W., Hao, Q., Harding, M.M., Nguti, N.D., Wilkinson, C., 1998. LAUEGEN version 6.0 and INTLDM. J. Appl. Crystallogr. 31, 496–502. https://doi.org/10.1107/S0021889897016683

Cecioni, S., Imberty, A., Vidal, S., 2015. Glycomimetics versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands. Chem. Rev. 115, 525–561. https://doi.org/10.1021/cr500303t

Chang, I.J., He, M., Lam, C.T., 2018. Congenital disorders of glycosylation. Ann. Transl. Med. 6, 477–477. https://doi.org/10.21037/atm.2018.10.45

Chemani, C., Imberty, A., de Bentzmann, S., Pierre, M., Wimmerova, M., Guery, B.P., Faure, K., 2009. Role of LecA and LecB Lectins in Pseudomonas aeruginosa-Induced Lung Injury and Effect of Carbohydrate Ligands. Infect. Immun. 77, 2065–2075. https://doi.org/10.1128/IAI.01204-08

Chen, C.P., Song, S.C., Gilboa-Garber, N., Chang, K.S.S., Wu, A.M., 1998. Studies on the binding site of the galactose-specific agglutinin PA-IL from Pseudomonas aeruginosa. Glycobiology 8, 7–16. https://doi.org/10.1093/glycob/8.1.7

Chorney, W., Scully, N.J., Crespi, H.L., Katz, J.J., 1960. The growth of algae in deuterium oxide. BBA - Biochim. Biophys. Acta 37, 280–287. https://doi.org/10.1016/0006-3002(60)90235-3

Ciche, T.A., Kim, K.S., Kaufmann-Daszczuk, B., Nguyen, K.C.Q., Hall, D.H., 2008. Cell invasion and matricide during Photorhabdus luminescens transmission by Heterorhabditis bacteriophora nematodes. Appl. Environ. Microbiol. 74, 2275–2287. https://doi.org/10.1128/AEM.02646-07

Cioci, G., Mitchell, E.P., Gautier, C., Wimmerová, M., Sudakevitz, D., Pérez, S., Gilboa-Garber, N., Imberty, A., 2003. Structural basis of calcium and galactose recognition by the lectin PA-IL of Pseudomonas aeruginosa. FEBS Lett. 555, 297–301. https://doi.org/10.1016/S0014-5793(03)01249-3

Clarke, D.J., 2020. Photorhabdus: A tale of contrasting interactions. Microbiol. (United Kingdom). https://doi.org/10.1099/mic.0.000907

Collins, P.M., Öberg, C.T., Leffler, H., Nilsson, U.J., Blanchard, H., 2012. Taloside Inhibitors of Galectin-1 and Galectin-3. Chem. Biol. Drug Des. 79, 339–346. https://doi.org/10.1111/j.1747-0285.2011.01283.x

Cordara, G., Krengel, U., 2013. Structure determination of lectins by X-ray crystallography - A hands-on approach. Carbohydr. Chem. 39, 222–246. https://doi.org/10.1039/9781849737173-00222

Cress, B.F., Bhaskar, U., Vaidyanathan, D., Williams, A., Cai, C., Liu, X., Fu, L., M-Chari, V., Zhang, F., Mousa, S.A., Dordick, J.S., Koffas, M.A.G., Linhardt, R.J., 2019. Heavy Heparin: A Stable Isotope-Enriched, Chemoenzymatically-Synthesized, Poly-Component Drug. Angew. Chemie Int. Ed. 58, 5962–5966. https://doi.org/10.1002/anie.201900768

Crocker, P.R., Paulson, J.C., Varki, A., 2007. Siglecs and their roles in the immune system. Nat. Rev. Immunol. https://doi.org/10.1038/nri2056

Cummings, R.D., Liu, F., 2009. Chapter 33 Galectins. Essentials Glycobiol. 2nd Ed. 1–12.

Cystic Fibrosis Foundation, 2020. 2019 PATIENT REGISTRY ANNUAL DATA

REPORT.

Daborn, P.J., Waterfield, N., Silva, C.P., Au, C.P.Y., Sharma, S., Ffrench-Constant, R.H., 2002. A single Photorhabdus gene, makes caterpillars floppy (mcf), allows Escherichia coli to persist within and kill insects. Proc. Natl. Acad. Sci. U. S. A. 99, 10742–10747. https://doi.org/10.1073/pnas.102068099

Dahms, N.M., Hancock, M.K., 2002. P-type lectins. Biochim. Biophys. Acta - Gen. Subj. https://doi.org/10.1016/S0304-4165(02)00317-3

Davies, R.C., Neuberger, A., Wilson, B.M., 1969. The dependence of lysozyme activity on pH and ionic strength. BBA - Enzymol. 178, 294–305. https://doi.org/10.1016/0005-2744(69)90397-0

De Hoff, P.L., Brill, L.M., Hirsch, A.M., 2009. Plant lectins: The ties that bind in root symbiosis and plant defense. Mol. Genet. Genomics. https://doi.org/10.1007/s00438-009-0460-8

De Meirelles, J.L., Nepomuceno, F.C., Peña-García, J., Schmidt, R.R., Pérez-Sánchez, H., Verli, H., 2020. Current Status of Carbohydrates Information in the Protein Data Bank. J. Chem. Inf. Model. 60, 684–699. https://doi.org/10.1021/acs.jcim.9b00874

Dean, M., Sung, V.W., 2018. Review of deutetrabenazine: A novel treatment for chorea associated with Huntington's disease. Drug Des. Devel. Ther. https://doi.org/10.2147/DDDT.S138828

Deisenhofer, J., Epp, O., Miki, K., Huber, R., Michel, H., 1985. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3 resolution. Nature 318, 618–624. https://doi.org/10.1038/318618a0

Delaine, T., Collins, P., MacKinnon, A., Sharma, G., Stegmayr, J., Rajput, V.K., Mandal, S., Cumpstey, I., Larumbe, A., Salameh, B.A., Kahl-Knutsson, B., van Hattum, H., van Scherpenzeel, M., Pieters, R.J., Sethi, T., Schambye, H., Oredsson, S., Leffler, H., Blanchard, H., Nilsson, U.J., 2016. Galectin-3-Binding Glycomimetics that Strongly Reduce Bleomycin-Induced Lung Fibrosis and Modulate Intracellular Glycan Recognition. ChemBioChem 17, 1759–1770. https://doi.org/10.1002/cbic.201600285

Derzelle, S., Turlin, E., Duchaud, E., Pages, S., Kunst, F., Givaudan, A., Danchin, A., 2004. The PhoP-PhoQ Two-Component Regulatory System of Photorhabdus luminescens Is Essential for Virulence in Insects. J. Bacteriol. 186, 1270–1279. https://doi.org/10.1128/JB.186.5.1270-1279.2004

Díaz, E.M., Vicente-Manzanares, M., Sacristan, M., Vicente, C., Legaz, M.E., 2011. Fungal lectin of Peltigera canina induces chemotropism of compatible Nostoc cells by constriction-relaxation pulses of cyanobiont cytoskeleton. Plant Signal. Behav. 6, 1525–1536. https://doi.org/10.4161/psb.6.10.16687

Diggle, S.P., Stacey, R.E., Dodd, C., Cámara, M., Williams, P., Winzer, K., 2006. The

galactophilic lectin, LecA, contributes to biofilm development in Pseudomonas aeruginosa. Environ. Microbiol. 8, 1095–1104. https://doi.org/10.1111/j.1462-2920.2006.001001.x

Domenici, L., Monti, M., Bracchi, C., Giorgini, M., Colagiovanni, V., Muzii, L., Benedetti Panici, P., 2016. D-mannose: a promising support for acute urinary tract infections in women. A pilot study. Eur. Rev. Med. Pharmacol. Sci. 20, 2920–2925.

