

Induction of Rare Conformation of Oligosaccharide by Binding to Calcium-dependent Bacterial Lectin: X-ray Crystallography and Modelling Study

Martin Lepsik ^{a, *}, Roman Sommer ^{b, c, d}, Sakonwan Kuhaudomlarp ^a, Mickaël Lelimosin ^a, Emanuele Paci ^e, Annabelle Varrot ^a, Alexander Titz ^{b, c, d}, Anne Imberty ^{a, *}

^a Université Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France

^b Chemical Biology of Carbohydrates, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, D-66123 Saarbrücken, Germany

^c Deutsches Zentrum für Infektionsforschung (DZIF), Standort Hannover-Braunschweig, Germany

^d Department of Pharmacy, Saarland University, D-66123 Saarbrücken, Germany

^e Astbury Centre & School of Molecular and Cellular Biology, University of Leeds, Leeds, U.K.

* Corresponding author. Martin Lepsik (martin.lepsik@cermav.cnrs.fr), Anne Imberty (anne.imberty@cermav.cnrs.fr)

ABSTRACT

Pathogenic micro-organisms utilize protein receptors (lectins) in adhesion to host tissues, a process that in some cases relies on the interaction between lectins and human glycoconjugates. Oligosaccharide epitopes are recognized through their three-dimensional structure and their flexibility is a key issue in specificity. In this paper, we analysed by X-ray crystallography the structures of the LecB lectin from two strains of *Pseudomonas aeruginosa* in complex with Lewis x oligosaccharide present on cell surfaces of human tissues. An unusual conformation of the glycan was observed in all binding sites with a non-canonical *syn* orientation of the *N*-acetyl group of *N*-acetyl-glucosamine. A PDB-wide search revealed that such an orientation occurs only in 4 % of protein/carbohydrate complexes. Theoretical chemistry calculations showed that the observed conformation is unstable in solution but stabilised by the lectin. A reliable description of LecB/Lewis x complex by force field-based methods had proven as especially challenging due to the special feature of the binding site, two closely apposed Ca^{2+} ions which induce strong charge delocalisation. By comparing various force-field parametrisations, we propose a general strategy which will be useful in near future for designing carbohydrate-based ligands (glycodrugs) against other calcium-dependent protein receptors.

KEYWORDS

Lectin, carbohydrate; calcium ion; Quantum effect; Molecular Dynamics; *N*-Acetyl

INTRODUCTION

Pseudomonas aeruginosa is a gram-negative bacterium, which acts as an opportunistic pathogen responsible for severe bronchopulmonary infections notably in cystic fibrosis patients. *P. aeruginosa* disposes of many soluble virulence factors, including two lectins, LecA (PA-IL) and LecB (PA-IIL), with specificity for galactose (Gal) and fucose (Fuc), respectively [1, 2]. Structurally, LecB is a homotetramer with four binding sites, each of which contains two Ca^{2+} atoms [3] that mediate the binding of Fuc (Fig. 1A, 1B) accompanied by a strong charge delocalisation [4]. The high affinity to Fuc results in strong attachment to branched fucosylated oligosaccharides present in the human blood group epitopes such as the AB(H) and Lewis series (Fig. 1C) [5].

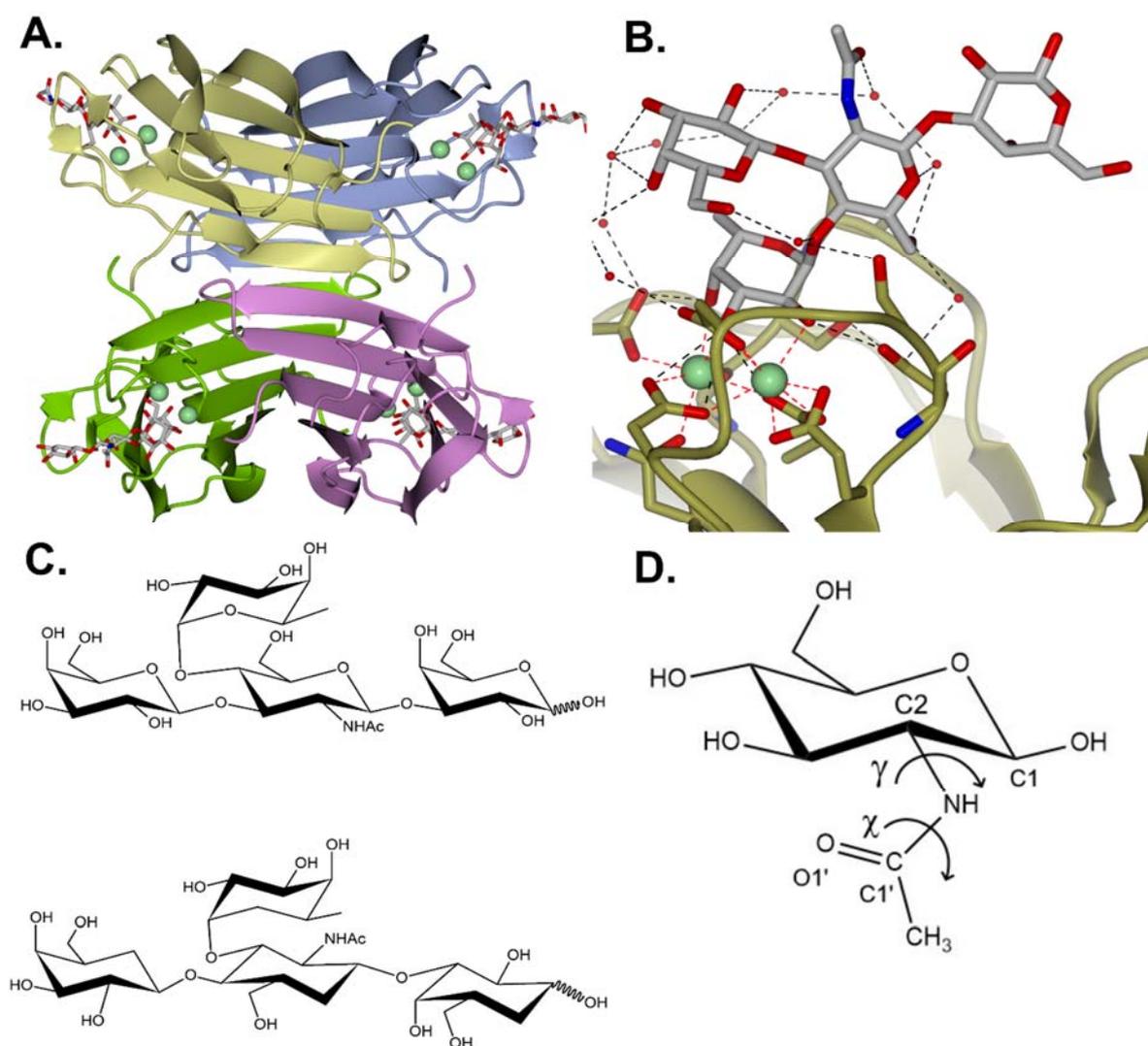


Figure 1. Structure of LecB and fucosylated ligands. A Overall view of LecB_{PA14} tetramer (coloured cartoon) with two calcium ions (green spheres) and one Le^a tetrasaccharide bound

(sticks) in each binding site (PDB: 5A6Z)[6]. **B** Zoom in the binding site displaying the orientation of the *N*-acetyl group towards the solvent. **C** The chemical structures of Le^a (top) and Le^x (bottom) tetrasaccharides. **D** The chemical structure of D-GlcNAc with definitions of γ and χ dihedrals [7].

Among human epitopes, LecB demonstrated higher affinity for Lewis a oligosaccharide (Le^a: Fuc α (1-4) [Gal β (1-3)] GlcNAc) which was rationalised by X-ray structure of the complex [4]. In the binding site, Fuc coordinates the two calcium ions via its hydroxyl groups, which moreover act as hydrogen bond donors to acidic amino acids of the binding site. Furthermore, *N*-acetyl-glucosamine (GlcNAc) creates an additional hydrogen bond with the protein surface. Different *P. aeruginosa* strains have been investigated [6, 8], since the classical PAO1 laboratory strain is not the only one involved in clinical infection. LecB from PA14 (LecB_{PA14}) and PAO1 (LecB_{PAO1}) strains differ by 13 % in their sequences but both bind to Le^a with high affinity as determined by titration microcalorimetry (K_d of 78 nM and 170 nM, respectively) [6]. Binding to Lewis x epitope (Le^x: Fuc α (1-3) [Gal β (1-4)] GlcNAc) is almost as efficient as to Le^a with affinities to LecB_{PA14} and LecB_{PAO1} of 90 and 400 nM, respectively) [4, 6]. This was unexpected since the orientation of GlcNAc in Le^x is different from that in Le^a and the *N*-acetyl was predicted to create steric conflict with the protein surface [3].

