A fucose-binding superlectin from *Enterobacter cloacae* with high Lewis and ABO blood group antigen specificity

Received for publication, October 1, 2024, and in revised form, December 20, 2024 Published, Papers in Press, December 30, 2024,
https://doi.org/10.1016/j.jbc.2024.108151

Ghamdan Beshr[1,2,3], Asfandyar Sikandar[1], Julia Gläser[1,2], Mario Fares[1,2,3], Roman Sommer[1,2,3], Stefanie Wagner[1,2], Jesko Köhnke[4,*], and Alexander Titz[1,2,3,*]

From the [1]Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, Saarbrücken, Germany; [2]Deutsches Zentrum für Infektionsforschung (DZIF), Standort Hannover-Braunschweig; [3]Department of Chemistry, PharmaScienceHub (PSH), Saarland University, Saarbrücken, Germany; [4]Institut für Lebensmittelchemie, Leibniz Universität Hannover, Hannover, Germany

Reviewed by members of the JBC Editorial Board. Edited by Robert Haltiwanger

Bacteria frequently employ carbohydrate-binding proteins, so-called lectins, to colonize and persist in a host. Thus, bacterial lectins are attractive targets for the development of new anti-infectives. To find new potential targets for anti-infectives against pathogenic bacteria, we searched for homologs of *Pseudomonas aeruginosa* lectins and identified homologs of LecA in *Enterobacter* species. Here, we recombinantly produced and biophysically characterized a homolog that comprises one LecA domain and one additional, novel protein domain. This protein was termed *Enterobacter cloacae* lectin A (EclA) and found to bind L-fucose. Glycan array analysis revealed a high specificity for the LewisA antigen and the type II H-antigen (blood group O) for EclA, while related antigens LewisX, Y, and B, as well as blood group A or B were not bound. We developed a competitive binding assay to quantify blood group antigen-binding specificity in solution. Finally, the crystal structure of EclA could be solved in complex with methyl α-L-selenofucoside. It revealed the unexpected binding of the carbohydrate ligand to the second domain, which comprises a novel fold that dimerizes *via* strand-swapping resulting in an intertwined beta sheet.

Bacterial infections are increasingly threatening as a consequence of antimicrobial resistance development against standard-of-care antibiotics, which have been on the market for many decades. Therefore, new treatment options are needed to fight the inevitable spread of multidrug resistant pathogens rendering current medicines ineffective (1). The World Health Organization has classified, especially, Gram-negative bacteria from the resistance-prone ESKAPE group as outstandingly important due to a lack of treatments with new modes of action, able to circumvent established antimicrobial resistance (2).

For the critical priority Gram-negative bacterium *Pseudomonas aeruginosa*, many new approaches are actively being studied, such as new classes of antibiotics and alternative approaches interfering with bacterial virulence and pathogenicity (1, 3). Among the so-called pathoblockers or antivirulence agents (4, 5), interference with biofilm formation—a major determinant of drug resistance—is of particular interest to restore the pathogen's susceptibility to treatments (6). Numerous bacteria employ carbohydrate-binding proteins, so-called lectins, to adhere to host tissue and establish and maintain biofilms (7–9). It has been demonstrated that the *Pseudomonas* lectins LecA and LecB can be targeted with defined carbohydrates, small glycomimetics, and dendrimers to prevent biofilms and interfere with bacterial virulence (10–21).

Inspired by these data, we searched for orthologs of LecA in other bacteria. We previously identified many LecA orthologs in the genomes of several *Photorhabdus*, *Xenorhabdus*, and *Enterobacter* species and experimentally characterized PllA from *Photorhabdus luminescens*, an insect pathogen (22).

The *Enterobacter cloacae* complex is part of the ESKAPE pathogens infecting humans and comprises numerous species classified within 12 genetic clusters (2, 23, 24). Bacteria associated with this complex are usually found in soil, sewage, and drinking water reservoirs and are also part of the human intestinal microbiota (25). *E. cloacae* and *Enterobacter hormaechei* constitute the most frequently isolated *Enterobacter* species from human clinical specimens (24). Outbreaks of *E. cloacae* and the trigger to change from a commensal bacterium to a virulent one can be sporadic and often happens in the intensive care units of hospitals and causes either localized infections, such as lung, wound, central nervous system, urinary tract, and catheter, or orthopedic implant-associated infections, or systemic infections, such as bacteremia and sepsis (25, 26). It is also well-established that *E. cloacae* is a culprit of outbreaks and fatalities in neonatal intensive care units of hospitals (27). In their virulence arsenal, *E. cloacae* species employ many virulence factors like enterotoxins, α-hemolysin, and thiol-activated pore-forming cytotoxins similar to Shiga-like toxin II after adhesion to epithelial cells (28). In some clinical strains of *E. cloacae*, type III secretion system was also seen to be employed to destroy phagocytes and epithelial cells to facilitate host colonization (29). Further, a type VI secretion system also showed to be instrumental in biofilm formation

* For correspondence: Alexander Titz, alexander.titz@uni-saarland.de, alexander.titz@helmholtz-hzi.de; Jesko Köhnke, jesko.koehnke@foh.uni-hannover.de.

and adherence to epithelial cells (30). Like other Gram-negative bacilli, the virulence of *E. cloacae* also depends on the presence of its outer membrane lipopolysaccharide, which can help in avoiding opsonophagocytosis or initiate an inflammation cascade in the host cell leading to sepsis (31). *E. cloacae* are multidrug resistant due to their chromosomally encoded and induced and/or constitutively expressed AmpC β-lactamase with increased resistance rates following treatment with either β-lactams or first, second, and third generation cephalosporins (31). Fourth generation cephalosporins may be suitable against the AmpC β-lactamase strains if extended-spectrum β-lactamases are not present, in which case a combination therapy of colistin or aminoglycosides and carbapenems in a double regimen is used (31). Due to the prevalence of extended-spectrum β-lactamases and carbapenemases in this species, *E. cloacae* has become the third most common broad-spectrum Enterobacteriaceae involved in nosocomial infections, following *Escherichia coli* and *Klebsiella pneumoniae* (32). For these reasons, the World Health Organization identified carbapenem-resistant Enterobacteriaceae as a critical priority on its list of antibiotic-resistant bacteria in 2017, highlighting the urgent need for the development of new antibiotics. The characterization of LecA homologs in *Enterobacter* spp. as possible antivirulence targets is therefore of interest.

Here, we report the identification, biophysical and structural characterization of the first two-domain ortholog of LecA found in *E. cloacae* subsp. *cloacae* (type strain: ATCC 13047), an important member of the *E. cloacae complex*. This strain was isolated from human cerebrospinal fluid and is the first completely sequenced member of the *E. cloacae* species. It possesses many virulence properties, encodes more than 50 antibiotic resistance genes, and has been extensively studied and used as a reference strain (33–36).

The lectin termed *Enterobacter cloacae* lectin A (EclA) consists of an N-terminal LecA domain and a C-terminal domain reminiscent of carbohydrate binding modules. In this work, we established the L-fucose binding of EclA *via* its C-terminal domain while we did not succeed in identifying a ligand for the LecA-homologous N terminus despite intense efforts. EclA forms homodimers resulting in the presentation of two N termini toward one end and two C termini toward the opposite end in its structure. This orientation suggests a function as a cross-linker for two carbohydrate ligands, one of which remains elusive. The unprecedented high specificity of EclA's C-terminal domain for L-fucosides in mammalian blood group H-type II and LewisA antigens over related A/B-antigens or isomeric LewisX suggests a link to host binding specificity and pathophysiology.

Results

Recently, we identified various orthologs of LecA in the genomes of diverse Gram-negative bacterial species, including the insect pathogens *P. luminescens* and *Xenorhabdus* spp. and the opportunistic human pathogen *Enterobacter* spp. (22). Here, we have further analyzed the sequences of LecA orthologs from the ESKAPE pathogen *Enterobacter* spp. Like LecA, most of them are short proteins comprised of only one LecA domain. Surprisingly, in the genome of the human spinal cord infection isolate of *E. cloacae* subsp. *cloacae*, we identified a fusion protein, which we termed EclA. Its gene *eclA* encodes for the protein EclA (ECL_04191, GenBank: ADF63724.1) composed of 283 amino acids with a molecular weight of 30.9 kDa. The protein sequence alignment of EclA with its orthologs LecA and PllA shows that the N-terminal domain of EclA is similar to the PA-1L (or LecA) family domain present in LecA and PllA (Fig. 1). This domain spans from Trp20 to Glu133, and the amino acid involved only in calcium binding is conserved, while those involved in galactose-binding in LecA and PllA are not conserved. The second, C-terminal domain of EclA is recognized as a carbohydrate-binding domain by bioinformatics sequence analysis tools.

Since EclA is the first example of a LecA lectin domain fused to another domain, we set out to analyze this protein in depth. Full-length EclA was cloned into the pET22 vector for recombinant cytosolic expression of the native protein in *E. coli*. The protein was highly expressed and present in the soluble fraction after cell lysis. The apparent molecular weight of EclA corresponded to the predicted value of 30.9 kDa as determined by denaturing gel electrophoresis (Fig. S1). Despite the fact that LecA and PllA can both be purified on D-galactose-modified affinity resins, EclA did not bind to this resin and analogous purification failed. Therefore, EclA was finally purified from the soluble cell extract fraction by gel filtration on a superdex matrix (Fig. S1).

Carbohydrate-binding specificity

To determine the ligand-binding specificity of EclA, we analyzed its thermal denaturation in the presence of different carbohydrates (Fig. 2). Among a set of 10 different monosaccharides tested, a detectable shift in EclA's melting temperature (T_m = 55.2 °C) was only observed in presence of the deoxyhexose L-fucose (T_m = 57.2 °C), whereas all nine other carbohydrates including D-galactose did not influence EclA's melting point. The result that EclA bound to L-fucose, *i.e.*, 6-deoxy-L-galactose, was surprising as LecA and PllA bind to its hydroxylated enantiomer D-galactose. EclA was subsequently purified on a Sephadex affinity resin modified with L-fucose, and a yield of 30 mg EclA was obtained per liter bacterial culture (Fig. S2).

To gain deeper insights into the carbohydrate-binding specificity, we tested EclA on a glycan array. To this end, EclA was fluorescein-labeled with fluorescein isothiocyanate (FITC) and analyzed on the Consortium for Functional Glycomics' (CFG's) Core H mammalian glycan array with 585 different carbohydrate epitopes (Fig. 3). Binding of EclA was detected for a range of fucosylated oligosaccharides and the monosaccharide L-fucose on the array. In addition, the binding of EclA-FITC to sulfated sialyl-LacNAc was detected, the only bound ligand without a fucose residue.