Drickamer, K., 1999. C-type lectin-like domains. Curr. Opin. Struct. Biol. https://doi.org/10.1016/S0959-440X(99)00009-3

Drickamer, K., Taylor, M.E., 2015. Recent insights into structures and functions of C-type lectins in the immune system. Curr. Opin. Struct. Biol. https://doi.org/10.1016/j.sbi.2015.06.003

Du Clos, T.W., 2013. Pentraxins: Structure, Function, and Role in Inflammation. ISRN Inflamm. 2013, 1–22. https://doi.org/10.1155/2013/379040

Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Givaudan, A., Taourit, S., Bocs, S., Boursaux-Eude, C., Chandler, M., Charles, J.F., Dassa, E., Derose, R., Derzelle, S., Freyssinet, G., Gaudriault, S., Médigue, C., Lanois, A., Powell, K., Siguier, P., Vincent, R., Wingate, V., Zouine, M., Glaser, P., Boemare, N., Danchin, A., Kunst, F., 2003. The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens. Nat. Biotechnol. 21, 1307–1313. https://doi.org/10.1038/nbt886

Egorova, K.S., Toukach, P. V., 2018. Glycoinformatics: Bridging Isolated Islands in the Sea of Data. Angew. Chemie Int. Ed. 57, 14986–14990. https://doi.org/10.1002/anie.201803576

Eierhoff, T., Bastian, B., Thuenauer, R., Madl, J., Audfray, A., Aigal, S., Juillot, S., Rydell, G.E., Muller, S., De Bentzmann, S., Imberty, A., Fleck, C., Romër, W., 2014. A lipid zipper triggers bacterial invasion. Proc. Natl. Acad. Sci. U. S. A. 111, 12895–12900. https://doi.org/10.1073/pnas.1402637111

Eleftherianos, I., Boundy, S., Joyce, S.A., Aslam, S., Marshall, J.W., Cox, R.J., Simpson, T.J., Clarke, D.J., Ffrench-Constant, R.H., Reynolds, S.E., 2007. An antibiotic produced by an insect-pathogenic bacterium suppresses host defenses through phenoloxidase inhibition. Proc. Natl. Acad. Sci. U. S. A. 104, 2419–2424. https://doi.org/10.1073/pnas.0610525104

Eleftherianos, I., Shokal, U., Yadav, S., Kenney, E., Maldonado, T., 2017. Insect immunity to entomopathogenic nematodes and their mutualistic bacteria, in: Current Topics in Microbiology and Immunology. Springer Verlag, pp. 123–156. https://doi.org/10.1007/82_2016_52

Endo, Y., Tsurugi, K., 1987. RNA N-glycosidase activity of ricin A-chain. Mechanism of action of the toxic lectin ricin on eukaryotic ribosomes. J. Biol. Chem. 262,

8128–8130. https://doi.org/10.1016/S0021-9258(18)47538-2

Faltinek, L., Fujdiarová, E., Melicher, F., Houser, J., Kašáková, M., Kondakov, N., Kononov, L., Parkan, K., Vidal, S., Wimmerová, M., 2019. Lectin PLL3, a Novel Monomeric Member of the Seven-Bladed β-Propeller Lectin Family. Molecules 24, 4540. https://doi.org/10.3390/molecules24244540

Feldman, M., Bryan, R., Rajan, S., Scheffler, L., Brunnert, S., Tang, H., Prince, A., 1998. Role of flagella in pathogenesis of Pseudomonas aeruginosa pulmonary infection. Infect. Immun. 66, 43–51. https://doi.org/10.1128/iai.66.1.43-51.1998

Filloux, A., 2011. Protein secretion systems in Pseudomonas aeruginosa: An essay on diversity, evolution, and function. Front. Microbiol. https://doi.org/10.3389/fmicb.2011.00155

Fisher, S.J., Blakeley, M.P., Cianci, M., McSweeney, S., Helliwell, J.R., 2012. Protonation-state determination in proteins using high-resolution X-ray crystallography: Effects of resolution and completeness. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 800–809. https://doi.org/10.1107/S0907444912012589

Fisher, Z., Kovalevsky, A.Y., Mustyakimov, M., Silverman, D.N., McKenna, R., Langan, P., 2011. Neutron structure of human carbonic anhydrase II: A hydrogen-bonded water network "switch" is observed between pH 7.8 and 10.0. Biochemistry 50, 9421–9423. https://doi.org/10.1021/bi201487b

Freeze, H.H., 1995. Lectin Affinity Chromatography. Curr. Protoc. Protein Sci. 00. https://doi.org/10.1002/0471140864.ps0901s00

Fujdiarová, E., Houser, J., Dobeš, P., Paulíková, G., Kondakov, N., Kononov, L., Hyršl, P., Wimmerová, M., 2020. Heptabladed β-propeller lectins PLL2 and PHL from *Photorhabdus* spp. recognize *O* -methylated sugars and influence the host immune system. FEBS J. febs.15457. https://doi.org/10.1111/febs.15457

Funken, H., Bartels, K.M., Wilhelm, S., Brocker, M., Bott, M., Bains, M., Hancock, R.E.W., Rosenau, F., Jaeger, K.E., 2012. Specific Association of Lectin LecB with the Surface of Pseudomonas aeruginosa: Role of Outer Membrane Protein OprF. PLoS One 7, 46857. https://doi.org/10.1371/journal.pone.0046857

Gabius, H.-J., 1997. Animal Lectins. Eur. J. Biochem. 243, 543–576. https://doi.org/10.1111/j.1432-1033.1997.t01-1-00543.x

Galli, S., 2014. X-ray Crystallography: One Century of Nobel Prizes. https://doi.org/10.1021/ed500343x

Garber, N., Guempel, U., Gilboa-Garber, N., Royle, R.J., 1987. Specificity of the fucose-binding lectin of *Pseudomonas aeruginosa*. FEMS Microbiol. Lett. 48, 331–334. https://doi.org/10.1111/j.1574-6968.1987.tb02619.x

Gatley, S.J., Wess, M.M., Govoni, P.L., Wagner, A., Katz, J.J., Friedman, A.M., 1986. Deuterioglucose: Alteration of Biodistribution by an Isotope Effect. J. Nucl. Med.

27.

Gerlits, O., Wymore, T., Das, A., Shen, C.H., Parks, J.M., Smith, J.C., Weiss, K.L., Keen, D.A., Blakeley, M.P., Louis, J.M., Langan, P., Weber, I.T., Kovalevsky, A., 2016. Long-Range Electrostatics-Induced Two-Proton Transfer Captured by Neutron Crystallography in an Enzyme Catalytic Site. Angew. Chemie - Int. Ed. 55, 4924–4927. https://doi.org/10.1002/anie.201509989

Gerlits, O.O., Coates, L., Woods, R.J., Kovalevsky, A., 2017. Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A As Revealed by Neutron Crystallography. Biochemistry 56, 4747–4750. https://doi.org/10.1021/acs.biochem.7b00654

Gerrard, J., Waterfield, N., Vohra, R., Ffrench-Constant, R., 2004. Human infection with Photorhabdus asymbiotica: An emerging bacterial pathogen. Microbes Infect. https://doi.org/10.1016/j.micinf.2003.10.018

Gilboa-Garber, N., Sudakevitz, D., Sheffi, M., Sela, R., Levene, C., 1994. PA-I and PA-II lectin interactions with the ABO(H) and P blood group glycosphingolipid antigens may contribute to the broad spectrum adherence of Pseudomonas aeruginosa to human tissues in secondary infections. Glycoconj. J. 11, 414–417. https://doi.org/10.1007/BF00731276

Gillon, E., Varrot, A., Imberty, A., 2020. LecB, a high affinity soluble fucose-binding lectin from pseudomonas aeruginosa, in: Methods in Molecular Biology. Humana Press Inc., pp. 475–482. https://doi.org/10.1007/978-1-0716-0430-4_46

Glick, J., Garber, N., 1983. The intracellular localization of Pseudomonas aeruginosa lectins. J. Gen. Microbiol. 129, 3085–3090. https://doi.org/10.1099/00221287-129-10-3085

Glusker, J.P., Carrell, H.L., Kovalevsky, A.Y., Hanson, L., Fisher, S.Z., Mustyakimov, M., Mason, S., Forsyth, T., Langan, P., 2010. Using neutron protein crystallography to understand enzyme mechanisms. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 1257–1261. https://doi.org/10.1107/S0907444910027915

Goldstein, I.J., Hughes, R.C., Monsigny, M., Osawa, T., Sharon, N., 1980. What should be called a lectin? [2]. Nature. https://doi.org/10.1038/285066b0

Grishin, A. V, Krivozubov, M.S., Karyagina, A.S., Gintsburg, A.L., 2015. Pseudomonas Aeruginosa Lectins As Targets for Novel Antibacterials. Acta Naturae 7, 29–41.