The Lewis epitopes are the targets for viral receptors, such as those of noroviruses, and for a number of bacterial lectins [5, 9]. These oligosaccharides are usually described as rigid due to their branched structure which results in stacking between Fuc and Gal rings, steric hindrance of the *N*-acetyl group of *N*-acetylglucosamine (GlcNAc) and the presence of an unconventional CH \cdots O hydrogen bond [10-12]. The conformation in solution of Lewis oligosaccharides is referred to as the “closed” conformation, which is the one recognized by most protein receptors. However, a class of lectins from fungi and bacteria has been reported to induce a conformation change in Le^x and a series of “open” conformations were observed by X-ray crystallography and rationalised by molecular dynamics (MD) calculations [13]. Characterising the interaction of bacterial lectins with their natural ligands is of interest for the development of anti-infective compounds that are able to inhibit the adhesion of bacteria or the formation of biofilm [14]. For example, the structure of LecB complexed with Le^a was used for the design of fucose-derived glycomimetics with aromatic aglycone mimicking the GlcNAc ring [15] and also for the design of mannose/fucose-derived glycomimetics carrying sulphonamide substituents [16, 17].

We describe here the crystal structures of LecB from PAO1 and PA14 strains in complex with Le^x and analyse the conformation of GlcNAc (Fig. 1D) by quantum mechanical (QM) and molecular dynamics (MD) calculations. The presence of the two bridging calcium ions and water molecules in the binding site necessitates special care for force-field parameterization as

previously demonstrated [18, 19] and thus several approaches are compared here. The proposed strategy will be of general interest in the future for the design of active compounds against calcium-containing lectins from pathogens.

RESULTS

Crystal structures of LecB complexed with Le^x

LecB_{PA14} and LecB_{PA01} were co-crystallised with the Le^x tetrasaccharide, resulting in crystals in P2₁ space group diffracting to 1.6 Å and 1.8 Å, respectively. The structures were solved by molecular replacement and details are listed in Table S1. Clear electron density was observed for the whole oligosaccharides in all four binding sites of both structures, except for the reducing-end Gal in LecB_{PA01} (Fig. S1). Consistent with previous findings, both structures are homotetramers with four binding sites, each of which contains two Ca²⁺ atoms [3, 6]. In the same manner as for the complexes with Fuc [3] or Le^a [4, 6], the calcium ions mediate the binding of Fuc residue of Le^x via three hydroxyl oxygens (O2, O3 and O4). Additionally, the three hydroxyl groups of Fuc can form several direct hydrogen bonds with amino acid residues within the binding site (N21:O, D96:OD1, D99:OD2, D101:OD2 and G114*:OXT). The recognition of the Fuc residue by LecB_{PA14} and LecB_{PA01} are in complete agreement with previous observations [3, 6].

In order to rationalise the recognition of Le^x by the LecB variants (LecB_{PA14} and LecB_{PA01}) in relation to the amino acid mutations in their binding sites (S97/A23 vs G97/S23, respectively), a detailed comparison of their complexes with Le^x was performed. The serine at position 97 in LecB_{PA14} provides a hydrophilic environment and organises waters W1, W2, and W3 to form several hydrogen-bond bridges between the protein and the carbohydrate (Fig. 2A). As a result of the S97G variation in LecB_{PA01}, W2 and W3 are absent in the LecB_{PA01}/Le^x complex and the water-mediated interaction is replaced by a direct hydrogen bond between the *N*-acetyl of the GlcNAc and the D96 side chain (Fig. 2B). The W1 interactions with the carbohydrate are retained in both proteins. From the protein side, W1 is held by an H-bond from the N98:N backbone.

The A23 variant in LecB_{PA14} constitutes (together with T45) a hydrophobic pocket which accepts the Fuc:C6 methyl, as described previously for LecB_{PA14}/Le^a [6]. Further toward the bulk solvent, a water network links the carbohydrate with the protein; W4, present in this water network, forms a direct contact with the non-reducing Gal (Fig. 2A, 2C). In LecB_{PA01}, S23 was found in two alternate conformations, one of which provides a direct H-bond to GlcNAc:O1' (Fig. 2B, 2D). The presence of the S23 polar side chain disrupts the water network observed in the LecB_{PA14} complex.

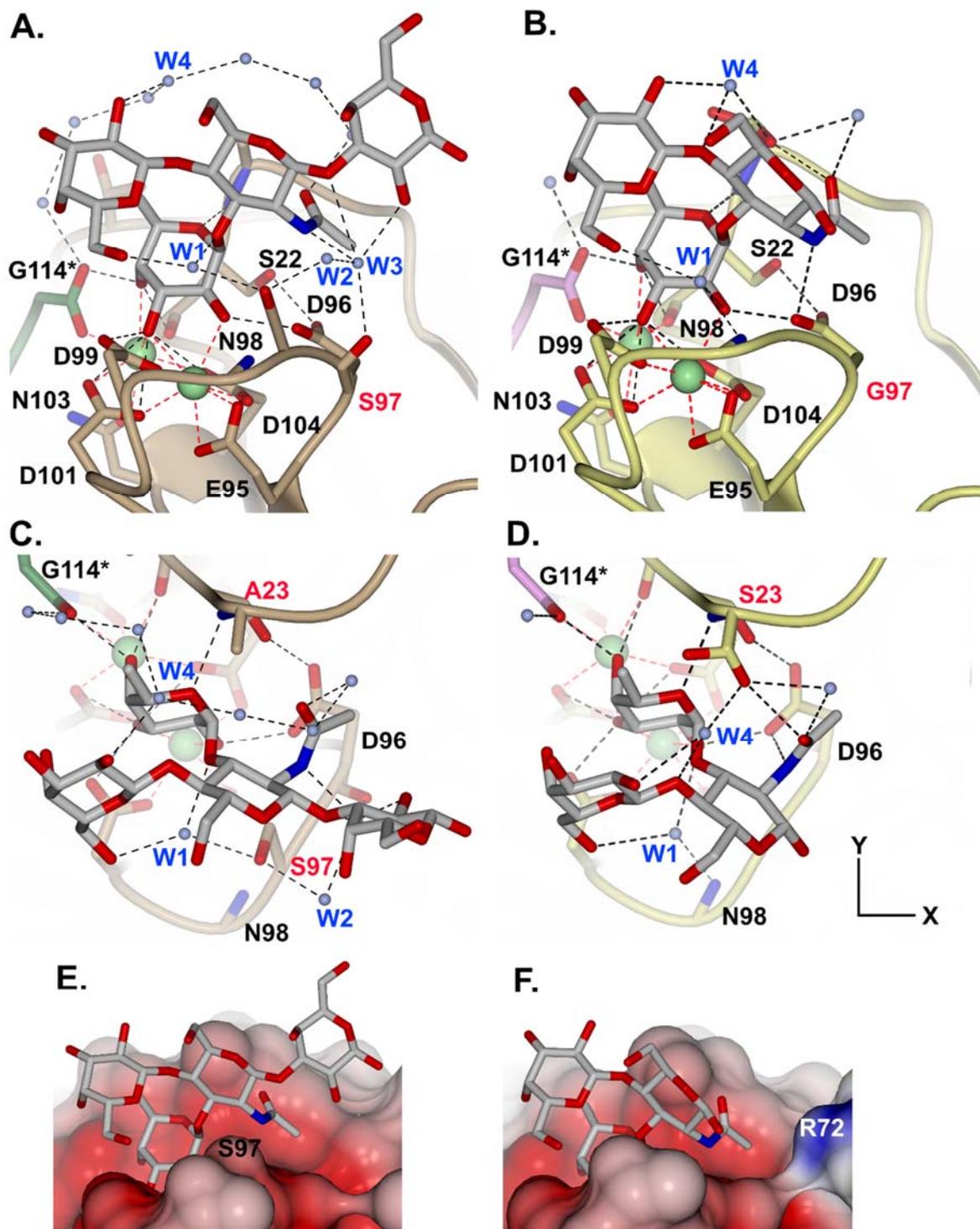


Figure 2. Comparison of the crystallographic structures of Le^x in complex with LecB_{PA14} (A, C, PDB: 5A70) and LecB_{PA01} (B, D, PDB: 6R35). The views in C, D correspond to a 90° counterclockwise rotation of A, B, respectively, around the x axis. Ca²⁺ atoms are represented as green spheres. The protein backbones of LecB_{PA14} and LecB_{PA01} are coloured in light brown and yellow, respectively. G114* is the C-terminal residue of the adjacent monomer and is coloured in green (in LecB_{PA14}) or pink (in LecB_{PA01}). Water molecules are shown as ice-blue spheres. Carbon atoms of Le^x are coloured in light grey (note that in LecB_{PA01}, the reducing galactose residue could not be modelled due to poor electron density; Fig. S1). E and F are surface representations of A and B, respectively. The surfaces are coloured by electrostatic