Careful analysis of the glycan array binding data for EclA revealed that fucosylated oligosaccharide ligands recognized by EclA were highly specific for only the blood group antigens

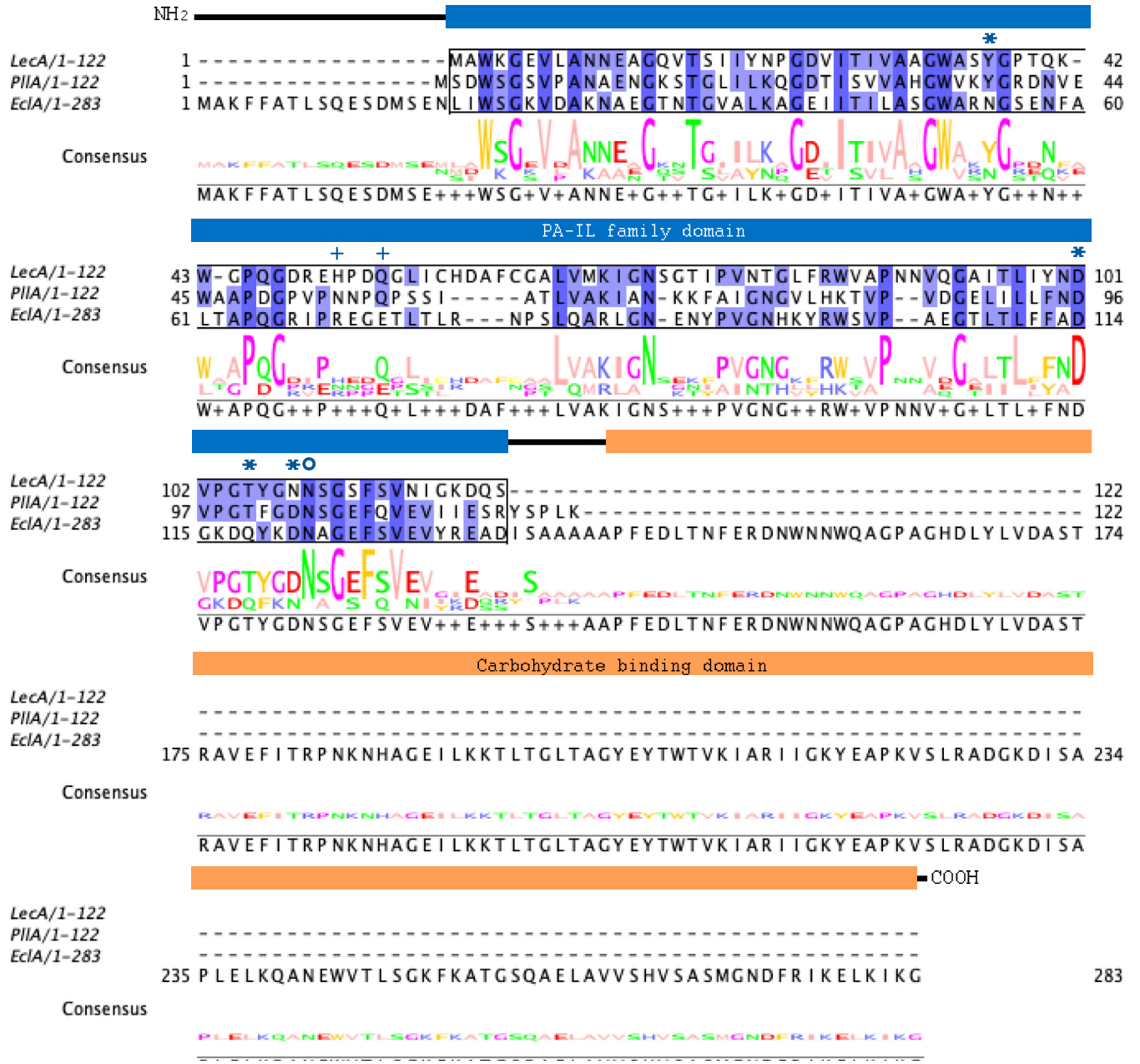


Figure 1. Sequence alignment of LecA, PllA, and EclA. Protein sequences of LecA from *P. aeruginosa* PAO1, PllA from *P. luminescens subsp. laumondii* TTO1, and EclA from *E. cloacae subsp. cloacae* ATCC13047 were aligned in a multiple sequence alignment. The PA-1L (LecA) family domain is indicated as *blue box*, and the second domain of EclA is indicated as *orange box*. Amino acids are colored according to percentage identity (the higher the identity, the darker the color). *Asterisks*: Amino acids involved in ligand and calcium binding; *open circle*: amino acid involved in calcium binding; *crosses*: amino acids involved in ligand binding. EclA, *Enterobacter cloacae* lectin A.

LewisA and the H-antigen on a type II core (Fig. 4). The selectivity of EclA for the LewisA antigen (Fuc-α-1,4-(Gal-β-1,3-)-GlcNAc) over its regioisomeric antigen LewisX (Fuc-α-1,3-(Gal-β-1,4-)-GlcNAc) is particularly remarkable. Additional sialylation was tolerated, and LewisA and sialyl LewisA are generally recognized with high apparent affinity. Interestingly, out of 10 LewisX and 8 sialyl LewisX structures on the array, only ligand no. 24 carrying two additional sulfate residues was strongly recognized. Mono-sulfated epitopes were weakly recognized, and no nonsulfated epitopes were bound. While α-2,3-sialylation of LewisA or LewisX has no influence on the specificity of EclA, an additional fucosylation resulting in LewisB and LewisY, respectively, abolished binding also for LewisB, a congener of LewisA.

The observed selectivity of EclA for the H-type II antigen (Fuc-α-1,2-Gal-β-1,4-GlcNAc) over isomeric H-type I (Fuc-α-1,2-Gal-β-1,3-GlcNAc) is an interesting observation since the structural difference between those epitopes is located rather remote from the recognized fucose moiety, *i.e.*, in the linking position of galactose to the GlcNAc residue. Blood group epitopes A or B are only occasionally recognized, and a clear pattern is absent.

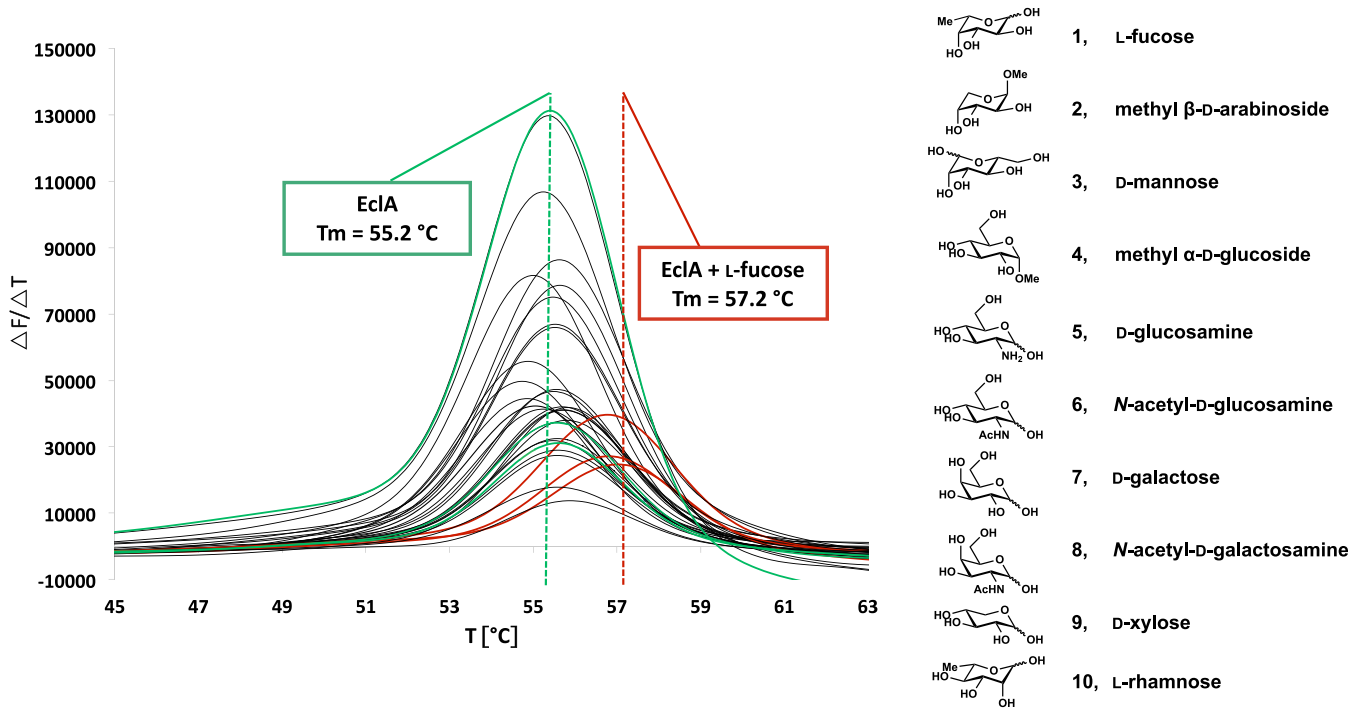


Figure 2. Thermal shift analysis of EclA ligand-binding specificity indicates its binding to L-fucose. EclA, *Enterobacter cloacae* lectin A.

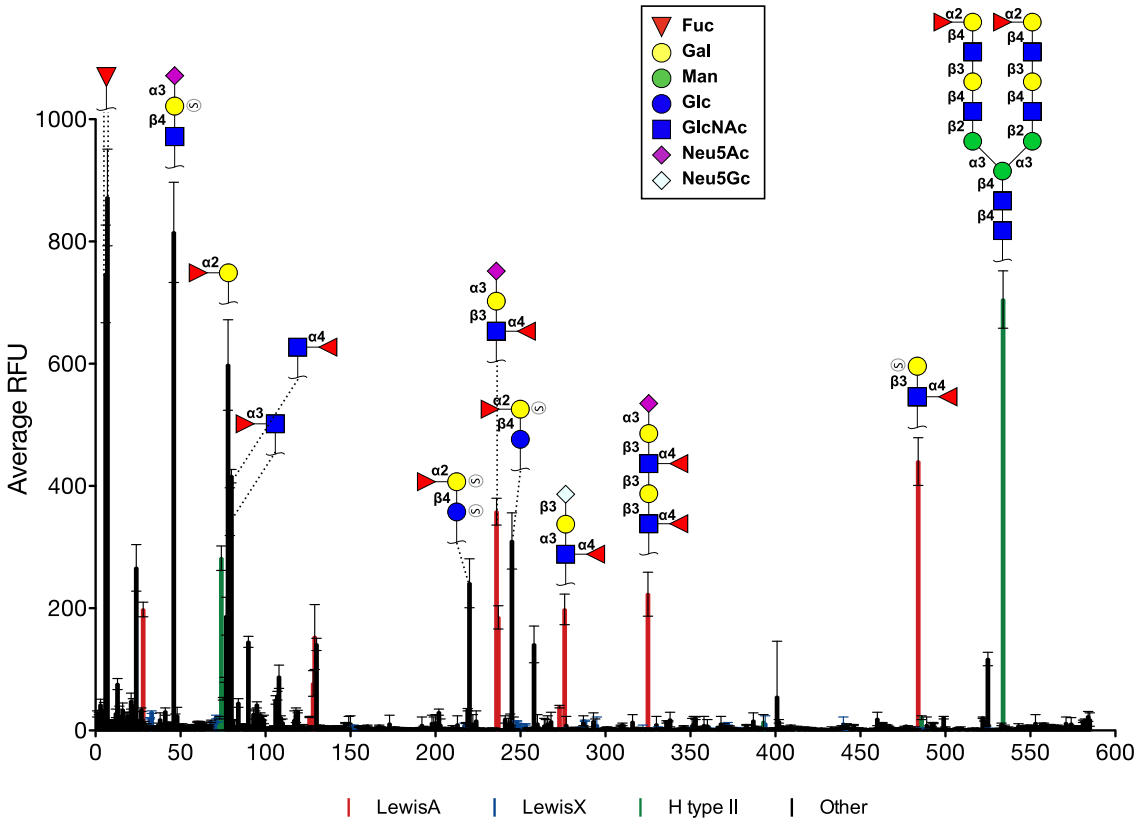


Figure 3. Glycan specificity analysis of fluorescein-labeled EclA on the CFG Core H glycan array. (S) = sulfate. Error bars indicate standard deviation from technical replicates on the glycan array. CFG, Consortium for Functional Glycomics; EclA, *Enterobacter cloacae* lectin A.