Gupta, M., Zha, J., Zhang, X., Jung, G.Y., Linhardt, R.J., Koffas, M.A.G., 2018. Production of Deuterated Cyanidin 3- O-Glucoside from Recombinant Escherichia coli. ACS Omega 3, 11643–11648. https://doi.org/10.1021/acsomega.8b01134

Gupta, S.K., Berk, R.S., Masinick, S., Hazlett, L.D., 1994. Pili and lipopolysaccharide of Pseudomonas aeruginosa bind to the glycolipid asialo GM1. Infect. Immun. 62, 4572–4579. https://doi.org/10.1128/iai.62.10.4572-4579.1994

Habash, J., Raftery, J., Nuttall, R., Price, H.J., Wilkinson, C., Kalb, A.J., Helliwell, J.R., 2000. Direct determination of the positions of the deuterium atoms of the bound water in concanavalin A by neutron Laue crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 56, 541–550. https://doi.org/10.1107/S0907444900002353

Habash, J., Raftery, J., Weisgerber, S., Cassetta, A., Lehmann, M.S., Høghøj, P., Wilkinson, C., Campbell, J.W., Helliwell, J.R., 1997. Neutron Laue diffraction study of concanavalin A- The proton of Asp28. J. Chem. Soc. - Faraday Trans. 93, 4313–4317. https://doi.org/10.1039/a704143h

Haertlein, M., Moulin, M., Devos, J.M., Laux, V., Dunne, O., Trevor Forsyth, V., 2016. Biomolecular Deuteration for Neutron Structural Biology and Dynamics, in: Methods in Enzymology. pp. 113–157. https://doi.org/10.1016/bs.mie.2015.11.001

Hapeshi, A., Benarroch, J.M., Clarke, D.J., Waterfield, N.R., 2019. Iso-propyl stilbene: A life cycle signal? Microbiol. (United Kingdom) 165, 516–526. https://doi.org/10.1099/mic.0.000790

Hardman, K.D., Ainsworth, C.F., 1972. Structure of Concanavalin a at 2.4-å Resolution. Biochemistry 11, 4910–4919. https://doi.org/10.1021/bi00776a006

Hassan, M.A.A., Rouf, R., Tiralongo, E., May, T.W., Tiralongo, J., 2015. Mushroom lectins: Specificity, structure and bioactivity relevant to human disease. Int. J. Mol. Sci. https://doi.org/10.3390/ijms16047802

Hauck, D., Joachim, I., Frommeyer, B., Varrot, A., Philipp, B., Möller, H.M., Imberty, A., Exner, T.E., Titz, A., 2013. Discovery of two classes of potent glycomimetic inhibitors of pseudomonas aeruginosa LecB with distinct binding modes. ACS Chem. Biol. 8, 1775–1784. https://doi.org/10.1021/cb400371r

Hazes, B., Sastry, P.A., Hayakawa, K., Read, R.J., Irvin, R.T., 2000. Crystal structure of Pseudomonas aeruginosa PAK pilin suggests a main-chain-dominated mode of receptor binding. J. Mol. Biol. 299, 1005–1017. https://doi.org/10.1006/jmbi.2000.3801

Heim, J.B., Hodnik, V., Heggelund, J.E., Anderluh, G., Krengel, U., 2019. Crystal structures of cholera toxin in complex with fucosylated receptors point to importance of secondary binding site. Sci. Rep. 9, 1–14. https://doi.org/10.1038/s41598-019-48579-2

Helliwell, J.R., 2020. Fundamentals of neutron crystallography in structural biology, in: Methods in Enzymology. Academic Press Inc., pp. 1–19. https://doi.org/10.1016/bs.mie.2020.01.006

Hooper, L. V., Gordon, J.I., 2001. Glycans as legislators of host-microbial interactions: Spanning the spectrum from symbiosis to pathogenicity. Glycobiology. https://doi.org/10.1093/glycob/11.2.1R

Houser, J., Komarek, J., Kostlanova, N., Cioci, G., Varrot, A., Kerr, S.C., Lahmann,

M., Balloy, V., Fahy, J. V., Chignard, M., Imberty, A., Wimmerova, M., 2013. A soluble fucose-specific lectin from Aspergillus fumigatus conidia - Structure, specificity and possible role in fungal pathogenicity. PLoS One 8, 83077. https://doi.org/10.1371/journal.pone.0083077

Houser, J., Kozmon, S., Mishra, D., Hammerová, Z., Wimmerová, M., Koča, J., 2020. The CH–π Interaction in Protein–Carbohydrate Binding: Bioinformatics and In Vitro Quantification. Chem. – A Eur. J. 26, 10769–10780. https://doi.org/10.1002/chem.202000593

Howard, E.I., Sanishvili, R., Cachau, R.E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T.R., Joachimiak, A., Podjarny, A., 2004. Ultrahigh resolution drug design I: Details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. Proteins Struct. Funct. Genet. 55, 792–804. https://doi.org/10.1002/prot.20015

Hudlicky, T., Pitzer, K.K., Stabile, M.R., Thorpe, A.J., Whited, G.M., 1996. Biocatalytic syntheses of protected D-mannose-d5, D-mannose-d7, D-mannitol-2,3,4,5,6-d5, and D-mannitol-1,1,2,3,4,5,6,6-d8. J. Org. Chem. 61, 4151–4153. https://doi.org/10.1021/jo951666s

Hunter, S.W., Fujiwara, T., Brennan, P.J., 1982. Structure and antigenicity of the major specific glycolipid antigen of Mycobacterium leprae. J. Biol. Chem. 257, 15072–15078.

Imberty, A., H. Prestegard, J., 2015. Structural Biology of Glycan Recognition, Essentials of Glycobiology. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/GLYCOBIOLOGY.3E.030

Imberty, A., Mitchell, E.P., Wimmerová, M., 2005. Structural basis of high-affinity glycan recognition by bacterial and fungal lectins. Curr. Opin. Struct. Biol. 15, 525–534. https://doi.org/10.1016/j.sbi.2005.08.003

Imberty, A., Varrot, A., 2008. Microbial recognition of human cell surface glycoconjugates. Curr. Opin. Struct. Biol. 18, 567–576. https://doi.org/10.1016/j.sbi.2008.08.001

Imberty, A., Wimmerová, M., Mitchell, E.P., Gilboa-Garber, N., 2004. Structures of the lectins from Pseudomonas aeruginosa: Insights into the molecular basis for host glycan recognition. Microbes Infect. https://doi.org/10.1016/j.micinf.2003.10.016

Jančaříková, G., Houser, J., Dobeš, P., Demo, G., Hyršl, P., Wimmerová, M., 2017. Characterization of novel bangle lectin from Photorhabdus asymbiotica with dual sugar-binding specificity and its effect on host immunity. PLoS Pathog. 13. https://doi.org/10.1371/journal.ppat.1006564

Jennings, L.K., Storek, K.M., Ledvina, H.E., Coulon, C., Marmont, L.S., Sadovskaya, I., Secor, P.R., Tseng, B.S., Scian, M., Filloux, A., Wozniak, D.J., Howell, P.L.,

Parsek, M.R., 2015. Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the Pseudomonas aeruginosa biofilm matrix. Proc. Natl. Acad. Sci. U. S. A. 112, 11353–11358. https://doi.org/10.1073/pnas.1503058112

Johannes, L., Jacob, R., Leffler, H., 2018. Galectins at a glance. J. Cell Sci. 131. https://doi.org/10.1242/jcs.208884

Joyce, S.A., Brachmann, A.O., Glazer, I., Lango, L., Schwär, G., Clarke, D.J., Bode, H.B., 2008. Bacterial biosynthesis of a multipotent stilbene. Angew. Chemie - Int. Ed. 47, 1942–1945. https://doi.org/10.1002/anie.200705148