potential (red; negative, blue; positive). Note that R72 is present in both LecB variants, but the side chain of R72 was modelled in LecB_{PAO1} only and thus is only visible in **F**. The amino acid variants are labelled in red. The figure was rendered in CCP4MG.

Due to the amino acid variations, some differences are observed in the interaction between Le^x and the two LecB variants. While the Fuc binding is virtually unchanged, the Φ , Ψ torsion angles of the α Fuc1-3GlcNAc glycosidic linkage in chains A and C differ by up to 16.5° and 11.5°, between LecB_{PAO1} and LecB_{PA14} (Table S2), which results in a slight change in the interactions. However, in chain D, these torsions are very similar, with changes of 0.9° and 0.5° (Table S2). The main observed differences are in the interaction between GlcNAc and the protein surface. In LecB_{PAO1}, the NH of GlcNAc is close enough to the Asp96 side chain to form a direct hydrogen bond (Fig. 2B, D). In LecB_{PA14}, the GlcNAc ring is pushed slightly away and instead of a direct hydrogen bond, a water bridge via W3 is observed between NH of GlcNAc and the backbone oxygen of Asp96 (Fig. 2A, C). In addition, the S97G mutation in LecA_{PAO1} provides a hydrophobic patch, enabling a closer contact of the GlcNAc ring with the protein surface, which is not observed in LecB_{PA14} due to the polar OH group of S97 side chain (Fig. 2E and 2F).

In the crystal structures of both LecB variants interacting with Le^x, an unusual conformation is observed for the *N*-acetyl group of all four Le^x molecules of the asymmetric unit. In all cases, the C2-N bond adopts a non-canonical *syn* orientation of the γ dihedral angle (Fig. 1D, Table 1). The electron density supports this finding unequivocally (Fig. S1). The χ torsion angles always remain in the classical *trans* conformation (Fig. 1D, Table 1). In all the binding sites, the unusual conformation of the *N*-acetyl brings the bulky acetyl group away from the protein surface, thus avoiding steric conflict. The NH moiety points in the direction of Asp96, establishing a direct hydrogen bond with the Asp96 side-chain carboxylate in LecB_{PAO1}, and a water bridged one with Asp96 backbone in LecB_{PA14} (Table 1). The presence of S97 in the LecB_{PA14}/Le^x complex and the accompanying hydration therefore results in a slightly different position and hydrogen bond network from the LecB_{PAO1}/Le^x complex.

Table 1. Geometric characteristics of the *N*-acetyl of Le^x in the crystal structures of complexes with LecB_{PA14} and LecB_{PAO1}. γ and χ torsion angles (°) are defined as γ (H2 – C2 – N – H), χ (H – N – C1' – O1') and **d1** (Å) is the GlcNAc:N \cdots D96:OD distance.

chain	γ		χ		d1	
	LecB _{PA14}	LecB _{PAO1}	LecB _{PA14}	LecB _{PAO1}	LecB _{PA14}	LecB _{PAO1}
A	-6.8	-8.3	172.4	177.1	3.7	3.0
B	-9.0	-11.7	174.1	169.5	3.6	3.1
C	-6.9	-13.3	171.7	161.3	3.5	3.1
D	-1.8	12.4	169.1	149.7	3.5	3.4

Analysis of GlcNAc *N*-acetyl group in the Protein Data Bank

In order to evaluate the frequency of the non-canonical orientation of *N*-acetyl, the GlcNAc residues were analysed in the Protein Data Bank (PDB). Such a study was previously performed for glycoproteins and protein/glycan complexes [20], but with no quality check, either on the crystallographic resolution or on the electron density. The search (as of March 17, 2019) yielded altogether 381 conformers, with 338 in the canonical γ -*anti*/ χ -*trans* conformation, i.e. 88.7 % of occurrences, with dihedrals ranging from 139° to 209° (-151°) for γ and 140° to 210° (-150°) for χ (Fig. 3).

The other observed structures can be assigned to two non-canonical conformations, *syn/trans* or *anti/cis*, and one outlier *syn/cis*. These minor populations were further analysed by direct examination of the electron density maps using Coot [21] and the structures which were not supported by electron density were discarded. For *syn/trans* torsion combination, 16 cases were confirmed including 8 from this work, i.e. 4 % occurrence with γ ranging from -10° to 11° and χ from 172° to 185° (-175°) (Fig. 3). For *anti/cis*, only six cases were retained, i.e. 1.6 % occurrence with γ ranging from 165° to 188° (-172°) and χ from -21° to 1° (Fig. 3). The *syn/cis* outlier was not supported by electron density and as such it was discarded.

Analysis of the PDB complemented with examination of the electron density maps confirmed eight cases plus another eight observed here crystallographically (i.e. 4% of the torsions studied) of the *N*-acetyl γ dihedral in *syn* conformation.

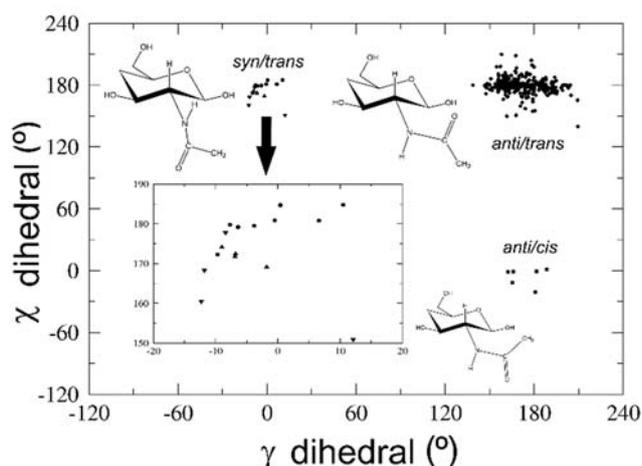


Figure 3. Analysis of γ and χ dihedral angles of *N*-acetyl in β -D-GlcNAc in glycoproteins and protein/carbohydrate complexes in the PDB (each point represents one monosaccharide in a PDB structure). The relative orientation of the functional groups of *N*-acetyl in the minor populations is shown schematically. Inset: Zoom in the γ -*syn*/ χ -*trans* region with the crystal

structures determined in this work labelled as triangle up (LecB_{PA14}/Le^x) and triangle down (LecB_{PA01}/Le^x).

Conformational behaviour of the *N*-Acetyl group in Le^x

The conformational behaviour of the Le^x trisaccharide was previously studied by us [13] by use of molecular dynamics using the GLYCAM06-j force field [22] in explicit water (TIP3P model) [23]. Several independent trajectories of 1 to 10 μ s were produced, starting from the Le^x solution shape with the *N*-acetyl group in the canonical *anti/trans* conformation. The behaviour of the *N*-acetyl group was analysed for the present work and no variations were observed for the *N*-acetyl group in these trajectories (data not shown).