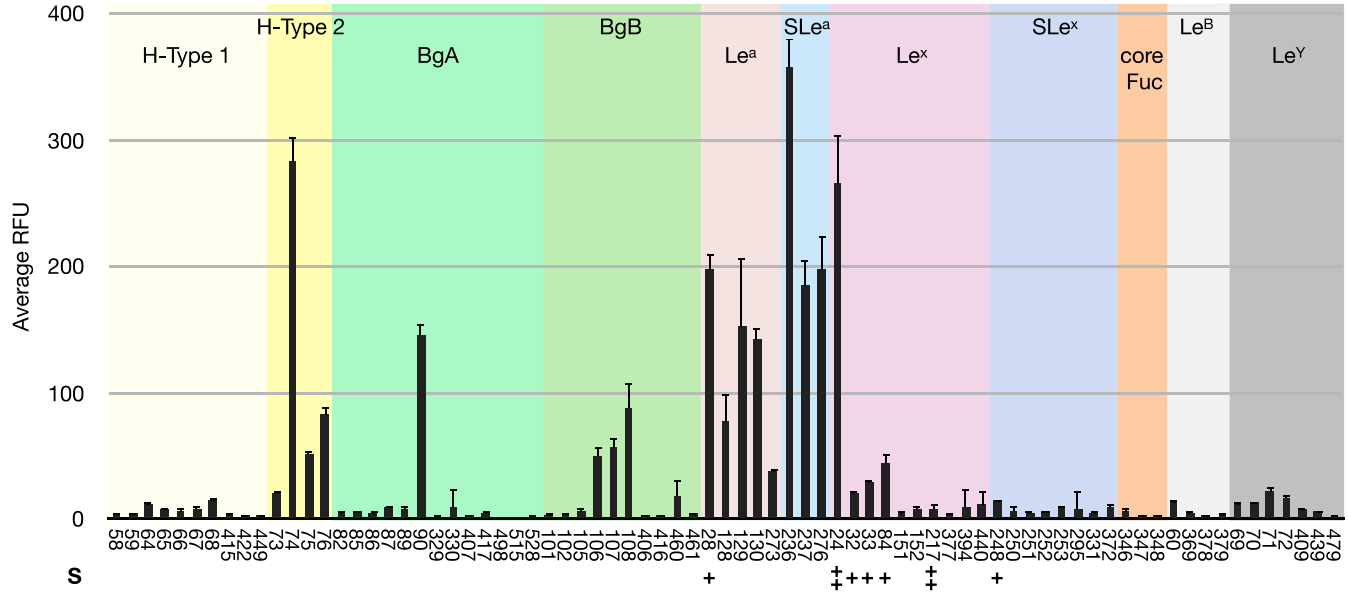


Figure 4. Analysis of blood group epitope specificity of EclA on the CFG Core H glycan array. Only unique and singly/monovalently displayed epitopes are compared here. Single (+) or double (++) sulfation (S) of carbohydrate epitopes is indicated. Error bars indicate standard deviation from technical replicates on the glycan array. CFG, Consortium for Functional Glycomics; EclA, *Enterobacter cloacae* lectin A.

As mentioned, some carbohydrate ligands were recognized by EclA on the array that were sulfated. An analysis of all sulfated ligands on the array revealed the known specificity pattern and 3′-sulfated LewisA (no. 28, 484) and di(6,6′)- and mono(6 or 6′)-sulfated H-type II-like antigens (reducing end GlcNAc replaced by Glc, no. 220, 245, 258) were recognized by EclA. Among all sulfated glycans on the chip, only two more ligands were recognized: the above mentioned disulfated LewisX (no. 24) and the only binder devoid of fucose, 3′-sialyl-6′-sulfo-LacNAc (no. 46) carrying also two negative charges.

Initial crystallization attempts of full-length EclA with a fucoside indicated ligand binding in the C-terminal domain (see below). Therefore, to further dissect the binding specificities of the individual N- and C-terminal domains, we cloned and recombinantly produced them individually in *E. coli*: N-terminally His_6-tagged EclA N terminus (termed EclA-N-tag, amino acids 2–140) and the native EclA C terminus (termed EclA-C, amino acids 140–283).

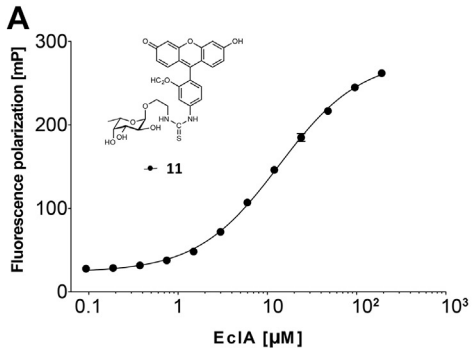
Both protein constructs were subsequently purified to homogeneity using either Ni(II)-NTA affinity resin with imidazole elution or fucosylated-sepharose affinity resin with L-fucose elution (Figs. S3 and S4).

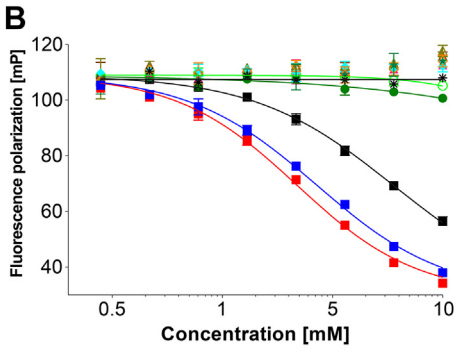
To quantify the binding of EclA with its carbohydrate ligands in solution, we developed a fluorescence polarization-based binding and displacement assay for EclA, in analogy to our previously described competitive binding assays for the bacterial lectins LecA (37), LecB (38, 39), PllA (22), BC2L-A (40), and BambL (41). Since EclA showed fucose-binding, the previously reported fluorescein-linked fucoside **11** (38) was incubated with increasing concentrations of native EclA. Fluorescence polarization was monitored, and fitting the obtained data resulted in K_Ds of 13.7 μM and 17.5 μM (two independent replicates) (Fig. 5*A*). The experiment was also performed with EclA-C, and a similar affinity was determined (K_D = 21.7 μM, Fig. S5). Both constructs, EclA and EclA-C, were subsequently used to screen for inhibitors in a competitive binding assay (Fig. 5, *B* and *C*). The same monosaccharides previously used in the thermal shift assay were tested. In addition, methyl α- and β-L-fucosides were included to assess linkage specificity. The fucose specificity from our previous assays was confirmed, and only L-fucose (IC_{50} = 4.0–4.7 mM), methyl α-L-fucoside (IC_{50} = 1.3–1.4 mM), and its anomer methyl β-L-fucoside (IC_{50} = 1.5–1.8 mM) were inhibitors of, both, EclA and EclA-C. These data suggest that the hydrophobic methyl group of the aglycon is beneficial for binding, while α- or β-linkage shows no influence on binding.

Next, we also tested the human blood group antigens H-type I/II, A-type I/II, B-type-I/II, LewisA, LewisX, LewisB, and LewisY in this competitive binding assay in solution (Fig. 6). The data obtained corresponded to the semiquantitative data obtained from the glycan array analysis: LewisA and H-type II antigens were inhibitors of EclA and EclA-C with IC_{50}s between 390 and 540 μM, whereas their isomeric epitopes LewisX and H-type I showed an at least 7-fold weaker inhibition, and the IC_{50}s were at or above 3.75 mM.

In contrast to the surface binding in the glycan array, the difucosylated Lewis relatives LewisB and LewisY were also binders of EclA and its C-terminal domain in solution (IC_{50}s between 530 and 690 μM). The blood group antigens A and B were only weak inhibitors of EclA and EclA-C with IC_{50}s at or above 3.75 mM, which corresponds to the glycan array. Interestingly, we also observed an effect of type-II over type-I core.

To shed light on the ligand binding preference of the N-terminal domain of EclA, we tested full-length EclA and both individual domains on a glycan array with a more diverse set of





C

	Compound	EclA IC_{50} [mM]	EclA-C IC_{50} [mM]
1	L-fucose	4.0 ± 1.6	4.7 ± 1.1
2	methyl β-D-arabinoside	n.i.	n.i.
3	D-mannose	n.i.	n.i.
4	methyl α-D-glucoside	n.i.	n.i.
5	D-glucosamine	n.i.	n.i.
6	*N*-acetyl-D-glucosamine	n.i.	n.i.
7	D-galactose	n.i.	n.i.
8	*N*-acetyl-D-galactosamine	n.i.	n.i.
9	D-xylose	n.i.	n.i.
10	L-rhamnose	n.i.	n.i.
12	methyl α-L-fucoside	1.4 ± 0.6	1.3 ± 0.5
13	methyl β-L-fucoside	1.8 ± 0.9	1.5 ± 0.1

Figure 5. EclA ligand binding analyzed by fluorescence polarization. *A* and *B*, direct binding of EclA to FITC-labeled fucoside **11** (*A*) and competitive inhibition (*B*) using a set of monosaccharides and methyl glycosides **1** to **10**, **12**, **13**. *C*, IC_{50}s and standard deviations for the inhibition of EclA or EclA-C derived from three independent titrations of triplicates each. Representative binding and inhibition data from one technical triplicate are depicted in (*A*) and (*B*). FITC, fluorescein isothiocyanate; EclA, *Enterobacter cloacae* lectin A; EclA-C, EclA C terminus.

carbohydrate epitopes than the CFG mammalian glycan array. To this end, we labeled all three protein constructs with Cy3 and analyzed them on the Semiotik glycan array. This array contains 400 glycans and 200 bacterial polysaccharides. Full-length EclA and its C-terminal domain showed binding specificity for fucose, LewisA, and H-type antigens, while no ligand specific to the N-terminal domain of EclA could be identified, leaving its function enigmatic (Fig. S6).