Kadam, R.U., Garg, D., Schwartz, J., Visini, R., Sattler, M., Stocker, A., Darbre, T., Reymond, J.L., 2013. CH-π "t-shape" interaction with histidine explains binding of aromatic galactosides to Pseudomonas aeruginosa lectin LecA. ACS Chem. Biol. 8, 1925–1930. https://doi.org/10.1021/cb400303w

Kalb, J.A., Myles, D.A.A., Habash, J., Raftery, J., Helliwell, J.R., 2001. Neutron Laue diffraction experiments on a large unit cell: Concanavalin A complexed with methyl-α-D-glucopyranoside. J. Appl. Crystallogr. 34, 454–457. https://doi.org/10.1107/S0021889801007245

Kanai, T., Amakawa, M., Kato, R., Shimizu, K., Nakamura, K., Ito, K.I., Hama, Y., Fujimori, M., Amano, J., 2009. Evaluation of a new method for the diagnosis of alterations of Lens culinaris agglutinin binding of thyroglobulin molecules in thyroid carcinoma. Clin. Chem. Lab. Med. 47, 1285–1290. https://doi.org/10.1515/CCLM.2009.277

Katz, J.J., Crespi, H.L., 1966. Deuterated organisms: Cultivation and uses. Science (80-.). 151, 1187–1194. https://doi.org/10.1126/science.151.3715.1187

Kent, B., Hunt, T., Darwish, T.A., Hauß, T., Garvey, C.J., Bryant, G., 2014. Localization of trehalose in partially hydrated DOPC bilayers: insights into cryoprotective mechanisms. J. R. Soc. Interface 11, 20140069. https://doi.org/10.1098/rsif.2014.0069

Kerem, B.S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., Tsui, L.C., 1989. Identification of the cystic fibrosis gene: Genetic analysis. Science (80-.). 245, 1073–1080. https://doi.org/10.1126/science.2570460

Kilpatrick, D.C., 2002. Animal lectins: A historical introduction and overview. Biochim. Biophys. Acta - Gen. Subj. https://doi.org/10.1016/S0304-4165(02)00308-2

Koch, H.J., Stuart, R.S., 1978. The catalytic C-deuteration of some carbohydrate derivatives. Carbohydr. Res. 67, 341–348. https://doi.org/10.1016/S0008-6215(00)84123-3

Koch, M., Wright, K.E., Otto, O., Herbig, M., Salinas, N.D., Tolia, N.H., Satchwell, T.J., Guck, J., Brooks, N.J., Baum, J., 2017. Plasmodium falciparum erythrocyte-

binding antigen 175 triggers a biophysical change in the red blood cell that facilitates invasion. Proc. Natl. Acad. Sci. U. S. A. 114, 4225–4230. https://doi.org/10.1073/pnas.1620843114

Koruza, K., Lafumat, B., Végvári, Knecht, W., Fisher, S.Z., 2018. Deuteration of human carbonic anhydrase for neutron crystallography: Cell culture media, protein thermostability, and crystallization behavior. Arch. Biochem. Biophys. 645, 26–33. https://doi.org/10.1016/j.abb.2018.03.008

Koruza, K., Mahon, B.P., Blakeley, M.P., Ostermann, A., Schrader, T.E., McKenna, R., Knecht, W., Fisher, S.Z., 2019. Using neutron crystallography to elucidate the basis of selective inhibition of carbonic anhydrase by saccharin and a derivative. J. Struct. Biol. 205, 147–154. https://doi.org/10.1016/j.jsb.2018.12.009

Kovalevsky, A., Gerlits, O., Beltran, K., Weiss, K.L., Keen, D.A., Blakeley, M.P., Louis, J.M., Weber, I.T., 2020. Proton transfer and drug binding details revealed in neutron diffraction studies of wild-type and drug resistant HIV-1 protease, in: Methods in Enzymology. Academic Press Inc., pp. 257–279. https://doi.org/10.1016/bs.mie.2019.12.002

Kovalevsky, A.Y., Hanson, L., Fisher, S.Z., Mustyakimov, M., Mason, S.A., Trevor Forsyth, V., Blakeley, M.P., Keen, D.A., Wagner, T., Carrell, H.L., Katz, A.K., Glusker, J.P., Langan, P., 2010. Metal Ion Roles and the Movement of Hydrogen during Reaction Catalyzed by D-Xylose Isomerase: A Joint X-Ray and Neutron Diffraction Study. Structure 18, 688–699. https://doi.org/10.1016/j.str.2010.03.011

Kovalevsky, A.Y., Katz, A.K., Carrell, H.L., Hanson, L., Mustyakimov, M., Zoe Fisher, S., Coates, L., Schoenborn, B.P., Bunick, G.J., Glusker, J.P., Langan, P., 2008. Hydrogen location in stages of an enzyme-catalyzed reaction: Time-of-flight neutron structure of D-xylose isomerase with bound D-xylulose. Biochemistry 47, 7595–7597. https://doi.org/10.1021/bi8005434

Kumar, A., Sýkorová, P., Demo, G., Dobeš, P., Hyršl, P., Wimmerová, M., 2016. A Novel Fucose-binding Lectin from Photorhabdus luminescens (PLL) with an Unusual Heptabladed β-Propeller Tetrameric Structure. J. Biol. Chem. 291, 25032–25049. https://doi.org/10.1074/jbc.M115.693473

Kumar, K.K., Sridhar, J., Murali-Baskaran, R.K., Senthil-Nathan, S., Kaushal, P., Dara, S.K., Arthurs, S., 2019. Microbial biopesticides for insect pest management in India: Current status and future prospects. J. Invertebr. Pathol. https://doi.org/10.1016/j.jip.2018.10.008

Kung, V.L., Ozer, E.A., Hauser, A.R., 2010. The Accessory Genome of Pseudomonas aeruginosa. Microbiol. Mol. Biol. Rev. 74, 621–641. https://doi.org/10.1128/mmbr.00027-10

Labaude, S., Griffin, C.T., 2018. Transmission success of entomopathogenic nematodes used in pest control. Insects. https://doi.org/10.3390/insects9020072

Laine, R.A., 1994. Invited Commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems. Glycobiology 4, 759–767. https://doi.org/10.1093/glycob/4.6.759

Lambert, P.A., 2002. Mechanisms of antibiotic resistance in Pseudomonas aeruginosa, in: Journal of the Royal Society of Medicine, Supplement. Royal Society of Medicine Press, pp. 22–26.

Lamblin, G., Degroote, S., Perini, J.M., Delmotte, P., Scharfman, A., Davril, M., Lo-Guidice, J.M., Houdret, N., Dumur, V., Klein, A., Roussel, P., 2001. Human airway mucin glycosylation: A combinatory of carbohydrate determinants which vary in cystic fibrosis. Glycoconj. J. https://doi.org/10.1023/A:1020867221861

Langan, P., Sangha, A.K., Wymore, T., Parks, J.M., Yang, Z.K., Hanson, B.L., Fisher, Z., Mason, S.A., Blakeley, M.P., Forsyth, V.T., Glusker, J.P., Carrell, H.L., Smith, J.C., Keen, D.A., Graham, D.E., Kovalevsky, A., 2014. L-Arabinose binding, isomerization, and epimerization by D-Xylose isomerase: X-Ray/Neutron crystallographic and molecular simulation study. Structure 22, 1287–1300. https://doi.org/10.1016/j.str.2014.07.002

Lanne, B., Cîopraga, J., Bergström, J., Motas, C., Karlsson, K.A., 1994. Binding of the galactose-specific Pseudomonas aeruginosa lectin, PA-I, to glycosphingolipids and other glycoconjugates. Glycoconj. J. 11, 292–298. https://doi.org/10.1007/BF00731201

Laughlin, R.S., Musch, M.W., Hollbrook, C.J., Rocha, F.M., Chang, E.B., Alverdy, J.C., 2000. The key role of Pseudomonas aeruginosa PA-I lectin on experimental gut-derived sepsis. Ann. Surg. 232, 133–142. https://doi.org/10.1097/00000658-200007000-00019