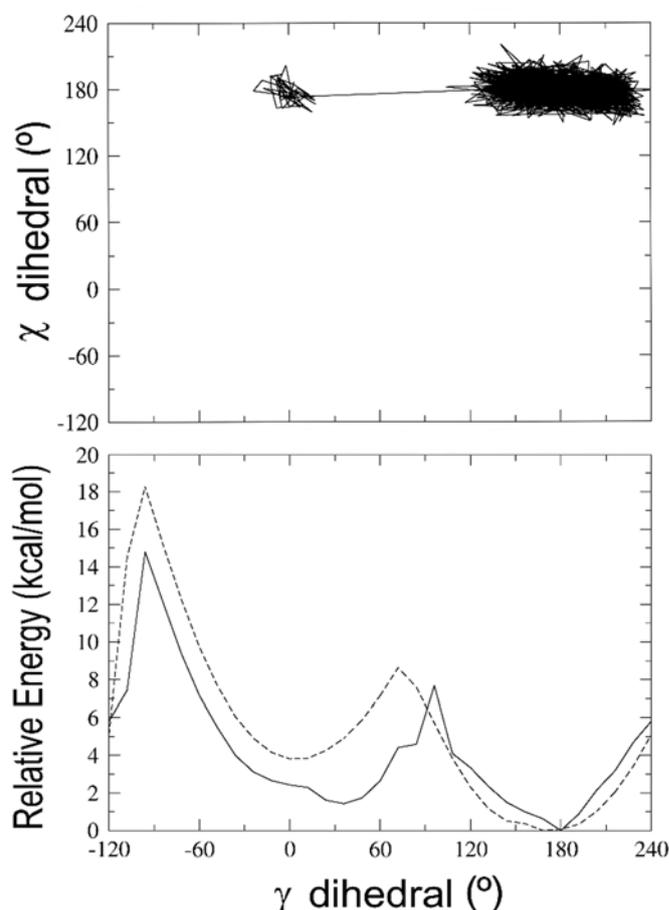


Figure 4. Conformational stability of the *N*-acetyl γ dihedral of the Le^x tetrasaccharide. **A.** MD/explicit solvent. *Syn/trans* ($0/180^\circ$) changed to *anti/trans* ($180/180^\circ$) after 50 ns and stayed until 1.4 μ s. **B.** The relaxed rotational scan of the *N*-acetyl γ dihedral in Le^x tetrasaccharide using quantum mechanics (solid line) and molecular mechanics (dashed line) calculations in implicit solvent (for details, see text).

In the present work, a 1.4 μ s MD simulation of Le^x tetrasaccharide in TIP3P water was started from the γ -*syn* conformation from the LecB_{PA14}/Le^x crystal structure. We observed a rapid conversion within 50 ns to γ -*anti* and its persistence until the end of the simulation (Fig. 4, top). The χ dihedral stayed in all the cases in *trans* in range from 147° to 221° (-139°). Therefore, the MD of the Le^x in aqueous solution confirms the stability of the major *anti/trans* conformation observed in most of the protein/carbohydrate complexes in the PDB. However, a validation of barriers to *syn/anti* rotation is necessary in order to obtain trajectories with relevant occupancies of the different states.

The relaxed rotational scans of the γ dihedral in Le^x were calculated at two levels (Fig. 4, bottom). First, for reference, we used quantum mechanics (QM; density functional theory with empirical dispersion (DFT-D3)[24] with B-LYP functional and double- ζ basis set (DZVP) of a triple- ζ quality [25] coupled with implicit solvent COSMO model [26] in Turbomole [27]. Second, for comparison with the MD simulations, we used molecular mechanics (MM; GLYCAM06-j)[22] coupled with generalised Born (GB) implicit solvent using the *igb5* parametrisation [28] in AMBER18 [29]. Overall, both levels of theory give similar shapes of the curves. The locations and relative stabilities of the minima in Le^x mirror the populations found in the PDB (c.f. Fig. 3) and MD (Fig. 4, top). The global minimum is at γ -*anti* (180°) and the local minimum at γ -*syn* (0°; relative energy to the minimum of the method is 3.8 kcal/mol for MM and 36° and 1.4 kcal/mol for QM). Barriers to rotation of the *N*-acetyl γ dihedral from *syn* to *anti* are asymmetric. The high barriers at -96° of 14.5 kcal/mol (i.e. 18.3 – 3.8 kcal/mol at 0°) in MM and 13.4 kcal/mol (i.e. 14.8 – 1.4 kcal/mol at 36°) in QM are brought about by the repulsions between the GlcNAc:O1' carbonyl oxygen and GlcNAc:O3, which also induces unfavourable amide bond non-planarity (Fig. S2A). The smaller barrier of 4.8 kcal/mol (i.e. 8.6 – 3.8 kcal/mol at 0°) at 72° for MM and 6.3 kcal/mol (i.e. 7.7 - 1.4 kcal/mol at 36°) at 96° for QM is reduced due to the stabilising H-bonds between the GlcNAc *N*-acetyl and the adjacent monosaccharide units (GlcNAc:N-H \cdots Fuc:O2 and reducing Gal:O2-H2O \cdots GlcNAc:O1'; Fig. S2B). For a backward rotation (i.e. *anti*-to-*syn*), the barrier is higher – 8.6 or 7.7 kcal/mol in MM or QM, respectively, which suggests why we could not observe any such transition in 1.4 μ s MD.

The χ dihedral angle stayed *trans* in all the γ rotational profile optimisations which is not surprising given the high rotational barriers of 19-20 kcal/mol calculated from NMR experiments [20]. Additionally, we optimised Le^x tetrasaccharide with *anti/trans* and *anti/cis* conformations at the QM/COSMO level and the latter was higher in energy by 2.6 kcal/mol. The reasons behind this finding are the intrinsic low stability of the χ *cis* isomer as found for GlcNAc by NMR [20] and also the lack of stabilising intramolecular H-bonds.

Molecular dynamics of LecB/Le^x complex

To gain insight into the stabilisation of the rare γ dihedral in the crystallographic LecB_{PA14}/Le^x complex, we performed 600 ns MD simulations using the standard force fields for the protein (ff14SB) [30], carbohydrates (GLYCAM06-j) [22], water (TIP3P) [23] and Ca²⁺ ions [31] (Setup 1, Table 2). In line with previous observations [18, 19], the distance between the two calcium ions increased from the crystallographic value of 3.74 Å (Fig. 5A, yellow) to 4.74 Å (average distance of the 150-600 ns window, Fig. 5A, black) due to a too high repulsion. A more advanced set of calcium parameters employing the C_4 charge-induced dipole term [32] was thereafter used (Setup 2) but with a similar outcome - the average distance was 4.73 Å (200-600 ns window, Fig. 5A, red). To model the known charge transfer effects [4, 33] implicitly, we used calcium parametrisation including effective polarisation [34] (Setup 3). The agreement with the experiment was excellent with average values of 3.80±0.15 Å (Fig. 5A, green). Inspired by [35], we additionally scaled the Ca²⁺-coordinating carboxylates (Setup 4) [36], which gave a fair agreement with the crystallographic distance (4.07±0.14 Å; Fig. 5A, blue). Nonbonded parameters for Ca²⁺, K⁺ and Cl⁻ ions in different MD setups are gathered in Table S3 and available in Bryce’s AMBER parameter database [37]. Partial charges on the Ca²⁺-coordinating carboxylates in different MD setups are gathered in Table S4, [36] and available in Bryce’s AMBER parameter database [37].

Table 2. Summary of the tested MD protocols with different parameterization of calcium ions: classical nonbonded set optimised for structural studies (12-6-IOD) [31], nonbonded set with added charge-induced dipole term (12-6-4) [32] or scaled charge set (Electronic Continuum Correction with Rescaling; ECCR) [34].

Setup	Ca ²⁺	Protein	Water
1	12-6-IOD	ff14SB	TIP3P
2	12-6-4	ff14SB	TIP3P
3	ECCR	ff14SB	TIP3P
4	ECCR	ff14SB/scaled carboxylates	TIP3P
5	ECCR	ff14SB	SPC/E
6	ECCR	ff14SB	OPC3
7	ECCR	ff14SB	TIP4P

Further, we analysed the behaviour of the γ torsion throughout the MD using Setups 3 and 4, since they best maintained the Ca²⁺...Ca²⁺ distance. Moreover, we tested three other water models in combination with the scaled Ca²⁺ parameters [34] because we knew from the crystal structure that water networks bridge LecB with Le^x. Besides the classical TIP3P model (Setup 3), these additional water models are: another three-site rigid SPC/E model [38]

(Setup 5), a more recent three-site OPC3 model [39] (Setup 6) and TIP4P four-site model [40] (Setup 7).