LecA and PllA are single domain proteins and form homotetrameric assemblies in solution, which are also observed in their crystal structures (22, 42). In contrast, their ortholog EclA consists of two separate domains, with one corresponding to a LecA domain. Thus, the elucidation of the oligomeric states of EclA is of interest. To this end, we analyzed EclA, EclA-N-tag, and EclA-C by dynamic light scattering (DLS, Fig. 7). The gene of EclA possesses two ATG codons at the 5′ end of the coding sequence which presumably led to the two protein species with very small apparent molecular weight differences as detected by SDS-PAGE (Figs. S1, S2 and S7). To test a homogeneous protein, we cloned the

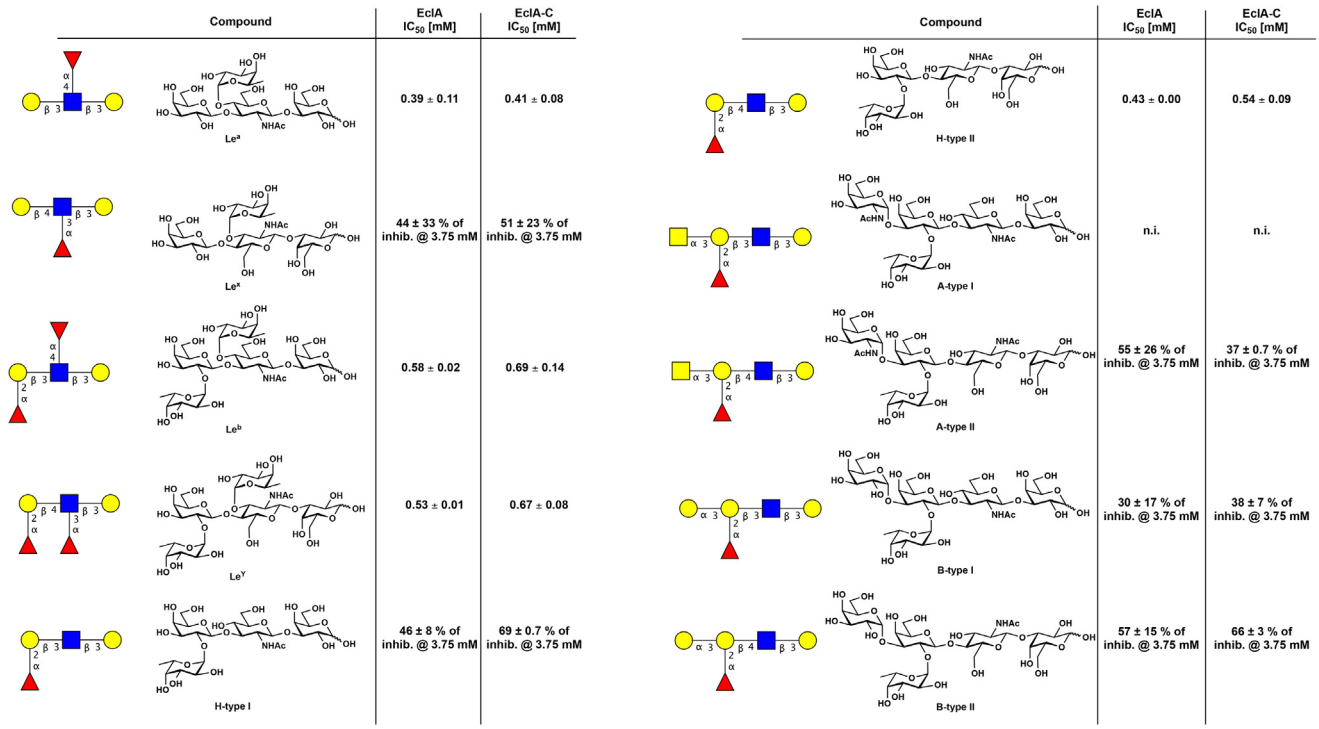


Figure 6. Competitive binding of human blood group antigens to EclA and EclA-C. IC_{50}s and standard deviations are from three independent titrations of triplicates each. EclA, *Enterobacter cloacae* lectin A; EclA-C, EclA C terminus.

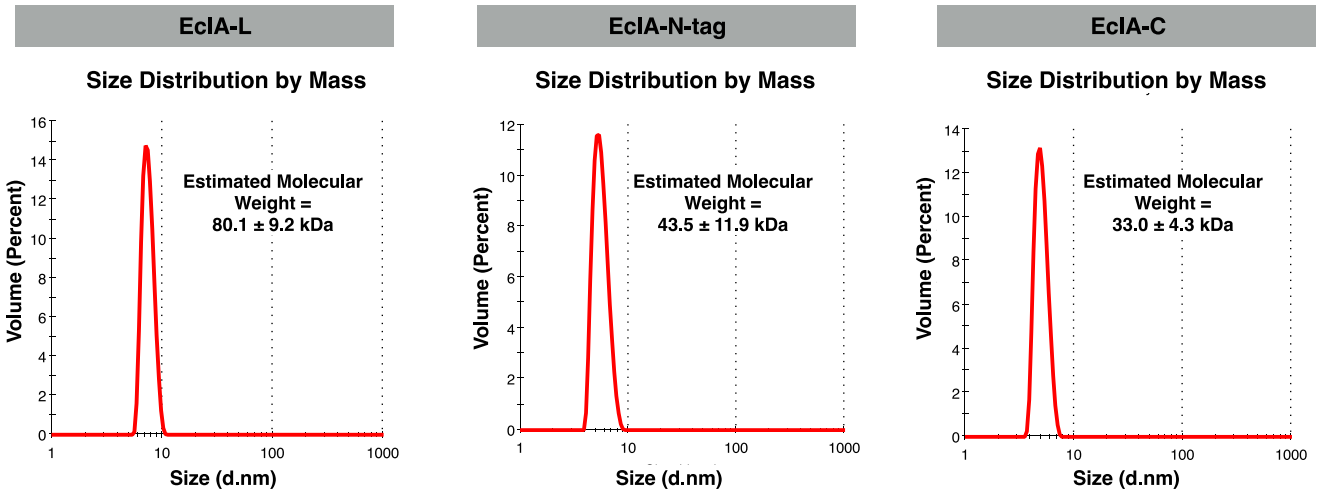


Figure 7. Dynamic light scattering analysis of EclA-L (EclA-M14A), EclA-N-tag, and EclA-C. EclA-L, EclA-M14 A; EclA-N-tag, EclA N terminus; EclA-C, EclA C terminus.

single amino acid variant of EclA-M14A, termed EclA-L, and expressed and purified it as reported for EclA (Fig. S7). The obtained estimated molecular weights of the protein species in solution showed an observed molecular weight of 80.1 kDa for full-length EclA-L (MW = 30.9 kDa), 43.5 kDa for the N-terminal domain (EclA-N-tag, MW = 16.0 kDa), and 33.0 kDa for the C-terminal domain (EclA-C, MW = 15.9 kDa). The observed molecular weight of EclA-C has a rather narrow distribution and suggests dimer formation. EclA-L and EclA-N-tag both have a larger dispersion of molecular weights with a value close to 2.5-fold of their calculated molecular weight. DLS measures the hydrodynamic radius and the molecular mass of the proteins is calculated assuming an ideal spherical particle, which explains these deviations. Taken together, all three constructs showed experimental values for their molecular weight between 2.1- and 2.7-fold of their calculated molecular weights, suggesting dimer formation in all three cases.

Crystallography

In order to shed light on structure and ligand binding of the two EclA carbohydrate recognition domains, crystallographic studies were first carried out on full-length EclA. We obtained crystals in several conditions that were all of similar morphology and diffracted well to beyond 2.5 Å. The datasets could be processed in $P2_1$ with excellent statistics and showed no obvious abnormalities. Molecular replacement with Phaser (43) using a monomer of the lectin PllA (PDB ID 5ofz), which is homologous to the N-terminal half of EclA (EclA-N), resulted in a solution containing a dimer of the N-terminal domain, with the expected dimeric arrangement of lectin domains and a TFZ score of 12.3, indicating a correct solution. Unexpectedly, we were unable to resolve the C-terminal domain of EclA, even though the crystal packing of the N-terminal domains clearly left sufficient space for it. Extensive trials to refine the full structure failed, and we thus resorted to treating the N- and C-terminal domains of EclA separately.

The crystal structure of EclA-N-tag was determined in space group $P2_1$ and revealed two dimers in the asymmetric unit, which matched the partial solution obtained for the full-length protein (Fig. 8). The overall structure of EclA-N-tag is highly similar to the structure of the *P. luminescens* lectin PllA (Cα rmsd of 0.47 Å over 93 atoms; Fig. 8*B*), with each monomer adopting a jelly-roll type β-sandwich fold as previously described (22). The crystals were obtained in the presence of 60 mM calcium chloride dihydrate in the crystallization buffer, revealing unambiguous electron density for the canonical calcium ion at the putative carbohydrate-binding site in each of the four protomers in the asymmetric unit. In all protomers, the calcium ion is coordinated by the side chain oxygens of Asp114, Asp121, Asn122, and the main chain oxygens of Asn54 and Gln118 (Fig. S8*A*). The coordination pattern mirrors those observed in other known calcium-binding lectins (Fig. S8, *A* and *D*).

One structural difference that stands out between the two structures is the size of the ligand-binding pockets: it is much wider in EclA-N than in PllA, which is the result of a movement of a loop (residues 70–82; Fig. 8, *B* and *C*). This could be a crystallization artefact, as H-bond interactions between this loop and symmetry mates can be observed. An overlay of all four EclA-N molecules shows an almost identical orientation of this loop, which may indicate a biological significance (Fig. S8*B*). This is supported by the involvement of residues present on the corresponding loops of other lectins in ligand-binding (Fig. S8*C*). However, in the absence of any known carbohydrate ligand, it is difficult to ascertain the role of the observed wide pocket and the loop in ligand-binding specificity of EclA-N.

The C-terminal construct, EclA-C, purified in the presence of excess calcium ions, yielded very well diffracting crystals belonging to space-group $P2_12_12_1$. Molecular replacement was unsuccessful; therefore, we expressed and crystallized selenomethionine-labeled EclA-C. We were able to obtain an initial solution with two dimers in the asymmetric unit.

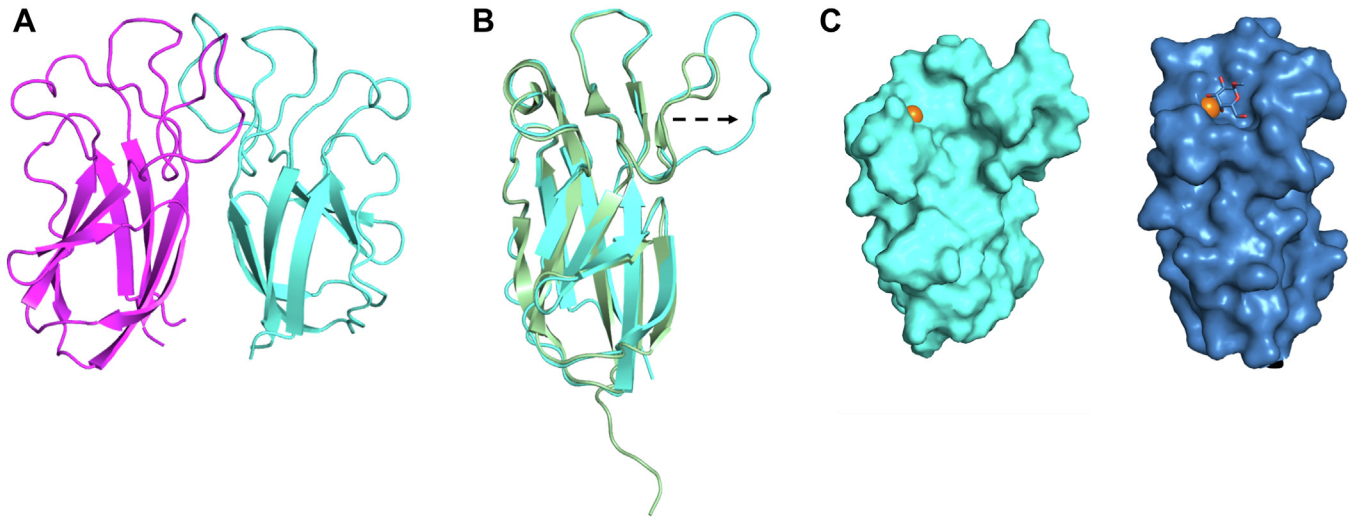


Figure 8. Crystal structure of the N-terminal domain of EclA (EclA-N-tag) and comparison with its ortholog PllA. *A*, dimer of EclA-N-tag; *B*, superposition of the monomers of EclA-N-tag and PllA indicate the extended loop in EclA leading to opening of the putative binding site. *C*, surface illustration of the crystal structure of the N-terminal domain of EclA (EclA-N-tag, *left*) and comparison with its ortholog PllA with bound galactose (*right*). Bound ligand is shown as *stick* and Ca^{+2} ions as *orange spheres*.