Lee, J., Zhang, L., 2014. The hierarchy quorum sensing network in Pseudomonas aeruginosa. Protein Cell 6, 26–41. https://doi.org/10.1007/s13238-014-0100-x

Lee, K.K., Sheth, H.B., Wong, W.Y., Sherburne, R., Paranchych, W., Hodges, R.S., Lingwood, C.A., Krivan, H., Irvin, R.T., 1994. The binding of Pseudomonas aeruginosa pili to glycosphingolipids is a tip-associated event involving the C-terminal region of the structural pilin subunit. Mol. Microbiol. 11, 705–713. https://doi.org/10.1111/j.1365-2958.1994.tb00348.x

Lesman-Movshovich, E., Lerrer, B., Gilboa-Garber, N., 2003. Blocking of Pseudomonas aeruginosa lectins by human milk glycans. Can. J. Microbiol. 49, 230–235. https://doi.org/10.1139/w03-027

Lessie, T.G., Phibbs, P. V., 1984. Alternative pathways of carbohydrate utilization in pseudomonads. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev.mi.38.100184.002043

Lillehoj, E.P., Kim, B.T., Kim, K.C., 2002. Identification of Pseudomonas aeruginosa flagellin as an adhesin for Muc1 mucin. Am. J. Physiol. - Lung Cell. Mol. Physiol. 282. https://doi.org/10.1152/ajplung.00383.2001

Lis, H., Sharon, N., 1998. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. Chem. Rev. 98, 637–674. https://doi.org/10.1021/cr940413g

Liu, J.F., Harbeson, S.L., Brummel, C.L., Tung, R., Silverman, R., Doller, D., 2017. A Decade of Deuteration in Medicinal Chemistry, in: Annual Reports in Medicinal Chemistry. Academic Press Inc., pp. 519–542. https://doi.org/10.1016/bs.armc.2017.08.010

Loris, R., Hamelryck, T., Bouckaert, J., Wyns, L., 1998. Legume lectin structure. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. https://doi.org/10.1016/S0167-4838(97)00182-9

Lowery, R., Gibson, M.I., Thompson, R.L., Fullam, E., 2015. Deuterated carbohydrate probes as "label-free" substrates for probing nutrient uptake in mycobacteria by nuclear reaction analysis. Chem. Commun. 51, 4838–4841. https://doi.org/10.1039/c4cc09588j

Lukacs, G.L., Verkman, A.S., 2012. CFTR: Folding, misfolding and correcting the ΔF508 conformational defect. Trends Mol. Med. https://doi.org/10.1016/j.molmed.2011.10.003

Lv, Z., Chu, Y., Wang, Y., 2015. HIV protease inhibitors: A review of molecular selectivity and toxicity. HIV/AIDS - Res. Palliat. Care 7, 95–104. https://doi.org/10.2147/HIV.S79956

Ma, L., Lu, H., Sprinkle, A., Parsek, M.R., Wozniak, D.J., 2007. Pseudomonas aeruginosa Psl is a galactose- and mannose-rich exopolysaccharide, in: Journal of Bacteriology. American Society for Microbiology Journals, pp. 8353–8356. https://doi.org/10.1128/JB.00620-07

Machado, R.A.R., Wüthrich, D., Kuhnert, P., Arce, C.C.M., Thönen, L., Ruiz, C., Zhang, X., Robert, C.A.M., Karimi, J., Kamali, S., Ma, J., Bruggmann, R., Erb, M., 2018. Whole-genome-based revisit of photorhabdus phylogeny:Proposal for the elevation of most photorhabdus subspecies to the species level and description of one novel species photorhabdus bodei sp. nov., and one novel subspecies photorhabdus laumondii subsp. clarkei subsp. nov. Int. J. Syst. Evol. Microbiol. 68, 2664–2681. https://doi.org/10.1099/ijsem.0.002820

Magalhes, A., Ismail, M.N., Reis, C.A., 2010. Sweet receptors mediate the adhesion of the gastric pathogen Helicobacter pylori: Glycoproteomic strategies. Expert Rev. Proteomics. https://doi.org/10.1586/epr.10.18

Manzoni, F., Saraboji, K., Sprenger, J., Kumar, R., Noresson, A.L., Nilsson, U.J., Leffler, H., Fisher, S.Z., Schrader, T.E., Ostermann, A., Coates, L., Blakeley, M.P., Oksanen, E., Logan, D.T., 2016. Perdeuteration, crystallization, data

collection and comparison of five neutron diffraction data sets of complexes of human galectin-3C. Acta Crystallogr. Sect. D Struct. Biol. 72, 1194–1202. https://doi.org/10.1107/S2059798316015540

Manzoni, F., Wallerstein, J., Schrader, T.E., Ostermann, A., Coates, L., Akke, M., Blakeley, M.P., Oksanen, E., Logan, D.T., 2018. Elucidation of Hydrogen Bonding Patterns in Ligand-Free, Lactose- and Glycerol-Bound Galectin-3C by Neutron Crystallography to Guide Drug Design. J. Med. Chem. 61, 4412–4420. https://doi.org/10.1021/acs.jmedchem.8b00081

Mbata, G.N., Shapiro-Ilan, D.I., Alborn, H.T., Strand, M.R., 2019. Preferential infectivity of entomopathogenic nematodes in an envenomed host. Int. J. Parasitol. 49, 737–745. https://doi.org/10.1016/j.ijpara.2019.05.002

McEver, R.P., 2015. Selectins: Initiators of leucocyte adhesion and signalling at the vascular wall. Cardiovasc. Res. https://doi.org/10.1093/cvr/cvv154

McPherson, A., Gavira, J.A., 2014. Introduction to protein crystallization. Acta Crystallogr. Sect. FStructural Biol. Commun. https://doi.org/10.1107/S2053230X13033141

Meilleur, F., Myles, D.A.A., Blakeley, M.P., 2006. Neutron Laue macromolecular crystallography. Eur. Biophys. J. https://doi.org/10.1007/s00249-006-0074-6

Meusch, D., Gatsogiannis, C., Efremov, R.G., Lang, A.E., Hofnagel, O., Vetter, I.R., Aktories, K., Raunser, S., 2014. Mechanism of Tc toxin action revealed in molecular detail. Nature 508, 61–65. https://doi.org/10.1038/nature13015

Midtgaard, S.R., Darwish, T.A., Pedersen, M.C., Huda, P., Larsen, A.H., Jensen, G.V., Kynde, S.A.R., Skar-Gislinge, N., Nielsen, A.J.Z., Olesen, C., Blaise, M., Dorosz, J.J., Thorsen, T.S., Venskutonytė, R., Krintel, C., Møller, J. V., Frielinghaus, H., Gilbert, E.P., Martel, A., Kastrup, J.S., Jensen, P.E., Nissen, P., Arleth, L., 2018. Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering. FEBS J. 285, 357–371. https://doi.org/10.1111/febs.14345

Mingeot-Leclercq, M.P., Glupczynski, Y., Tulkens, P.M., 1999. Aminoglycosides: Activity and resistance. Antimicrob. Agents Chemother. https://doi.org/10.1128/aac.43.4.727

Mitchell, E., Houles, C., Sudakevitz, D., Wimmerova, M., Gautier, C., Pérez, S., Wu, A.M., Gilboa-Garber, N., Imberty, A., 2002. Structural basis for oligosaccharide-mediated adhesion of Pseudomonas aeruginosa in the lungs of cystic fibrosis patients. Nat. Struct. Biol. 9, 918–921. https://doi.org/10.1038/nsb865

Mitchell, E.P., Sabin, C., Šnajdrová, L., Pokorná, M., Perret, S., Gautier, C., Hofr, C., Gilboa-Garber, N., Koča, J., Wimmerová, M., Imberty, A., 2004. High affinity fucose binding of Pseudomonas aeruginosa lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational

chemistry approaches. Proteins Struct. Funct. Bioinforma. 58, 735–746. https://doi.org/10.1002/prot.20330

Möckl, L., 2020. The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization and Immune System Regulation. Front. Cell Dev. Biol. https://doi.org/10.3389/fcell.2020.00253