In Fig. 5C (green) we see that MD with Setup 3 maintained the N -acetyl group in the γ -*syn* conformation for most of the simulated time (85%). After 170 ns, it departed to γ -*anti* for 90 ns but thereafter it stayed at γ -*syn* until the end of the trajectory. In Setup 4 (Fig. 5C, blue), the γ -*syn* conformation was maintained for 88% of the simulation time, i.e. until 530 ns, then it changed to γ -*anti* and stayed until the end of the simulation [36]. It should be noted that a further extension to 1.4 μ s showed the persistence of the non-native γ -*anti* dihedral. For the three water models in combination with the scaled-charge Ca^{2+} model, the γ -*syn* conformation was retained for a larger proportion of time (99, 100 and 98%, for Setup 5, 6, and 7, respectively).

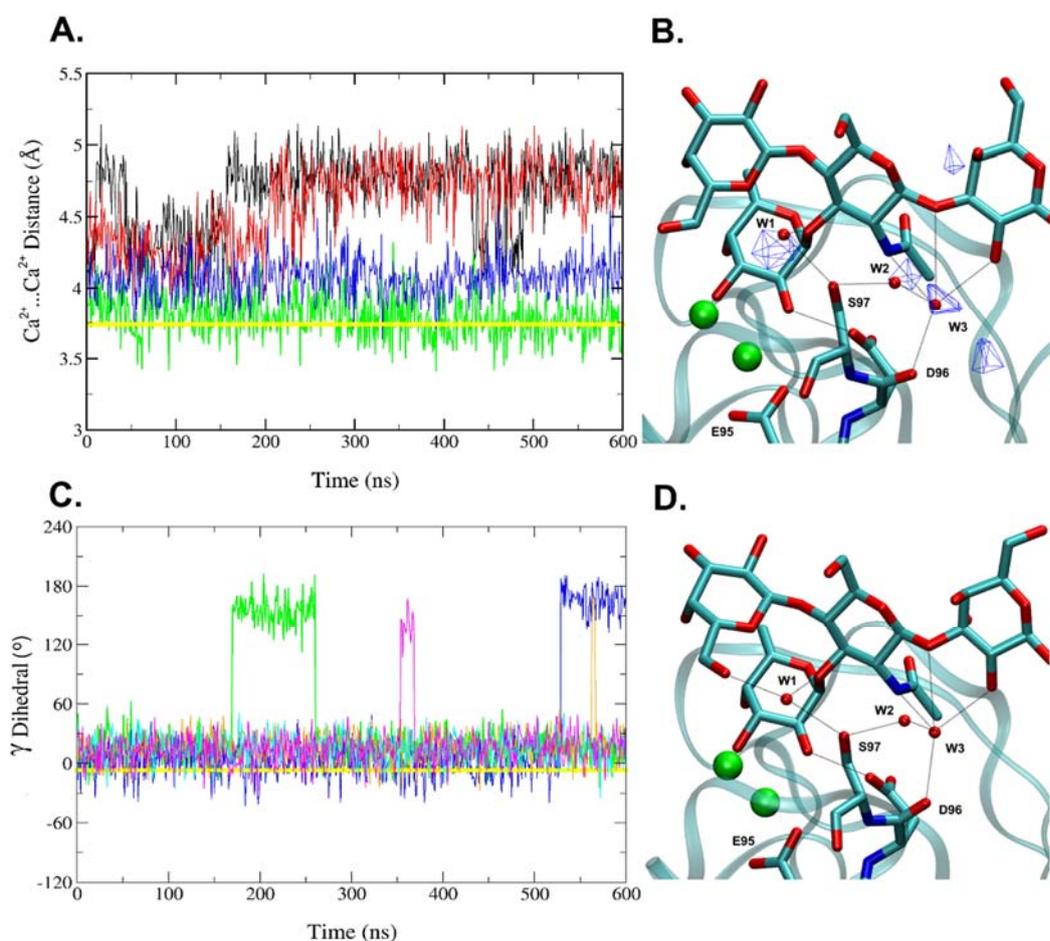


Figure 5. Analyses of 600 ns MD simulations of LecB_{PA14}/Le^x complex (A, B, C) and comparison with the crystal structure (D). A. The Ca²⁺...Ca²⁺ distance and C. the N -acetyl γ torsion. Curve colour coding: MD simulations with different setups: black – Setup 1; red – Setup 2; green – Setup 3; blue – Setup 4, orange – Setup 5, cyan – Setup 6, magenta – Setup 7.

The LecB_{PA14}/Le^x crystal structure (chain A) determined here – yellow. Structural view of the interaction between LecB (ribbon) and Le^x (sticks) from **B**, a representative snapshot from MD using Setup 4 and **D**, the LecB_{PA14}/Le^x crystal structure. Atom colour coding: cyan – carbon, red – oxygen, blue – nitrogen. Hydrogens are omitted from the MD snapshot for clarity. Dotted lines indicate H-bonding. Blue meshes show densities of water oxygen atoms contoured at 0.06 isosurface value. The densities are 0.112, 0.082 and 0.069 for W1, W2, and W3, respectively.

A more detailed insight into the reliability of the computational description of the LecB_{PA14}/Le^x complex is obtained by monitoring two distances. For the interactions of the *N*-acetyl, it is the GlcNAc:N···D96:OD distance (**d1**) which in the LecB_{PA14}/Le^x crystal structure is 3.7 Å in chain A (c.f. Table 1), i.e. it slightly exceeds the criterion for H-bond formation. Regarding the role of Ser97 in organizing the W1-W2-W3 water network, we monitor the Ser97:OG···D96:OD distance (**d2**) which in the crystal structure shows no signs of hydrogen bonding (4.6 Å). In all the MD protocols tested, **d1** varies between 2.8 and 4.5 Å but averages to 3.2 Å (Fig. S3), i.e. forming a direct hydrogen bond, not observed in the LecB_{PA14}/Le^x crystal structure. We note however that i) a direct H-bond has also been observed in the closely related LecB_{PA14}/Le^x crystal structure and ii) that even the direct H-bonding did not prevent transient formation of bifurcated H-bond toward W3 site, which in turn mediated the bridge to D96:O (c.f. Fig. 2). The **d2** distance evolved during MD towards a non-native strong H-bond (distance of 2.6 Å), except Setups 4 and 7. In Setup 7, a strong Ser97:OG···D96:OD H-bond (2.6 Å) was formed for the starting 400 ns and then was lost. But Ser97 then bulged out to the solvent to a non-native position. In Setup 4, Ser97:OG is not attracted so much by the scaled carboxylate of D96 and thus does not make the non-native H-bond. This results in a very good reproduction of the details of the carbohydrate binding site including the location of the Ca²⁺ ions and water networks in Setup 4 [36] as compared to the crystal structure (Fig. 5B, D). Taken together, the charge-scaling procedure of the calcium ions *together with* their coordinating carboxylates (Setup 4, ECCR2) provides the most balanced description of the dynamics in the LecB/Le^x complex. However, the protocol will require further refinement to be stable in μ s-long MD simulations.

Discussion and Conclusions

The properties of isolated GlcNAc have been previously studied by a combination of experimental (NMR) and computational (QM, MD) techniques, which yielded a molecular-level understanding of the structures, dynamics and energetics of several conformational families. Regarding the *N*-acetyl group in GlcNAc, the global minimum of the γ dihedral is predicted to be *anti* (161° and 180° for the α and β anomers, respectively) [41]. The shift in the minimum between the anomers is caused by intramolecular hydrogen bonding. In the β anomer, a minor population (13%) of γ -*syn* conformation was predicted from a 5 ns MD simulation

[41]. We note, however, that this percentage is dependent both, on the force-field type and version (not only for the carbohydrates but also for water) and on the length of the simulation; all these have been radically improved [42]. The χ amide bond dihedral was found by NMR to be predominantly in *trans* (98%) with a high barrier of 19-20 kcal/mol to conversion to χ *cis* [20].

In oligosaccharides containing GlcNAc, especially those where it serves as a branching point, the location and relative stabilities of the minima of the γ dihedral along with the transition barriers may change due to the molecular surroundings. Indeed, we found that a QM rotational profile of Le^x tetrasaccharide was governed by repulsive oxygen/oxygen contacts on one hand and stabilising O-H \cdots O hydrogen bonds, thus shaping the potential energy curve. The GLYCAM06 carbohydrate force field performed qualitatively well in locating the minima but overestimated their energy difference. This translated into the instability of the γ -*syn* dihedral in a 1.4 μ s-long MD of Le^x in explicit solvent. An additional contribution comes from the expected lower energy barriers in explicit solvent. Further relaxation of the system may be due to the pyranose ring chair pucker transitions from the most stable ⁴C₁ chair, which however occurs rarely [13, 43, 44].