Mirroring the issues of the full-length EclA construct, unambiguous, clear electron density could only be observed for one of the two dimers. We hypothesized that the intrinsic flexibility of the protein was the culprit. After collecting many datasets with different additives and ligands, we were able to obtain a dataset of EclA-C in complex with methyl α-L-selenofucoside (MeSeFuc) and were able to refine this structure fully. In the structure, this ligand appears to limit the flexibility of the dimer by cross-linking symmetry-mates *via* the carbohydrate ligand (MeSeFuc interacts with Glus178, Ile214, Ile215, and Arg275, Fig. S9).

In each of the promoters in the asymmetric unit, MeSeFuc is present in a small cleft, stabilized by a number of hydrophobic, metal, and H-bond interactions (Fig. 9). The cis-diol of the fucoside coordinates a metal that is most likely calcium as the protein was stored in a 2 mM calcium chloride buffer. In EclA-C, the calcium ion is tightly coordinated by Ile180, Thr181, Ser269, Gly271, and Asp273 (Figs. 9 and S10*A*). In addition to the calcium coordination, the 4-OH of the ligand is in hydrogen bonding distance of the backbone NH of Gly271 and the side chain carboxylate of Asp273. The carbohydrate residue forms one additional hydrogen bond *via* its 2-OH and a bound water molecule with the side chain of Ser269. Furthermore, extended hydrophobic contacts of the aglycon and the glycosidic selenium are formed with Met270 and Tyr218. The exocyclic methyl group of the fucoside is

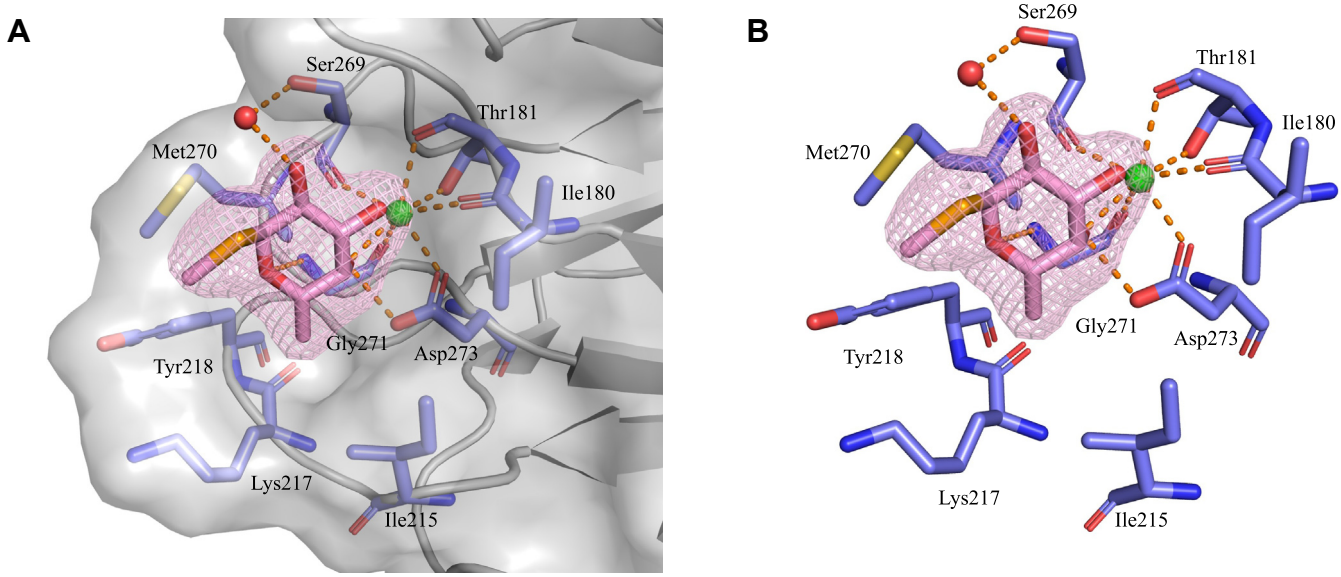


Figure 9. Crystal structure of EclA-C in complex with methyl α-L-selenofucoside. *A*, surface display of the protein illustrating the ligand's interactions in the binding pocket. *B*, coordination of the protein-bound calcium ion and hydrogen-bonding interactions of the fucoside. Polder map of the fucoside is contoured at 4 σ. EclA-C, EclA C terminus.

accommodated by a hydrophobic pocket formed by Ile215, Lys217, and Tyr218.

The dimer partners of EclA-C interact mainly *via* their N termini through a β-strand swap, which results in a twisted, antiparallel β-zipper (Fig. 10, *A* and *B*). This mode of interaction presumably results in a very unusual arrangement in the full-length protein (Fig. 11*A*). However, in the absence of structural data for the full-length EclA, it remains uncertain whether the β-strand swap is a crystallization artefact or not. Interestingly, EclA's AlphaFold model shows low confidence in this region (Fig. S13). A search for structural homologs of EclA-C using the DALI server (44) revealed the carbohydrate-binding module (CBM) CBM22-1 of *Clostridium thermocellum* endo-1,4-β-D-xylanase (Xyn10B, PDB ID 2W5F) to be the closest structural homolog (Z-score of 9.8; Fig. 10*C*). Akin to CBM22-1, each monomer of the EclA-C structure consists of a β-sandwich fold, comprised of two antiparallel β-sheets made from 7 β-strands (β4, β5, β6, β7, β9, β10, and β11; CBM22–1 numbering). In EclA-C, β4 is extended and connected to an additional β-strand (β3) that is not part of the β-sandwich (Figs. 10, *B* and *C*; S10*B*).

Despite the overall similar structure, the observed mode of Ca^{2+} ion and ligand binding in the structure of EclA-C is different compared other CBMs; the binding of the carbohydrate ligands is located at different areas of the protein and in the EclA structure primarily stabilized by the Ca^{2+} ion (Figs. 9, 10*D* and S11) rather than a network of H-bonds as observed, *e.g.*, CBM67, CBM60, and CBM36. As a result, the removal of Ca^{2+} had a drastic effect of the EclA-C ligand binding. Moreover, CBMs often contain a second, structural Ca^{2+} ion, which is absent in the EclA-C structure.

Bioinformatics

Following the discovery of EclA's unique structure and ligand-binding specificity, we next analyzed the abundance of EclA in other *Enterobacter* species and strains. The protein sequence of EclA from *E. cloacae* subsp. *cloacae* ATCC 13047

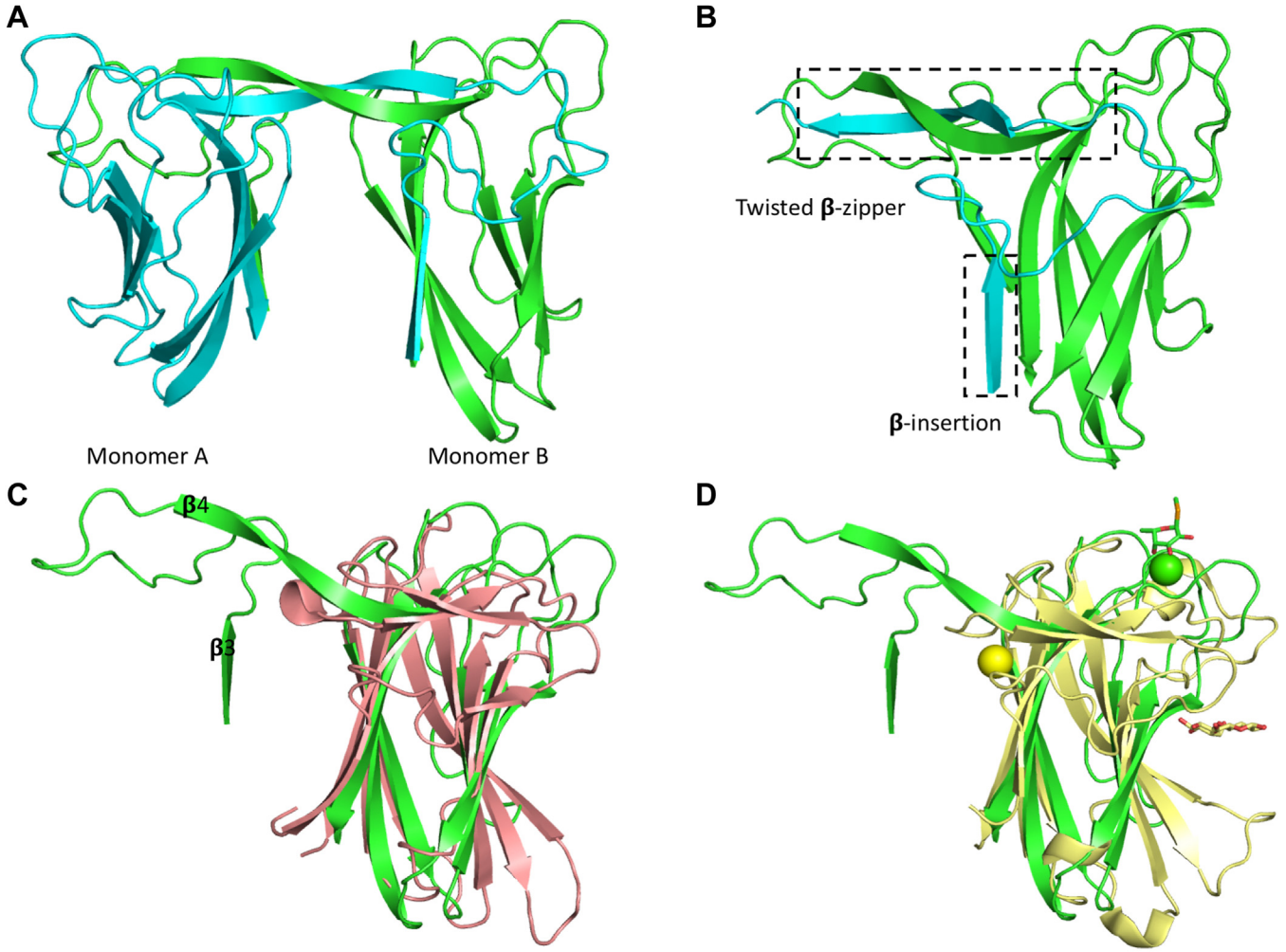


Figure 10. Structural analysis of the C-terminal domain of EclA (EclA-C). *A*, dimer comprising the asymmetric unit is shown as ribbon model with the monomers colored *cyan* and *green*. *B*, the twisted zipper formation and the additional β-insertion observed in the crystal structure of EclA-C. Only β-sheet (3 and 4) of monomer A are shown. *C*, superposition of EclA-C structure (*green*) with CBM22-1 (PDB ID: 2W5F; *salmon*). C α RMSDs approx. 3.6 Å over the entire length of the protein (monomer B). *D*, superposition of EclA-C structure (*green*) with CBM22-2 (PDB ID: 4XUQ (*yellow*)). Ligands are shown as *sticks* (methyl selenofucoside, *green* and xylotriose, *yellow*) and the Ca^{+2} ions as *spheres* (EclA-C, *green* and CBM22–2, *yellow*).