Mogayzel, P.J., Naureckas, E.T., Robinson, K.A., Brady, C., Guill, M., Lahiri, T., Lubsch, L., Matsui, J., Oermann, C.M., Ratjen, F., Rosenfeld, M., Simon, R.H., Hazle, L., Sabadosa, K., Marshall, B.C., Mueller, G., Hadjiliadis, D., Hoag, J.B., 2014. Cystic fibrosis foundation pulmonary guideline pharmacologic approaches to prevention and Eradication of Initial Pseudomonas aeruginosa Infection. Ann. Am. Thorac. Soc. https://doi.org/10.1513/AnnalsATS.201404-166OC

Moradali, M.F., Ghods, S., Rehm, B.H.A., 2017. Pseudomonas aeruginosa lifestyle: A paradigm for adaptation, survival, and persistence. Front. Cell. Infect. Microbiol. https://doi.org/10.3389/fcimb.2017.00039

Moulin, M., Strohmeier, G.A., Hirz, M., Thompson, K.C., Rennie, A.R., Campbell, R.A., Pichler, H., Maric, S., Forsyth, V.T., Haertlein, M., 2018. Perdeuteration of cholesterol for neutron scattering applications using recombinant Pichia pastoris. Chem. Phys. Lipids 212, 80–87. https://doi.org/10.1016/j.chemphyslip.2018.01.006

Mousavifar, L., Touaibia, M., Roy, R., 2018. Development of Mannopyranoside Therapeutics against Adherent-Invasive Escherichia coli Infections. Acc. Chem. Res. 51, 2937–2948. https://doi.org/10.1021/acs.accounts.8b00397

Mydock-McGrane, L.K., Hannan, T.J., Janetka, J.W., 2017. Rational design strategies for FimH antagonists: new drugs on the horizon for urinary tract infection and Crohn's disease. Expert Opin. Drug Discov. https://doi.org/10.1080/17460441.2017.1331216

Niimura, N., Bau, R., 2008. Neutron protein crystallography: Beyond the folding structure of biological macromolecules. Acta Crystallogr. Sect. A Found. Crystallogr. https://doi.org/10.1107/S0108767307043498

Niimura, N., Chatake, T., Kurihara, K., Maeda, M., 2004. Hydrogen and hydration in proteins. Cell Biochem. Biophys. 40, 351–69. https://doi.org/10.1385/CBB:40:3:351

Niimura, N., Minezaki, Y., Nonaka, T., Castagna, J.C., Cipriani, F., Høghøj, P., Lehmann, M.S., Wilkinson, C., 1997. Neutron Laue diffractometry with an imaging plate provides an effective data collection regime for neutron protein crystallography. Nat. Struct. Biol. https://doi.org/10.1038/nsb1197-909

Niimura, N., Podjarny, A., 2011. Neutron Protein Crystallography: Hydrogen, Protons, and Hydration in Bio-macromolecules, Neutron Protein Crystallography: Hydrogen, Protons, and Hydration in Bio-macromolecules. Oxford University

Press. https://doi.org/10.1093/acprof:oso/9780199578863.001.0001

Nizet, V., Varki, A., Aebi, M., 2017. Microbial Lectins: Hemagglutinins, Adhesins, and Toxins. https://doi.org/10.1101/GLYCOBIOLOGY.3E.037

O'Neill, H., Shah, R., Evans, B.R., He, J., Pingali, S.V., Chundawat, S.P.S., Jones, A.D., Langan, P., Davison, B.H., Urban, V., 2015. Production of Bacterial Cellulose with Controlled Deuterium-Hydrogen Substitution for Neutron Scattering Studies, in: Methods in Enzymology. Academic Press Inc., pp. 123–146. https://doi.org/10.1016/bs.mie.2015.08.031

Orlandi, P.A., Klotz, F.W., Haynes, J.D., 1992. A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of Plasmodium falciparum recognizes the terminal Neu5Ac(α2-3)Gal- sequences of glycophorin A. J. Cell Biol. 116, 901–909. https://doi.org/10.1083/jcb.116.4.901

Paës, G., Berrin, J.G., Beaugrand, J., 2012. GH11 xylanases: Structure/function/properties relationships and applications. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2011.10.003

Pang, Z., Raudonis, R., Glick, B.R., Lin, T.J., Cheng, Z., 2019. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2018.11.013

Passos da Silva, D., Matwichuk, M.L., Townsend, D.O., Reichhardt, C., Lamba, D., Wozniak, D.J., Parsek, M.R., 2019. The Pseudomonas aeruginosa lectin LecB binds to the exopolysaccharide Psl and stabilizes the biofilm matrix. Nat. Commun. 10. https://doi.org/10.1038/s41467-019-10201-4

Perret, S., Sabin, C., Dumon, C., Pokorná, M., Gautier, C., Galanina, O., Ilia, S., Bovin, N., Nicaise, M., Desmadril, M., Gilboa-Garber, N., Wimmerov´, M., Mitchell, E.P., Imberty, A., 2005. Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of *Pseudomonas aeruginosa*. Biochem. J. 389, 325–332. https://doi.org/10.1042/BJ20050079

Podjarny, A., Cachau, R.E., Schneider, T., Van Zandt, M., Joachimiak, A., 2004. Subatomic and atomic crystallographic studies of aldose reductase: Implications for inhibitor binding. Cell. Mol. Life Sci. https://doi.org/10.1007/s00018-003-3404-1

Poole, K., 2005. Aminoglycoside resistance in Pseudomonas aeruginosa. Antimicrob. Agents Chemother. https://doi.org/10.1128/AAC.49.2.479-487.2005

Poole, K., 2004. Resistance to β-lactam antibiotics. Cell. Mol. Life Sci. https://doi.org/10.1007/s00018-004-4060-9

Priem, B., Gilbert, M., Wakarchuk, W.W., Heyraud, A., Samain, E., 2002. A new fermentation process allows large-scale production of human milk oligosaccharides by metabolically engineered bacteria. Glycobiology 12, 235–240. https://doi.org/10.1093/glycob/12.4.235

Raetz, C.R.H., Whitfield, C., 2002. Lipopolysaccharide endotoxins. Annu. Rev. Biochem. https://doi.org/10.1146/annurev.biochem.71.110601.135414

Ramakrishnan, V., 2002. Ribosome structure and the mechanism of translation. Cell. https://doi.org/10.1016/S0092-8674(02)00619-0

Ratcliffe, N.A., Gagen, S.J., 1977. Studies on the in vivo cellular reactions of insects: An ultrastructural analysis of nodule formation in Galleria mellonella. Tissue Cell 9, 73–85. https://doi.org/10.1016/0040-8166(77)90050-7

Reichhardt, C., Wong, C., da Silva, D.P., Wozniak, D.J., Parsek, M.R., 2018. CDRA interactions within the pseudomonas aeruginosa biofilm matrix safeguard it from proteolysis and promote cellular packing. MBio 9. https://doi.org/10.1128/mBio.01376-18

Reily, C., Stewart, T.J., Renfrow, M.B., Novak, J., 2019. Glycosylation in health and disease. Nat. Rev. Nephrol. https://doi.org/10.1038/s41581-019-0129-4

Rhim, A.D., Stoykova, L.I., Trindade, A.J., Glick, M.C., Scanlin, T.F., 2004. Altered terminal glycosylation and the pathophysiology of CF lung disease. J. Cyst. Fibros. 3, 95–96. https://doi.org/10.1016/j.jcf.2004.05.021

Rodrigue, J., Ganne, G., Blanchard, B., Saucier, C., Giguère, D., Shiao, T.C., Varrot, A., Imberty, A., Roy, R., 2013. Aromatic thioglycoside inhibitors against the virulence factor LecA from Pseudomonas aeruginosa. Org. Biomol. Chem. 11, 6906–6918. https://doi.org/10.1039/c3ob41422a

Rosenbaum, D.M., Rasmussen, S.G.F., Kobilka, B.K., 2009. The structure and function of G-protein-coupled receptors. Nature. https://doi.org/10.1038/nature08144

Russell, M.W., Bobek, L.A., Brock, J.H., Hajishengallis, G., Tenovuo, J., 2005. Innate humoral defense factors, in: Mucosal Immunology, Two-Volume Set. Elsevier Inc., pp. 73–93. https://doi.org/10.1016/B978-012491543-5/50009-7

Sabin, C., Mitchell, E.P., Pokorná, M., Gautier, C., Utille, J.-P., Wimmerová, M., Imberty, A., 2006. Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: Thermodynamics data correlated with X-ray structures. FEBS Lett. 580, 982–987. https://doi.org/10.1016/j.febslet.2006.01.030

Saraboji, K., Håkansson, M., Genheden, S., Diehl, C., Qvist, J., Weininger, U., Nilsson, U.J., Leffler, H., Ryde, U., Akke, M., Logan, D.T., 2012. The carbohydrate-binding site in galectin-3 is preorganized to recognize a sugarlike framework of oxygens: ultra-high-resolution structures and water dynamics. Biochemistry 51, 296–306. https://doi.org/10.1021/bi201459p

Sauer, M.M., Jakob, R.P., Luber, T., Canonica, F., Navarra, G., Ernst, B., Unverzagt, C., Maier, T., Glockshuber, R., 2019. Binding of the Bacterial Adhesin FimH to Its Natural, Multivalent High-Mannose Type Glycan Targets. J. Am. Chem. Soc.