A PDB-wide search showed that GlcNAc-containing oligosaccharides bound to proteins exhibited one major (*anti/trans*) and two minor (*syn/trans* and *anti/cis*) populations (c.f. Fig. 3), for which we inspected the experimental electron densities and found that only 16 *syn/trans* cases (including 8 determined in this work, i.e. 4% of the torsions studied) could be confirmed unequivocally. In contrast, for the χ dihedral, in the six cases with good fit in the electron densities we could not confirm the minor *anti/cis* population because the *sp*² substituents could both be fitted upon 180° rotation. The two crystal structures of LecB/Le^x complexes determined here showed that the geometry and characteristics of the LecB binding site were sufficient to stabilise the rare γ -*syn* conformation of *N*-acetyl of GlcNAc. The free energy cost associated with this rearrangement was partly compensated for by the acquired hydrogen bonding (direct or water bridged) between the NH group of the *N*-acetyl and Asp96, as could be observed in the crystal structures.

The LecB carbohydrate-binding site presents a particularly challenging system for force-field calculations due to the charge delocalisation induced by the two closely positioned calcium ions [4]. Previous modelling studies of LecB complexed with monosaccharides indicated problems arising from modelling this special site using the standard 2+ charge model of calcium ion - the Ca²⁺ \cdots Ca²⁺ distance rose by up to 20 % with respect to the crystallographic one. [18, 19]. Furthermore, the present structures of LecB/Le^x complexes determined add another layer of complexity due to the rare γ torsion of GlcNAc, whose force-field parametrisation needs to be validated against QM data. Lastly, the presence of water networks mediating protein/carbohydrate interactions requires testing of several water models. A new

paradigm in modelling Ca^{2+} ions via effective polarisation had been proposed [45] and developed for inorganic salts, proteins and lipids [34, 35, 46]. The strategy is to scale calcium ion (and counterion) charges by 0.75, reparametrize Lennard-Jones parameters and optionally scale the charged groups in biomolecules. Herein, both these approaches were applied to the LecB/Le^x complex and yielded results superior to a standard treatment of calcium ions with 2+ partial charge. Besides the standard TIP3P water model, we tested three additional ones (SPC/E, OPC3 and TIP4P) because the crystal structures had shown that water networks bridge the protein/carbohydrate interface and thus the chosen water model may be of importance. Using the combination of TIP4P water model with scaling Ca^{2+} and counterions only (Setup 7), had a slightly better conservation of the γ dihedral than with the TIP3P water model (Setup 3) but after 400 ns the binding site geometry was impaired by a formation of non-native hydrogen bond. The most promising setup (Setup 4) utilised scaled-charge Ca^{2+} ions, counterions and Ca^{2+} -coordinating carboxylates and the TIP3P water model. Such an approach gave slightly longer but still acceptable $\text{Ca}^{2+}\cdots\text{Ca}^{2+}$ distance but less preservation of the rare unstable γ -syn dihedral of Le^x, especially on a μs time scale. On the contrary, due to the scaling of the binding site carboxylates, occurrence of non-native H-bonding seen in Setup 7 was avoided.

In conclusion, to the best of our knowledge, this is the first report on comprehensive testing of force-field parametrisations on protein with two closely positioned calcium ions in the binding site complexed with a carbohydrate in a rare conformation. By identifying one promising setup, we pave the way for further development, exploration and fine-tuning of parameters to obtain a general protocol for reliable description of such difficult systems where quantum phenomena play important roles. Even though further improvements of force fields are needed for a balanced handling of charge delocalization, the present study allowed for a better understanding of protein/carbohydrate interactions, which can in turn be applied for design of glycomimetics of therapeutical interest.

Methods

X-ray crystallography

Molecular cloning, expression and purification of LecB_{PA14} and LecB_{PA01} were performed as previously described [4, 6]. For crystallization, either LecB_{PA14} or LecB_{PA01} was dissolved in water (10 mg mL⁻¹) and incubated with 800 $\mu\text{g}\cdot\text{mL}^{-1}$ of Le^x tetrasaccharide (Elicityl) supplemented with 2 mM CaCl_2 for 30 minutes prior to crystallization experiment. Crystallization was performed by the hanging drop vapour diffusion method using 1 μL of the protein solution with Le^x and 1 μL of the reservoir solution containing 28% - 30% PEG8K,

2 mM CaCl₂, 0.2 M ammonium sulphate, and 0.1 M Tris-HCl pH 8.5 at 19 °C in a 24 well plate. Crystals were cryo-protected with well solution supplemented with 10 - 25% (v/v) glycerol and flash-cooled in liquid nitrogen. Data were collected at ESRF-BM30A (Grenoble, France) for LecB_{PA14} or at SOLEIL-PROXIMA1 (Saint Aubin, France) for LecB_{PAO1}, using an ADSC Q315 CCD detector and a Pilatus 6M hybrid photon counting detector (Dectris), respectively. The recorded data were indexed, integrated and scaled using XDS [47] and merged using AIMLESS [48]. The structures were solved by molecular replacement with PHASER [49] using PDB: 5A6Q for LecB_{PA14}/Le^x or PDB: 1W8H for LecB_{PAO1}/Le^x as a searching template. The model was finalised by further iterations of manual rebuilding in COOT [21] and restrained refinement in REFMAC5 [50]. Ligand libraries were created using Ligand Builder in Coot and built by Acedrg in CCP4 suite [51]. The geometries of the final models were validated with MOLPROBITY [52], wwPDB validation service (<https://validate.rcsb-1.wwpdb.org/>) and PDB-redo (<https://pdb-redo.eu/>) before submission to the Protein Data Bank. All structural figures were prepared using CCP4MG [53]. Data processing and refinement statistics are provided in Table S1 in the Supplementary Information.

Protein Data Bank (PDB) Search

The values of the *N*-acetyl γ and χ dihedrals of β -D-GlcNAc pyranose residues in protein/carbohydrate complexes were measured in the PDB database using the GlyTorsion tool [54]. To increase the confidence in the structural parameters, we set the X-ray resolution cutoff to 1.5 Å. NMR structures were filtered out for the lack of a quality validation criterion. To be able to observe the effect of the neighbouring carbohydrate moieties, we set the minimal oligosaccharide chain length to trisaccharide. The GlyTorsion tool measures the ω 2a (C1-C2-N-C1') and ω 2b (C2-N-C1'-O1') dihedrals ("X-ray definition", i.e. using heavy atoms only, c.f. Fig. 1D). These values were transformed into a more readily understandable "NMR definition" [7] using hydrogens added via LEaP program of AMBER18 [29]: γ dihedral (H2-C2-N-H) and χ dihedral (H-N-C1'-O1') values were obtained by adding the approximate value of 60° (deviations of several degrees from the exact values, data not shown) and the exact value of 180°, respectively, to the "X-ray definition" ones. This approach avoids the need of adding hydrogens to all the structures in a PDB-wide analysis.

Modelling

All the modelling was done on the A chain (or the A/B chain dimer) of LecB_{PA14}/Le^x structure (this work, PDB: 5A70) because of the high quality of the electron density maps of the whole Le^x tetrasaccharide (Fig. S1). Crystallographic water molecules were retained except those with partial occupancy which may induce clashes. Hydrogens were added using the LEaP program of AMBER18 [29]. As suggested by the previous crystal structure of LecB/Fuc complex and semiempirical quantum chemical calculations [4], the carboxylates in the active site were left

unprotonated. The molecular mechanics (MM) force fields used here were: ff14SB [30] for the protein and GLYCAM-06j [22] for the carbohydrate. Several Ca^{2+} parametrisations were used (Table 2): standard nonbonded +2 charge [31] optimized for structural studies, i.e. ion-oxygen distance (IOD) (Setup 1); more advanced setup with +2 charge on Ca^{2+} and employing the C_4 charge-induced dipole term (Setup 2) [32], a setup with charge on calcium and counterions scaled by 0.75 (Setup 3) [34] and lastly a setup inspired by [35] where Setup 3 was augmented with scaling Ca^{2+} -coordinating carboxylates by 0.75 (Setup 4; partial charges are in Table S4, [36] and available in Bryce's AMBER parameter database [37]. Several water models were used in conjunction with the scaled calcium and counterion setup: TIP3P [23] (Setup 3), SPC/E [38] (Setup 5), OPC3 [39] (Setup 6) and TIP4P [40] (Setup 7). Hydrogens were optimized using the generalized Born (GB) implicit solvent model [28] (ϵ_r of 78.5) by 5000 cycles of LBFGS optimization with zeroed dihedral barriers, followed by another 5000 LBFGS cycles in full force field.