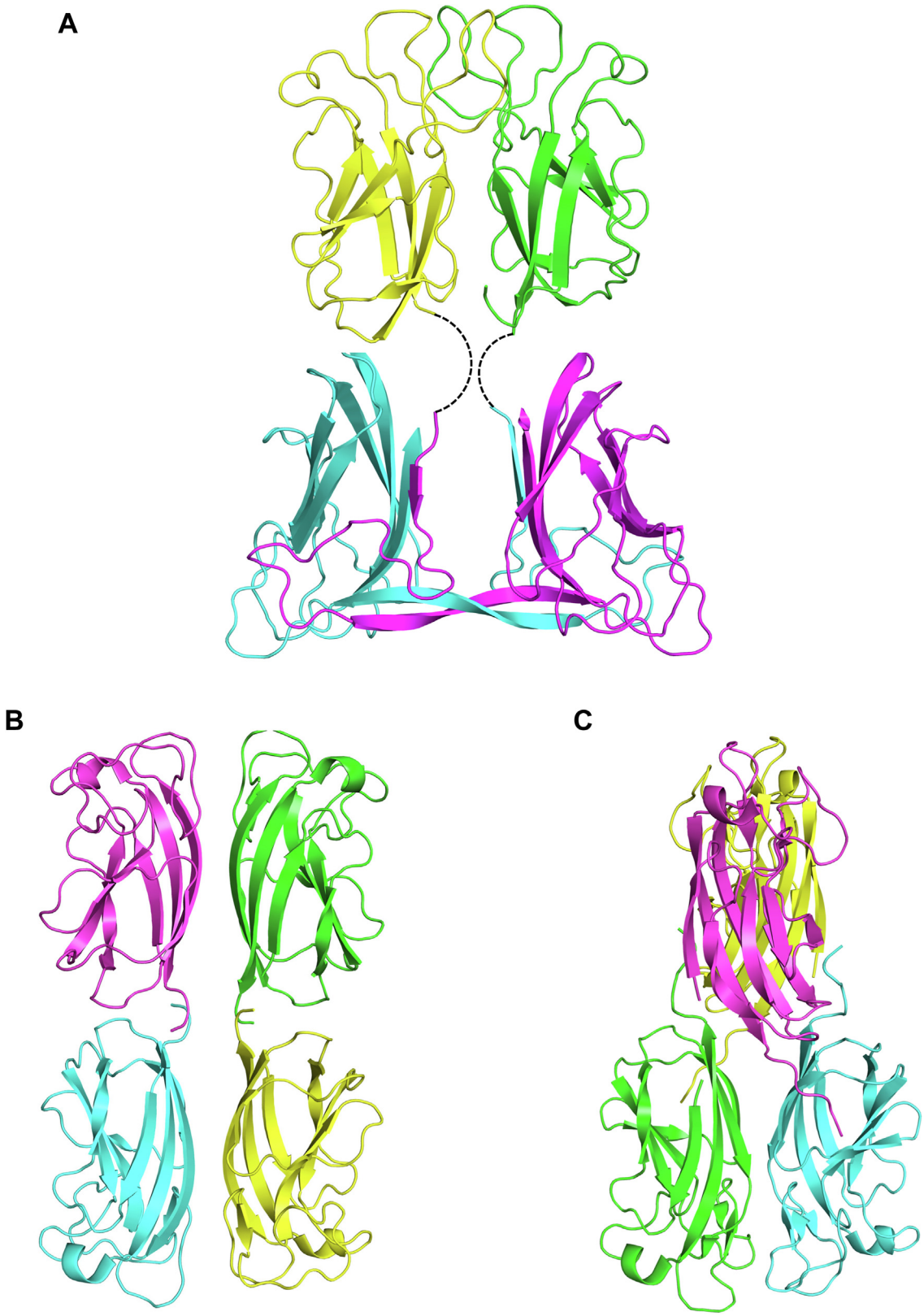


Figure 11. Comparison of quaternary structures of homodimer EclA with homotetramers LecA and PllA. *A*, model of the quaternary structure of EclA using the separate crystal structures of EclA-C and EclA-N-tag. *Dashed lines* indicate the missing 11 amino acids between N and C terminus. *B*, tetramer of LecA in its rectangular orientation (pdb code 4LKD (42)). *C*, tetramer of PllA with both dimers twisted by 90° to each other (pdb code 5OFZ (22)). EclA, *Enterobacter cloacae* lectin A; EclA-C, EclA C terminus; EclA-N-tag, EclA N terminus.

was used as template in a tblastn search in *Enterobacter* genomes. Nearly 50 homologs of EclA have been identified and are shown in a multiple sequence alignment (Fig. 12).

Interestingly, three different groups of homologs could be identified: (i) in *E. cloacae* strains, proteins of unknown function matching full-length EclA were identified (sequence identities: 99–85%) and in *E. ludwigii* strains, proteins harboring either one (ii) or the other (iii) of the two EclA domains were identified, but sequences covering both domains simultaneously were absent (Fig. 12*B*).

Moreover, each of the three complete *E. ludwigii* genomes harbors three proteins that match to one EclA domain (Fig. 12*B*): (i) one single domain protein of unknown function annotated as a PA-1L family protein with 45 to 39% sequence identities to EclA-N; (ii) a protein of unknown function, annotated as lectin harboring two domains, of which the N-terminal domain shares 42 to 38% sequence identity with EclA-N; and (iii) another protein of unknown function annotated as carbohydrate-binding protein harboring two domains, of which the C-terminal domain shares 35 to 33% sequence identity with EclA-C. The C-terminal domain of the 'lectin' hits found in *E. ludwigii* genomes is furthermore similar to its N-terminal domain, while the N-terminal domain of the 'carbohydrate-binding protein' is a von Willebrand Factor type A domain, named after its presence in von Willebrand Factor —a large glycoprotein found in blood plasma. Interestingly, the genes found in *E. ludwigii* genomes encoding the three proteins that share either the LecA domain or the C-terminal domain of EclA are in direct proximity to each other on the genome.

In a second tblastn search on orthologs of the EclA C-terminal domain, the taxid Enterobacteriaceae was excluded. The results showed mainly *Pseudomonas* genomes (ca. 83%, but not *P. aeruginosa*) and some *Serratia* genomes (ca. 17%) carrying one gene, which is of unknown function but annotated as "carbohydrate-binding protein" (*e.g.*, *Pseudomonas asplenii* strain ATCC 23835, LT629777.1, data not shown). The C-terminal domain of this protein shares a sequence identity with the C-terminal domain of EclA of 38 to 26%.

Discussion

Enterobacter spp. belong to the critical ESKAPE pathogens. In this work, we identified a new lectin, termed EclA, from *E. cloacae* as an ortholog of LecA from *P. aeruginosa*. Interestingly, EclA consists of two domains, an N-terminal LecA/PA-IL domain and a C-terminal carbohydrate-binding domain. The N-terminal domain is also homologous to PllA from *P. luminescens*, and both proteins, LecA and PllA, bind D-galactosides as ligands. To analyze EclA structure and ligand binding, we recombinantly expressed the protein and a first purification attempt on a galactose-affinity matrix surprisingly failed. Subsequent thermal shift analysis of EclA purified by gel filtration unexpectedly indicated the thermal stabilization of EclA in presence of L-fucose.

The binding specificity of EclA was analyzed on the mammalian glycan array of the CFG. Among the approximately 600 mammalian glycans, a highly specific binding preference of EclA could be established. In general, fucose-containing ligands were recognized by EclA, with an unusually high specificity for LewisA and H-type II blood group antigens. Their isomeric structures LewisX and H-type I were not bound. In addition, larger oligosaccharide analogs of LewisA and H-type II carrying additional fucose or Gal/GalNAc residues, such as LewisB and blood group A and B antigens, were also not bound. Interestingly, one additional oligosaccharide was bound by EclA that did not have a fucose residue but is negatively charged due to sulfation and sialylation.

We then cloned and expressed the genes of the individual domains of EclA to analyze their binding properties independently. The fucose-binding was unequivocally assigned to the C-terminal domain. Solution phase binding preferences have been established by fluorescence polarization using a fluorescence-labeled fucoside and various ABO and the Lewis human blood group antigens. The competitive binding assay confirmed the high specificity of EclA for LewisA over isomeric LewisX and for the H-type II antigen over isomeric H-type I antigen oligosaccharides.

In an attempt to identify the ligand of the N-terminal domain, EclA and its two individual domains were then further analyzed on a complementary glycan array from Semiotik comprising approximately 400 glycans and 200 polysaccharides. In general, this array also reproduced the established ligand binding preferences determined for the full-length EclA. The Semiotik array further confirmed EclA's preference for fucose, LewisA, and H-type II antigens located at the C-terminal domain, while unfortunately no specific ligand could be identified for the N-terminal domain, the ortholog of the galactose-binding relatives LecA and PllA.

To uncover the protein's structure, we then crystallized EclA first as the full-length protein and then both individual domains. The structure of the full-length protein crystals could unfortunately not be solved for both domains, in contrast to the structures of the individual domains. A small fucoside ligand was coordinated *via* its 3,4-dihydroxy cis-diol motif *via* one calcium ion by the C-terminal domain. Additional hydrogen bonds and lipophilic contacts of the ligand with EclA explain the binding specificity for fucose-containing ligands. Interestingly, the C-terminal domain formed dimers in the crystal with a very unusual dimerization mode *via* their N termini through a β-strand insertion and a twisted antiparallel β-zipper.