141, 936–944. https://doi.org/10.1021/jacs.8b10736

Sawama, Y., Yabe, Y., Iwata, H., Fujiwara, Y., Monguchi, Y., Sajiki, H., 2012. Stereo- and Regioselective Direct Multi-Deuterium-Labeling Methods for Sugars. Chem. - A Eur. J. 18, 16436–16442. https://doi.org/10.1002/chem.201202852

Scanlin, T.F., Wang, Y.M., Glick, M.C., 1985. Altered fucosylation of membrane glycoproteins from cystic fibrosis fibroblasts. Pediatr. Res. 19, 368–374. https://doi.org/10.1203/00006450-198519040-00011

Scharfman, A., Arora, S.K., Delmotte, P., Van Brussel, E., Mazurier, J., Ramphal, R., Roussel, P., 2001. Recognition of lewis x derivatives present on mucins by flagellar components of Pseudomonas aeruginosa. Infect. Immun. 69, 5243–5248. https://doi.org/10.1128/IAI.69.9.5243-5248.2001

Schönemann, W., Cramer, J., Mühlethaler, T., Fiege, B., Silbermann, M., Rabbani, S., Dätwyler, P., Zihlmann, P., Jakob, R.P., Sager, C.P., Smieško, M., Schwardt, O., Maier, T., Ernst, B., 2019. Improvement of Aglycone π-Stacking Yields Nanomolar to Sub-nanomolar FimH Antagonists. ChemMedChem 14, 749–757. https://doi.org/10.1002/cmdc.201900051

Schuster, M., Hawkins, A.C., Harwood, C.S., Greenberg, E.P., 2004. The Pseudomonas aeruginosa RpoS regulon and its relationship to quorum sensing. Mol. Microbiol. 51, 973–985. https://doi.org/10.1046/j.1365-2958.2003.03886.x

Sciacchitano, S., Lavra, L., Morgante, A., Ulivieri, A., Magi, F., De Francesco, G.P., Bellotti, C., Salehi, L.B., Ricci, A., 2018. Galectin-3: One molecule for an alphabet of diseases, from A to Z. Int. J. Mol. Sci. https://doi.org/10.3390/ijms19020379

Scotet, V., L'hostis, C., Férec, C., 2020. The changing epidemiology of cystic fibrosis: Incidence, survival and impact of the CFTRGene discovery. Genes (Basel). https://doi.org/10.3390/genes11060589

Sheth, H.B., Lee, K.K., Wong, W.Y., Srivastava, G., Hindsgaul, O., Hodges, R.S., Paranchych, W., Irvin, R.T., 1994. The pili of Pseudomonas aeruginosa strains PAK and PAO bind specifically to the carbohydrate sequence βGalNAc(1–4)βGal found in glycosphingolipids asialo-GM1 and asialo-GM2. Mol. Microbiol. 11, 715–723. https://doi.org/10.1111/j.1365-2958.1994.tb00349.x

Shigeki Shimba, S.S., Unno, K., Okada, S., 1990. Study on the Biodistribution of Deuterated Biomolecules in Mice Aiming at New Diagnostic Radio-Imaging Agents. Chem. Pharm. Bull. 38, 2610–2613. https://doi.org/10.1248/cpb.38.2610

Singh, R.S., Bhari, R., Kaur, H.P., 2011. Characteristics of yeast lectins and their role in cell-cell interactions. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2011.06.002

Sommer, R., Joachim, I., Wagner, S., Titz, A., 2013. New approaches to control infections: Anti-biofilm strategies against gramnegative bacteria. Chimia (Aarau). 67, 286–290. https://doi.org/10.2533/chimia.2013.286

Sommer, R., Wagner, S., Rox, K., Varrot, A., Hauck, D., Wamhoff, E.C., Schreiber, J., Ryckmans, T., Brunner, T., Rademacher, C., Hartmann, R.W., Brönstrup, M., Imberty, A., Titz, A., 2018. Glycomimetic, Orally Bioavailable LecB Inhibitors Block Biofilm Formation of Pseudomonas aeruginosa. J. Am. Chem. Soc. 140, 2537–2545. https://doi.org/10.1021/jacs.7b11133

Somvanshi, V.S., Kaufmann-Daszczuk, B., Kim, K.S., Mallon, S., Ciche, T.A., 2010. Photorhabdus phase variants express a novel fimbrial locus, mad, essential for symbiosis. Mol. Microbiol. 77, 1021–1038. https://doi.org/10.1111/j.1365-2958.2010.07270.x

Somvanshi, V.S., Sloup, R.E., Crawford, J.M., Martin, A.R., Heidt, A.J., Kim, K.S., Clardy, J., Ciche, T.A., 2012. A single promoter inversion switches photorhabdus between pathogenic and mutualistic states. Science (80-.). 336, 88–93. https://doi.org/10.1126/science.1216641

Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K., Palese, P., Paulson, J.C., Wilson, I.A., 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. J. Mol. Biol. 355, 1143–1155. https://doi.org/10.1016/j.jmb.2005.11.002

Stillmark, H., 1888. Ueber ricin: ein giftiges ferment aus den samen von Ricinus comm. 1. und einigen anderen euphorbiaceen. PhD Thesis. Schnakenburg's buchdruckr.

Stoitsova, S.R., Boteva, R.N., Doyle, R.J., 2003. Binding of hydrophobic ligands by Pseudomonas aeruginosa PA-I lectin. Biochim. Biophys. Acta - Gen. Subj. 1619, 213–219. https://doi.org/10.1016/S0304-4165(02)00496-8

Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowallk, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K.S., Wu, Z., Paulsen, I.T., Relzer, J., Saler, M.H., Hancock, R.E.W., Lory, S., Olson, M. V., 2000. Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature 406, 959–964. https://doi.org/10.1038/35023079

Strauch, O., Ehlers, R.U., 1998. Food signal production of Photorhabdus luminescens inducing the recovery of entomopathogenic nematodes Heterorhabditis spp. in liquid culture. Appl. Microbiol. Biotechnol. 50, 369–374. https://doi.org/10.1007/s002530051306

Su, J., Raghuwanshi, V.S., Raverty, W., Garvey, C.J., Holden, P.J., Gillon, M., Holt, S.A., Tabor, R., Batchelor, W., Garnier, G., 2016. Smooth deuterated cellulose films for the visualisation of adsorbed bio-macromolecules. Sci. Rep. 6, 1–11. https://doi.org/10.1038/srep36119

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D.L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., Ouellette, M., Outterson,