Quantum Chemical Calculations

N-Acetyl Rotational Scan. The γ dihedral in the Le^x tetrasaccharide was rotated to 0° using Cuby4 [55]. Thereafter it was rotated stepwise by 12° in either direction up to a dihedral of $-180/180^\circ$ and optimised in each step to tight convergence criteria (maximum energy difference of 0.6 cal/mol, maximum gradient of $0.12 \text{ kcal/mol/\AA}^2$), keeping the γ dihedral fixed. This was performed in Cuby4 [55] at MM/GB level (using GLYCAM06-j [22] for Le^x and *igb=5* option for generalised Born implicit solvation) using AMBER18 [29] or QM/COSMO level (DFT-D3 using B-LYP functional/DZVP basis set [25]) with COSMO [26] implicit solvent using Turbomole 7.0 [27]. It should be noted here that the gas-phase part of this setup has been shown to give triple- ζ quality results for noncovalent interactions while requiring only double- ζ timing [26]. Energies of the relaxed structures relative to the minimum of the respective method were plotted.

QM Charge Calculations. The $\text{LecB}_{\text{PA14}}$ A/B chain dimer with two Le^x tetrasaccharides, four calcium ions and the crystallographic water molecules had its hydrogen atoms optimized by use of MM/GB [28]. Thereafter, it was subjected to 600 cycles of further hydrogen optimization at the semiempirical QM (SQM) level using PM6-D3H4 [56] coupled with COSMO [26] using MOPAC2016 (<http://OpenMOPAC.net>). The resulting structure was subjected to another 600 cycles of hydrogen optimization at QM/SQM/COSMO level, QM being B-LYP/DZVP-D3 [25]. We used a subtractive mechanical embedding scheme with link hydrogen atom approach. The QM part consisted of residues within 4.8 \AA of the two calcium ions in the A chain, the Le^x tetrasaccharide and one crystallographic water molecule (W1), altogether 262 atoms. The SQM part comprised the whole system.

Molecular Dynamics

The LecB chain A/B dimer with two Le^x tetrasaccharides, four calcium ions and the crystallographic water molecules with MM/GB optimized hydrogens was immersed in an octahedral box of water molecules which extended at least 12 Å from the solute. K⁺/Cl⁻ counterions (scaled in Setups 3 to 7; their Lennard-Jones parameters are shown in Table S3 and available via [36]) were added to maintain neutrality and physiological concentration of 0.15 M. Bonds involving hydrogen atoms were constrained using SHAKE [57]. Hydrogen mass repartitioning to 3Da [58] was applied, which allowed us to use a longer integration time step of 4 fs. Initial stepwise relaxation and subsequent MD was performed according to the published protocol [59].

ASSOCIATED CONTENT

The crystallographic methodology and analyses, QM optimised structures, MD analyses and MD parameters are presented in the Supporting Information.

CONTRIBUTORS

M. Lepsik, AI and AT designed the study. RS and SK performed the crystallography work under the guidance of AV. M. Lepsik performed all calculations with the help of M. Lelimosin and EP. M. Lepsik, SK and AI wrote the manuscript.

ACKNOWLEDGEMENT

Dr. Martin Lepsik has received funding for this project from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 795605. The authors acknowledge support by the ANR PIA Glyco@Alps (ANR-15-IDEX-02), Labex ARCANE and CBH-EUR-GS (ANR-17-EURE-0003). Further, we acknowledge funding from the Helmholtz Association (grant no. VH-NG-934). We are grateful to synchrotron SOLEIL (Saint Aubin, France) and ESRF (Grenoble, France) for access and technical support at beamline PROXIMA 1 and BM30A, respectively. Part of the computations presented in this paper were performed using the Froggy platform of the CIMENT infrastructure which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA) and the Equip@Meso project (reference ANR-10-EQPX-29-01). The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the computer resources and technical support provided by EPCC at the

University of Edinburgh, Scotland. We thank Emilie Gillon for providing recombinantly produced LecB_{PAO1} for X-ray crystallographic experiments.

REFERENCES

- [1] N. Gilboa-Garber, *Pseudomonas aeruginosa* lectins, *Methods Enzymol.*, 83 (1982) 378-385.
- [2] A. Imberty, M. Wimmerova, E.P. Mitchell, N. Gilboa-Garber, Structures of the lectins from *Pseudomonas aeruginosa*: Insights into molecular basis for host glycan recognition, *Microb. Infect.*, 6 (2004) 222-229.
- [3] E. Mitchell, C. Houles, D. Sudakevitz, M. Wimmerova, C. Gautier, S. Pérez, A.M. Wu, N. Gilboa-Garber, A. Imberty, Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients, *Nat. Struct. Mol. Biol.*, 9 (2002) 918-921.
- [4] E.P. Mitchell, C. Sabin, L. Šnajdrová, M. Pokorná, S. Perret, C. Gautier, C. Hofr, N. Gilboa-Garber, J. Koča, M. Wimmerová, A. Imberty, 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*, 58 (2005) 735-748.
- [5] J.E. Heggelund, A. Varrot, A. Imberty, U. Krengel, Histo-blood group antigens as mediators of infections, *Curr. Opin. Struct. Biol.*, 44 (2017) 190-200.
- [6] R. Sommer, S. Wagner, A. Varrot, C. Nycholat, A. Khaledi, S. Häussler, J. Paulson, A. Imberty, A. Titz, The virulence factor LecB varies in clinical isolates: consequences for ligand binding and drug discovery, *Chem. Sci.*, 7 (2016) 4990-5001
- [7] X. Hu, I. Carmichael, A.S. Serianni, N-acetyl side-chains in saccharides: NMR J-coupling equations sensitive to CH-NH and NH-CO bond conformations in 2-acetamido-2-deoxy-aldohexopyranosyl rings, *J. Org. Chem.*, 75 (2010) 4899-4910.
- [8] A.M. Boukerb, A. Decor, S. Ribun, R. Tabaroni, A. Rousset, L. Commin, S. Buff, A. Doleans-Jordheim, S. Vidal, A. Varrot, A. Imberty, B. Cournoyer, Genomic rearrangements and functional diversification of *lecA* and *lecB* lectin coding regions impacting the efficacy of glycomimetics directed against *Pseudomonas aeruginosa*, *Frontiers Microb.*, 7 (2016) 811.
- [9] A. Imberty, A. Varrot, Microbial recognition of human cell surface glycoconjugates, *Curr. Opin. Struct. Biol.*, 18 (2008) 567-576.
- [10] R.U. Lemieux, K. Bock, L.T.J. Delbaere, S. Koto, V.S.R. Rao, The conformations of oligosaccharides related to the ABH and Lewis human blood group determinants, *Can. J. Chem.*, 58 (1980) 631-653.
- [11] S. Pérez, N. Mouhous-Riou, N.E. Nifant'ev, Y.E. Tsvetkov, B. Bachet, A. Imberty, Crystal and molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis x trisaccharide, *Glycobiology*, 6 (1996) 537-542.
- [12] M. Zierke, M. Smiesko, S. Rabbani, T. Aeschbacher, B. Cutting, F.H. Allain, M. Schubert, B. Ernst, Stabilization of branched oligosaccharides: Lewis(x) benefits from a nonconventional C-H...O hydrogen bond, *J. Am. Chem. Soc.*, 135 (2013) 13464-13472.
- [13] J. Topin, M. Lelimosin, J. Arnaud, A. Audfray, S. Pérez, A. Varrot, A. Imberty, The hidden conformation of Lewis x, a human histo-blood group antigen, is a determinant for recognition by pathogen lectins, *ACS Chem. Biol.*, 11 (2016) 2011-2020.
- [14] S. Cecioni, A. Imberty, S. Vidal, Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands, *Chem. Rev.*, 115 (2015) 525-561.
- [15] A. Imberty, Y.M. Chabre, R. Roy, Glycomimetics and glycodendrimers as high affinity microbial antiadhesins, *Chem. Eur. J.*, 14 (2008) 7490-7499.
- [16] D. Hauck, I. Joachim, B. Frommeyer, A. Varrot, B. Philipp, H.M. Möller, A. Imberty, T.E. Exner, A. Titz, Discovery of two classes of potent glycomimetic inhibitors of *Pseudomonas aeruginosa* LecB with distinct binding modes, *ACS Chem. Biol.*, 8 (2013) 1775-1784.
- [17] R. Sommer, S. Wagner, K. Rox, A. Varrot, D. Hauck, E.-C. Wamhoff, J. Schreiber, T. Ryckmans, T. Brunner, C. Rademacher, R.W. Hartmann, M. Brönstrup, A. Imberty, A. Titz, First glycomimetic, orally bioavailable LecB inhibitors block biofilm formation of *Pseudomonas aeruginosa*, *J. Am. Chem. Soc.*, 140 (2018) 2537-2545.
- [18] N.K. Mishra, P. Kulhánek, L. Šnajdrová, M. Petřek, A. Imberty, J. Koča, Molecular dynamics study of *Pseudomonas aeruginosa* lectin-II complexed with monosaccharides, *Proteins*, 72 (2008) 382-392.
- [19] R. Sommer, T.E. Exner, A. Titz, A biophysical study with carbohydrate derivatives explains the molecular basis of monosaccharide selectivity of the *Pseudomonas aeruginosa* lectin LecB, *PLoS One*, 9 (2014) e112822.
- [20] X. Hu, W. Zhang, I. Carmichael, A.S. Serianni, Amide cis-trans isomerization in aqueous solutions of methyl N-formyl-D-glucosaminides and methyl N-acetyl-D-glucosaminides: chemical equilibria and exchange kinetics, *J. Am. Chem. Soc.*, 132 (2010) 4641-4652.