The LecA domain located at the N terminus of EclA also formed dimers in the crystal structure with very high overall similarity to PllA/EclA. The major difference concerns a loop located in the binding site of LecA and PllA that participates in galactoside binding in both proteins. This loop is extended in EclA, leading to a large opening of the site. In addition, three arginines form an extended positively charged patch at the surface adjacent to the conserved calcium ion that is mediating galactose binding in LecA and PllA. This positively charged area might explain the binding to several charged carbohydrates that were not recognized by the C-terminal domain as their neutral congeners (*e.g.*, sulfated LewisX or sulfated sialyl LacNAc),

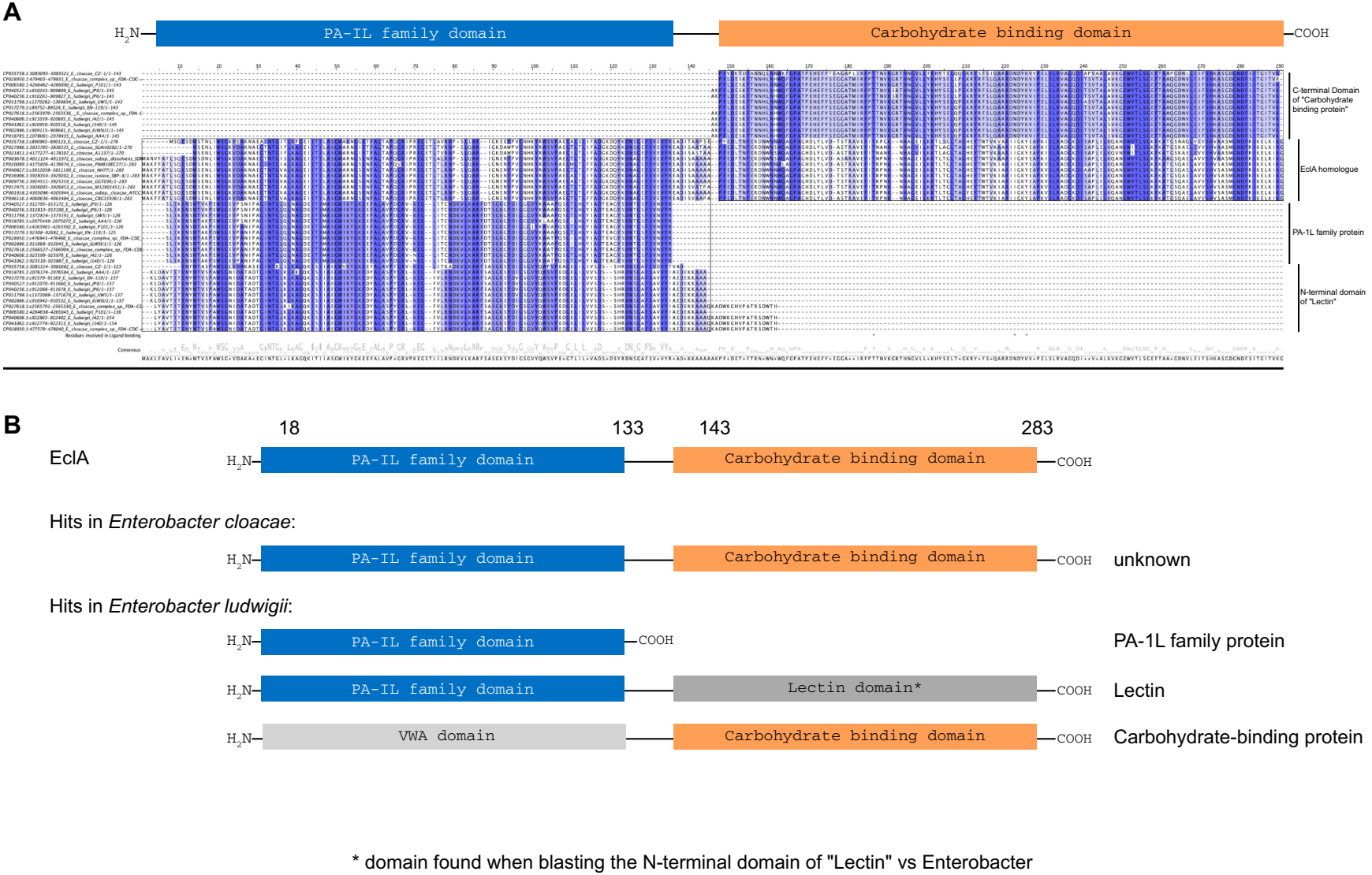


Figure 12. Multiple sequence alignment of EclA homologs in the *Enterobacter* taxid. *A*, protein sequences of hits from a tblastn search were aligned in a multiple sequence alignment. PA-1L (LecA) family domain is indicated as *blue box*, and the C-terminal domain of EclA is indicated as *orange box*. Amino acids are labeled according to percentage identity (the higher the identity, the darker the color). For illustration purposes, found hits were truncated, and only the sequence parts matching EclA were included in the alignment. *B*, schematic representation of proteins found with the BLAST search. In *E. cloacae* genomes, only genes of unknown function were found. In *E. ludwigii* genomes, three proteins were identified that match either the N-terminal or the C-terminal domain of EclA: an annotated lectin, an annotated carbohydrate-binding protein, and a member of the PA-1L protein family. EclA, *Enterobacter cloacae* lectin A.

although not all charged carbohydrates were recognized, and a clear understanding of these observed binding patterns is still missing. Further analysis for the identification of the binding specificity of the EclA N-terminal domain remains a challenge.

Lectins are often multivalent to increase binding strength through avidity. The spatial presentation of binding sites in lectins is important for their function. For surface binding, lectins present their carbohydrate-binding sites on the same side as in Shiga toxin, which invades gut epithelial cells after binding to glycolipids on a cell (45), whereas the interaction with two or more partners involves the presentation of their binding sites in opposing directions, which allows cross-linking of glycosylated binding partners as observed by the native hemagglutination of red blood cells for many lectins and demonstrated by atomic force microscopy for the crosslinking of artificial glycosystems by LecA (46).

Our data from X-ray crystallography and DLS strongly suggest that in EclA the two domain homodimers face in opposite directions, reminiscent of LecA and PllA. However, since LecA and PllA are homotetramers assembled as dimers of dimers of the same short single domain polypeptide, they present four identical binding sites. In contrast, EclA is a dimer of a two-domain polypeptide chain resulting in the presentation of two identical domains to each side. This orientation suggests the cross-linking of two distinct binding partners by EclA, one of which has been elucidated in this work, *i.e.*, the blood group antigens LewisA and H-type II.

In summary, we have presented a novel lectin termed EclA from the opportunistic human pathogen *E. cloacae*. The two-domain protein forms homodimers with an unusual structural assembly. While its C-terminal domain is specific for LewisA and H-type II oligosaccharides, ligand binding specificity of its LecA-homologous N-terminal domain remains enigmatic. The ortholog LecA is a known virulence factor for infections with *P. aeruginosa* and has been shown to bind to host glycolipids for cell entry. LecA is furthermore a key factor for biofilm formation. Thus, LecA has become a drug target for fighting infections with the high priority pathogen *P. aeruginosa* (12, 19). Whether EclA has an analogous role in infection of the ESKAPE member *E. cloacae* remains to be elucidated. Still, its mode of action will be more complex due to the two distinct domains.

Enterobacter species are frequently isolated from neonatal sepsis (24). Blood groups have been widely analyzed, and these cell surface carbohydrates are primary receptors for infections by viruses and bacteria (47, 48). Interestingly, LewisA is overrepresented in neonates. It is known that newborns are LewisA positive, and during childhood, most individuals develop into LewisB positive, a ligand that is recognized with much weaker affinity by EclA. LewisA is strongly expressed in the gastrointestinal tract, the urinary tract, and in the lungs; these are the prime loci for infection with *E. cloacae*. In contrast, epitopes of the ABO blood group system are expressed broadly across the body. Interestingly, Boral *et al.* reported that "the group O neonates had significantly more sepsis due to *E. cloacae*" (49). Therefore, EclA could potentially be a link between *Enterobacter* infections and neonatal sepsis, which merits further analysis. Furthermore, the type I core is present on tissue oligosaccharides, while type II core is found in secreted soluble blood group antigens (50). Possibly, EclA evaded its inhibition by soluble oligosaccharides and evolved specificity for the tissue resident type I core structure for efficient infection.

Experimental procedures

Methyl α-L-fucoside (**12**) and methyl β-L-fucoside (**13**) were purchased from Carbosynth. Methyl α-D-glucoside (**4**), *N*-acetylgalactosamine (**8**), *N*-acetylglucosamine (**6**), and D-xylose (**9**) were obtained from Sigma Aldrich Chemie; L-fucose (**1**), D-mannose (**3**), and D-galactose (**7**) were obtained from Dextra Laboratories; D-glucosamine (**5**) was obtained from Calbiochem; methyl β-D-arabinoside (**2**) was obtained from TCI Deutschland GmbH; L-rhamnose (**10**) was obtained from AppliChem; Histo-blood group antigens were purchased from Elicityl OligoTech.

Fluorescent ligand FITC-fucoside **11** was synthesized as described (38). Methyl α-L-selenofucoside was synthesized according to the previously published protocols from fucose (51, 52).

Cloning, expression, and purification of recombinant EclA, EclA-L, EclA-N-tag, and EclA-C

The synthetic gene of EclA (ECL_04191, GenBank: ADF63724.1) was cloned tag-free into the vector pET22b, and the resulting plasmid was purchased from Eurofins Genomics. The gene sequences of EclA-N-tag and EclA-C were amplified by PCR using Phusion polymerase (New England Biolabs) with the designed primers containing restriction sites and N-terminal His6-tag sequence if needed as listed in Table S1. The protein sequence of EclA is MAKFFATLSQESDMSENLIWSGKVDAKNAEGT NTGVALKAGEIITILASGWARNGSENFALTAPQGRIPRE GETLTLRNPSLQARLGNENYPVGNHKYRWSVPAEGTLT LFFADGKDQYKDNAGEFSVEVYREADISAAAAAPFEDLT NFERDNWNNWQAGPAGHDLYLVDASTRAVEFITRPNK NHAGEILKKTLTGLTAGYEYTWTVKIARIIGKYEAPKVS-LRADGKDISAPLELKQANEWVTLSGKFKATGSQAELAV VSHVSASMGNDFRIKELKIKG (283 aa, MW 30.9 kDa, calculated pI = 6.54). The protein sequence of EclA-L, where a second possible start codon at position 14 is mutated (M14A), in order to have only one isoform of EclA, is MAKFFATLSQESDASENLIWSGKVDAKNAE GTNTGVALKAGEIITILASGWARNGSENFALTAPQGRIP REGETLTLRNPSLQARLGNENYPVGNHKYRWSVPAEGT LTLFFADGKDQYKDNAGEFSVEVYREADISAAAAAPFED LTNFERDNWNNWQAGPAGHDLYLVDASTRAVEFITRP NKNHAGEILKKTLTGLTAGYEYTWTVKIARIIGKYEAPK VSLRADGKDISAPLELKQANEWVTLSGKFKATGSQAEL AVVSHVSASMGNDFRIKELKIKG (283 aa, MW 30.9 kDa, calculated pI = 6.54). The protein sequence of EclA-N-tag, based on EclA-L (see below), is MHHHHHHAKFFATLS-QESDASENLIWSGKVDAKNAEGTNTGVALKAGEIITILAS GWARNGSENFALTAPQGRIPREGETLTLRNPSLQARL

GNENYPVGNHKYRWSVPAEGTLTLFFADGKDQYKDNA GEFSVEVYREADISAAA (146 aa, MW 16.0 kDa, calculated pI = 5.95), and the protein sequence of EclA-C is MAAP-FEDLTNFERDNWNNWQAGPAGHDLYLVDASTRAVEFIT RPNKNHAGEILKKTLTGLTAGYEYTWTVKIARIIGKYEAP-KVSLRADGKDISAPLELKQANEWVTLSGKFKATGSQAEL AVVSHVSASMGNDFRIKELKIKG (144 aa, MW 15.9 kDa, calculated pI = 8.8).