K., Patel, J., Cavaleri, M., Cox, E.M., Houchens, C.R., Grayson, M.L., Hansen, P., Singh, N., Theuretzbacher, U., Magrini, N., Aboderin, A.O., Al-Abri, S.S., Awang Jalil, N., Benzonana, N., Bhattacharya, S., Brink, A.J., Burkert, F.R., Cars, O., Cornaglia, G., Dyar, O.J., Friedrich, A.W., Gales, A.C., Gandra, S., Giske, C.G., Goff, D.A., Goossens, H., Gottlieb, T., Guzman Blanco, M., Hryniewicz, W., Kattula, D., Jinks, T., Kanj, S.S., Kerr, L., Kieny, M.P., Kim, Y.S., Kozlov, R.S., Labarca, J., Laxminarayan, R., Leder, K., Leibovici, L., Levy-Hara, G., Littman, J., Malhotra-Kumar, S., Manchanda, V., Moja, L., Ndoye, B., Pan, A., Paterson, D.L., Paul, M., Qiu, H., Ramon-Pardo, P., Rodríguez-Baño, J., Sanguinetti, M., Sengupta, S., Sharland, M., Si-Mehand, M., Silver, L.L., Song, W., Steinbakk, M., Thomsen, J., Thwaites, G.E., van der Meer, J.W., Van Kinh, N., Vega, S., Villegas, M.V., Wechsler-Fördös, A., Wertheim, H.F.L., Wesangula, E., Woodford, N., Yilmaz, F.O., Zorzet, A., 2018. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect. Dis. 18, 318–327. https://doi.org/10.1016/S1473-3099(17)30753-3

Takaba, K., Tai, Y., Eki, H., Dao, H.-A., Hanazono, Y., Hasegawa, K., Miki, K., Takeda, K., 2019. Subatomic resolution X-ray structures of green fluorescent protein. IUCrJ 6, 387–400. https://doi.org/10.1107/S205225251900246X

Tielker, D., Hacker, S., Loris, R., Strathmann, M., Wingender, J., Wilhelm, S., Rosenau, F., Jaeger, K.E., 2005. Pseudomonas aeruginosa lectin LecB is located in the outer membrane and is involved in biofilm formation. Microbiology 151, 1313–1323. https://doi.org/10.1099/mic.0.27701-0

Varki, A., Kornfeld, S., 2017. Historical Background and Overview. https://doi.org/10.1101/GLYCOBIOLOGY.3E.001

Varrot, A., Basheer, S.M., Imberty, A., 2013. Fungal lectins: Structure, function and potential applications. Curr. Opin. Struct. Biol. https://doi.org/10.1016/j.sbi.2013.07.007

Varrot, A., Blanchard, B., Imberty, A., 2011. Lectin Binding and its Structural Basis, in: Carbohydrate Recognition. Wiley, pp. 329–347. https://doi.org/10.1002/9781118017586.ch13

Vasta, G.R., 2009. Roles of galectins in infection. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro2146

Vestweber, D., Blanks, J.E., 1999. Mechanisms that regulate the function of the selectins and their ligands. Physiol. Rev. https://doi.org/10.1152/physrev.1999.79.1.181

Vines, R.R., Ramakrishnan, G., Rogers, J.B., Lockhart, L.A., Mann, B.J., Petri, W.A., 1998. Regulation of adherence and virulence by the Entamoeba histolytica lectin cytoplasmic domain, which contains a β2 integrin motif. Mol. Biol. Cell 9, 2069–2079. https://doi.org/10.1091/mbc.9.8.2069

Wan, Q., Parks, J.M., Hanson, B.L., Fisher, S.Z., Ostermann, A., Schrader, T.E., Graham, D.E., Coates, L., Langan, P., Kovalevsky, A., 2015. Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography. Proc. Natl. Acad. Sci. U. S. A. 112, 12384–12389. https://doi.org/10.1073/pnas.1504986112

Wang, S., Liu, X., Liu, H., Zhang, L., Guo, Y., Yu, S., Wozniak, D.J., Ma, L.Z., 2015. The exopolysaccharide Psl-eDNA interaction enables the formation of a biofilm skeleton in Pseudomonas aeruginosa. Environ. Microbiol. Rep. 7, 330–340. https://doi.org/10.1111/1758-2229.12252

Waterfield, N.R., Bowen, D.J., Fetherston, J.D., Perry, R.D., Ffrench-Constant, R.H., 2001. The tc genes of Photorhabdus: A growing family. Trends Microbiol. https://doi.org/10.1016/S0966-842X(01)01978-3

Waterfield, N.R., Ciche, T., Clarke, D., 2009. Photorhabdus and a Host of Hosts . Annu. Rev. Microbiol. 63, 557–574. https://doi.org/10.1146/annurev.micro.091208.073507

Watson, J.D., Crick, F.H.C., 1953. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature 171, 737–738. https://doi.org/10.1038/171737a0

Weichert, S., Jennewein, S., Hüfner, E., Weiss, C., Borkowski, J., Putze, J., Schroten, H., 2013. Bioengineered 2′-fucosyllactose and 3-fucosyllactose inhibit the adhesion of Pseudomonas aeruginosa and enteric pathogens to human intestinal and respiratory cell lines. Nutr. Res. 33, 831–838. https://doi.org/10.1016/j.nutres.2013.07.009

Wilton, M., Charron-Mazenod, L., Moore, R., Lewenza, S., 2016. Extracellular DNA acidifies biofilms and induces aminoglycoside resistance in Pseudomonas aeruginosa. Antimicrob. Agents Chemother. 60, 544–553. https://doi.org/10.1128/AAC.01650-15

Winzer, K., Falconer, C., Garber, N.C., Diggle, S.P., Camara, M., Williams, P., 2000. The Pseudomonas aeruginosa lectins PA-IL and PA-IIL are controlled by quorum sensing and by RpoS. J. Bacteriol. 182, 6401–11.

Yamada, S., Suzuki, Y., Suzuki, T., Le, M.Q., Nidom, C.A., Sakai-Tagawa, Y., Muramoto, Y., Ito, M., Kiso, Maki, Horimoto, T., Shinya, K., Sawada, T., Kiso, Makoto, Usui, T., Murata, T., Lin, Y., Hay, A., Haire, L.F., Stevens, D.J., Russell, R.J., Gamblin, S.J., Skehel, J.J., Kawaoka, Y., 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. Nature 444, 378–382. https://doi.org/10.1038/nature05264

Yang, G., Dowling, A.J., Gerike, U., Ffrench-Constant, R.H., Waterfield, N.R., 2006. Photorhabdus virulence cassettes confer injectable insecticidal activity against the wax moth. J. Bacteriol. 188, 2254–2261. https://doi.org/10.1128/JB.188.6.2254-2261.2006

Yau, T., Dan, X., Ng, C.C.W., Ng, T.B., 2015. Lectins with potential for anti-cancer therapy. Molecules. https://doi.org/10.3390/molecules20033791

Yoshimura, F., Nikaido, H., 1982. Permeability of Pseudomonas aeruginosa outer membrane to hydrophilic solutes. J. Bacteriol. 152, 636–642.

Zahorska, E., Kuhaudomlarp, S., Minervini, S., Yousaf, S., Lepsik, M., Kinsinger, T., Hirsch, A.K.H., Imberty, A., Titz, A., 2020. A rapid synthesis of low-nanomolar divalent LecA inhibitors in four linear steps from d-galactose pentaacetate. Chem. Commun. 56, 8822–8825. https://doi.org/10.1039/d0cc03490h

Zhang, L., Shi, L., Shen, Y., Miao, Y., Wei, M., Qian, N., Liu, Y., Min, W., 2019. Spectral tracing of deuterium for imaging glucose metabolism. Nat. Biomed. Eng. 3, 402–413. https://doi.org/10.1038/s41551-019-0393-4

Zhou, J., Yang, W., Hu, Y., Höti, N., Liu, Y., Shah, P., Sun, S., Clark, D., Thomas, S., Zhang, H., 2017. Site-Specific Fucosylation Analysis Identifying Glycoproteins Associated with Aggressive Prostate Cancer Cell Lines Using Tandem Affinity Enrichments of Intact Glycopeptides Followed by Mass Spectrometry. Anal. Chem. 89, 7623–7630. https://doi.org/10.1021/acs.analchem.7b01493

Zupancic, M.L., Frieman, M., Smith, D., Alvarez, R.A., Cummings, R.D., Cormack, B.P., 2008. Glycan microarray analysis of Candida glabrata adhesin ligand specificity. Mol. Microbiol. 68, 547–559. https://doi.org/10.1111/j.1365-2958.2008.06184.x