- [21] P. Emsley, B. Lohkamp, W. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr. D. Biol. Crystallogr.*, 66 (2010) 486-501.
- [22] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field. *Carbohydrates, J. Comput. Chem.*, 29 (2008) 622-655.
- [23] W.L. Jorgensen, J. Chandrasekhar, J. Madura, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 79 (1983) 926–935.
- [24] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 132 (2010) 154104.
- [25] J. Hostas, J. Rezac, Accurate DFT-D3 calculations in a small basis set, *J. Chem. Theor. Comput.*, 13 (2017) 3575-3585.
- [26] A. Klamt, G. Schüümann., COSMO: A new approach to dielectric screening in solvents with explicit expressions for the dcreening energy and its gradient., *J. Chem. Soc. PerkinTrans. 2*, (1993) 799-805.
- [27] R. Ahlrichs, M. Bar, M. Haser, H. Horn, C. Kolmel, Electronic structure calculations on workstation computers: The program system Turbomole, *Chem. Phys. Lett.*, 162 (1989) 165-169.
- [28] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, *Proteins*, 55 (2004) 383-394.
- [29] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M.J. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc.*, 117 (1995) 5179–5197.
- [30] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theor. Comput.*, 11 (2015) 3696-3713.
- [31] P. Li, B.P. Roberts, D.K. Chakravorty, K.M. Merz, Jr., Rational design of Particle Mesh Ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent, *J. Chem. Theor. Comput.*, 9 (2013) 2733-2748.
- [32] P. Li, K.M. Merz, Jr., Taking into account the Ion-induced dipole interaction in the nonbonded model of ions, *J. Chem. Theor. Comput.*, 10 (2014) 289-297.
- [33] M. Lepsik, M.J. Field, Binding of calcium and other metal ions to the EF-hand loops of calmodulin studied by quantum chemical calculations and molecular dynamics simulations, *J Phys Chem B*, 111 (2007) 10012-10022.
- [34] M. Kohagen, M. Lepšík, P. Jungwirth, Calcium binding to calmodulin by molecular dynamics with effective polarization, *J. Phys. Chem. Lett.*, 5 (2014) 3964-3969.
- [35] T. Martinek, E. Duboue-Dijon, S. Timr, P.E. Mason, K. Baxova, H.E. Fischer, B. Schmidt, E. Pluharova, P. Jungwirth, Calcium ions in aqueous solutions: Accurate force field description aided by ab initio molecular dynamics and neutron scattering, *J. Chem. Phys.*, 148 (2018).
- [36] Dataset in Mendeley Library, <https://data.mendeley.com/datasets/hbkdvdkjk/1>
- [37] AMBER parameter database, <http://research.bmh.manchester.ac.uk/bryce/amber/>
- [38] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.*, 91 (1987) 6269–6271.
- [39] S. Izadi, A.V. Onufriev, Accuracy limit of rigid 3-point water models, *J. Chem. Phys.*, 145 (2016) 074501.
- [40] W.L. Jorgensen, J.D. Madura, Temperature and size dependence for Monte Carlo simulations of TIP4P water, *Mol. Phys.*, 56 (1985) 1381–1392.
- [41] M. Mobli, A. Almond, N-Acetylated amino sugars: the dependence of NMR 3J(HNH2)-couplings on conformation, dynamics and solvent, *Org. Biomol. Chem.*, 5 (2007) 2243-2251.
- [42] R.J. Woods, Predicting the structures of glycans, glycoproteins, and their complexes, *Chem. Rev.*, 118 8005-8024.
- [43] I. Alibay, K.K. Burusco, N.J. Bruce, R.A. Bryce, Identification of rare Lewis oligosaccharide conformers in aqueous solution using enhanced sampling molecular dynamics, *J. Phys. Chem. B*, 122 (2018) 2462-2474.
- [44] B.M. Sattelle, A. Almond, Is N-acetyl-D-glucosamine a rigid 4C1 chair?, *Glycobiology*, 21 (2011) 1651-1662.
- [45] I. Leontyev, A. Stuchebrukhov, Accounting for electronic polarization in non-polarizable force fields, *Phys. Chem. Chem. Phys.*, 13 (2011) 2613-2626.
- [46] J. Melcr, H. Martinez-Seara, R. Nencini, J. Kolafa, P. Jungwirth, O.H.S. Ollila, Accurate binding of sodium and calcium to a POPC bilayer by effective inclusion of electronic polarization, *J. Phys. Chem. B*, 122 (2018) 4546-4557.
- [47] W. Kabsch, XDS, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 66 (2010) 125-132.
- [48] P.R. Evans, An introduction to data reduction: Space-group determination, scaling and intensity statistics, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 67 (2011) 282-292.

- [49] A.J. McCoy, Solving structures of protein complexes by molecular replacement with Phaser, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 63 (2006) 32-41.
- [50] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 67 (2011) 355-367.
- [51] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 67 (2011) 235-242.
- [52] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 66 (2010) 12-21.
- [53] S. McNicholas, E. Potterton, K.S. Wilson, M.E.M. Noble, Presenting your structures: The CCP4mg molecular-graphics software, *Acta Crystallogr., Sect D: Biol. Crystallogr.*, 67 (2011) 386-394.
- [54] T. Lutteke, M. Frank, C.W. von der Lieth, Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB, *Nucleic Acids Res.*, 33 (2005) D242-246.
- [55] J. Rezac, Cuby: An integrative framework for computational chemistry, *J. Comput. Chem.*, 37 (2016) 1230-1237.
- [56] J. Rezac, P. Hobza, Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods, *J. Chem. Theor. Comput.*, 8 (2012) 141-151.
- [57] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comput. Phys.*, 23 (1977) 327-341.
- [58] C.W. Hopkins, S. Le Grand, R.C. Walker, A.E. Roitberg, Long-time-step molecular dynamics through hydrogen mass repartitioning, *J. Chem. Theor. Comput.*, 11 (2015) 1864-1874.
- [59] P. Srb, M. Svoboda, L. Benda, M. Lepsik, J. Tarabek, V. Sicha, B. Gruner, K. Grantz-Saskova, J. Brynda, P. Rezacova, J. Konvalinka, V. Veverka, Capturing a dynamically interacting inhibitor by paramagnetic NMR spectroscopy, *Phys. Chem. Chem. Phys.*, 21 (2019) 5661-5673.