After digestion of the expression vector pET22b(+) (Novagen) and the PCR product with NdeI and BamHI (New England Biolabs), ligation of the insert was performed using TakaRa DNA ligation kit (TAKARA Bio Inc.) resulting in plasmid pET22b-*eclA-N-tag* and pET22b-*eclA-C*. The sequences were confirmed by sequencing (GATC Biotech) with primers T7 promotor and T7 terminator (Table S1).

For expression, pET22b-*eclA*, pET22b-*eclA-N-tag*, and pET22b-*eclA-C* were transformed into chemically competent *E. coli* BL21(DE3), and the expression strain was selected on LB agar supplemented with ampicillin (100 μg mL^{-1}). LB supplemented with ampicillin (100 μg mL^{-1}) (2 L) was inoculated with a preculture and grown at 37 °C and 180 rpm to A_{600} of 0.5 to 0.6. Expression was induced by addition of IPTG (0.15 mM final concentration), and bacteria were then further cultured for 4 h at 30 °C and 180 rpm. The cells were harvested by centrifugation (9000*g*, 10 min), and the pellet was washed with TBS/Ca buffer (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, pH 7.4, supplemented with 100 μM $CaCl_2$). The cells were resuspended in 40 ml of TBS/Ca supplemented with PMSF (1 mM) and lysozyme (0.4 mg mL^{-1}) and subsequently disrupted by five cycles on a SONOPULS ultrasonic homogeniser (BANDELIN electronic, Germany). Cell debris was removed by centrifugation (19,000*g*, 60 min), and the supernatant was loaded onto a gel filtration column (see below) or a column containing fucose-coupled sepharose CL-6B (for EclA and EclA-C). The affinity resin was prepared using L-fucose in analogy to the procedure reported by Fornstedt and Porath (53). The column was washed with TBS/Ca buffer. EclA and EclA-C were then eluted by addition of 100 mM L-fucose to the TBS/Ca buffer. The eluted fractions were extensively dialyzed against distilled water and then TBS/Ca buffer. For the His-tagged protein (EclA-N-tag), the supernatant was loaded on a 5 ml HiTrap affinity column (GE Healthcare). The column was washed with TBS/Ca supplemented with 20 mM imidazole, and then EclA-N-tag was eluted by addition of 250 mM imidazole to the TBS/Ca buffer. The eluted fractions were thereafter desalted on a HiTrap desalting column (GE Healthcare). The concentrations of EclA, EclA-N-tag, and EclA-C were determined by UV absorbance at 280 nm using calculated molar extinction coefficients of 50,420, 22,460, and 27,960 M^{-1} cm^{-1}, respectively. The yield of purified EclA, EclA-N-tag, and EclA-C was 15, 5, and 12 mg per liter of culture volume, respectively.

EclA-L (M14A) was generated by QuikChange site-directed mutagenesis (54) using plasmid pET22b-*eclA* as the template and phusion polymerase (New England Biolabs). The PCR forward and reverse primers GB32 and GB33 (Table S1, MWG-Biotech AG) were designed to introduce point mutation at the desired position. The reaction for either a forward or a reverse primer was separately performed in a 25 l volume using a PCR thermocycler (Professional thermocycler, Analytik Jena). Thereafter, the separately amplified DNA strands were combined and reannealed to give pET22b-*eclA-L*. The parent template was digested by DpnI restriction enzymes (New England Biolabs) (>1 h at 37 °C), and then 10 l of the reaction was transformed (electroporation, 1800 V, 25 μF, 200 Ω, 5.2 ms) into *E. coli* XL1-blue cells and plated onto LB agar supplemented with 100 μg mL^{-1} ampicillin. After overnight incubation at 37 °C, a clone was selected, grown in liquid culture and the plasmid isolated (GenEluteTM Plasmid miniprep Kit, SIGMA). Finally, the isolated plasmid was confirmed by sequencing at GATC.

Gel filtration

A HiLoad 16/600 Superdex 200 pg column (GE Healthcare) was equilibrated with TBS/Ca buffer (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, pH 7.4, supplemented with 2 mM $CaCl_2$) at a flow rate of 1 ml/min. A calibration curve was used from our previous study. Thereafter, EclA was loaded on the column and analyzed with the same flow rate.

DLS measurements

DLS measurements were performed on a Zetasizer Nano-ZS (Malvern Instruments). Stock solutions were filtered with a syringe filter before measurements. Either 100 μM of EclA-L or 200 μM of EclA-N-tag and EclA-C in TBS/Ca (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, pH 7.4, supplemented with 1 mM $CaCl_2$) was measured at 25 °C.

Fluorescence labeling of EclA constructs and glycan array analysis

Protein (1.4 ml of 60 μM EclA in Na_2CO_3 buffer, pH 9.3) was incubated at room temperature under shaking (500 rpm) with FITC (66 l of 3 mg mL^{-1}, in sodium carbonate buffer, pH 9.3) for 1 h. Purification of the labeled protein was performed as described above for unlabeled EclA; the protein concentration was determined as described previously for LecB-PA14 (39) using an extinction coefficient 50420 M^{-1} cm^{-1}.

FITC-labeled EclA was tested on the CFG mammalian glycan array (Core H), version 5.3, containing 585 printed glycans in replicates of 6. Standard procedures of Core H (details see https://research.bidmc.org/ncfg) were run at 5 and 50 μg mL^{-1} protein based on the protocol by Blixt *et al.* (55). Raw data of the EclA-binding experiments at 5 and 50 μg mL^{-1} are available as in Table S3.

For analysis on the Semiotik glycan array, EclA-L, EclA-C, and EclA-N-tag were individually labeled using Cy3-NHS ester (Toronto Research Chemicals). Briefly, 700 μg of protein in sodium carbonate buffer (190 μl, pH 8.3) was mixed with 10 μl Cy3-NHS ester solution (4.7 mg/ml in DMSO). The reaction was incubated for 4 h at r.t. under exclusion of light. The labeled proteins were desalted, and excess dye was removed in spin columns using TBS/Ca. The labeled proteins were used without further purification. The samples were analyzed on the

Semiotik glycan array as described (56). Array slides were incubated with protein samples at 50 μg/ml (in PBS supplemented with 0.05% Tween 20) for 1 h at 37 °C and shaking (32–34 rpm). Slides were analyzed after washing on a PerkinElmer ScanArray Gx (gain 60, laser power 80, resolution 20 μm). The raw data are depicted in the Supporting Information as Table S4.

Thermal shift assay

The thermal shift assay was conducted using a StepOnePlus Real Time PCR device (Applied Biosystems). Ten microliters of a solution of each tested compound at 20 mM in TBS/Ca was transferred in triplicate into white semiskirted RT-PCR 96-well plate (Axon-lab). Thereafter, 10 μl of a solution of 10 μM of EclA and 10X of SYPRO Orange dye in TBS/Ca was added to each well. The plate was centrifuged for 1 min at 4000 rpm and covered by RT-PCR 96-well plate foil (Axon-lab). The final concentrations of EclA, SYPRO orange, and compounds were 5 to 10 μM, 5-10X, and 10 mM, respectively. The plate was heated from 20 to 99 °C with a heating rate of 0.5 °C min^{-1}. Fluorescence was monitored at Ex/Em: 490/530 nm.

Direct titration of fluorescent ligand 11 with EclA or EclA-C

A serial dilution of EclA or EclA-C (10 μl, 0.4–400 μM) was transferred in triplicates to a black 384-well plate (Greiner Bio-One, catalog no. 781900), and then 10 μl of 20 nM of fluorescent ligand **11** in TBS/Ca buffer was added to each well. The plate was centrifuged for 1 min at 4000 rpm. After incubation at r.t. for 1 h, blank corrected fluorescence intensity was recorded using a PheraStar FS microplate reader (BMG Labtech GmbH) with excitation filters at 485 nm and emission filters at 535 nm, and fluorescence polarization was calculated. The data were analyzed using a four-parameter fit of the MARS Data Analysis Software (BMG Labtech GmbH, Germany). A minimum of three independent experiments on three plates was performed.

Competitive binding assay: Displacement of fluorescent ligand 11 from EclA, EclA-L, or EclA-C

Each tested monosaccharide and the Histo-blood group antigens were serially diluted from 0.07 to 20 mM and 0.03 to 3.5 mM, respectively. Thereafter, 10 μl of each concentration was transferred to a 384-well plate (Greiner Bio-One, catalog no. 781900) in triplicates. Then, 10 μl of 20 μM of EclA or EclA-C and 10 nM of fluorescent ligand **11** in TBS/Ca were added to each well. The plate was centrifuged for 1 min at 4000 rpm. After incubation at r.t. for 1 h, blank corrected fluorescence intensity was recorded using a PheraStar FS microplate reader (BMG Labtech GmbH) with excitation filters at 485 nm and emission filters at 535 nm, and fluorescence polarization was calculated. The data were analyzed using a four-parameter fit of the MARS Data Analysis Software (BMG Labtech GmbH). A minimum of two independent experiments on two plates was performed for each compound.

Crystallization

Crystals of EclA-N (12 mg/ml) were obtained in 0.06 M divalents (magnesium chloride hexahydrate and calcium chloride dihydrate), 0.1 M Tris and bicine, pH 8.5, 40% v/v glycerol, and 20% w/v PEG 4000. EclA-C (8 mg/ml) was pre-incubated with methyl α-L-selenofucoside (1–2 mM) prior to setting up crystallization trials. The optimized crystals were obtained in 100 to 280 mM disodium phosphate and 16 to 22% w/v PEG 3350 and cryoprotected with 30% v/v glycerol. Data were collected at SLS (EclA N terminus; beamline X06DA) and ESRF (EclA-C terminus; beamline MASSIF-3). EclA-N-tag structure was determined using PHASER Molecular Replacement (43) with PllA (PDB ID 5OFZ) as a search model, whereas the EclA-C structure was determined using the anomalous signals from the Se-containing sugar. Data were processed using XDS (57) or Xia2 (58), and the initial models obtained using PHENIX AutoSol (59) were rebuilt manually with COOT (60) and refined using PHENIX (61). Both structures were validated using MolProbity (62), and the images presented were created using PyMOL (Schrödinger, LLC) and Ligplot+ (63). Protein crystallography data collection and refinement statistics are depicted in Table S2.

Bioinformatics

The protein sequence of EclA from *E. cloacae* subsp. *cloacae* ATCC 13047 was used in a tblastn search in the *Enterobacter* taxid using the Basic Local Alignment Search Tool on the NCBI/NLM/NIH website (https://blast.ncbi.nlm.nih.gov/Blast.cgi). In a second tblastn search, the taxid *Enterobacter* was excluded from the search, and only the C-terminal domain of EclA was used as a query.

Sequence alignments were performed using the multiple sequence alignment tool Clustal Omega (64). Amino acids are labeled according to percentage identity (the higher the identity, the darker the color) using Jalview (65) software.

Data availability

All data are contained within the manuscript and the supporting information. Additionally, crystal structure data are deposited at the Protein Data Bank (www.rcsb.org) under accession numbers 6YF6 and 6YGQ.

Supporting information—This article contains supporting information.

Acknowledgments—We acknowledge Dr Eike Siebs (HIPS, Saarbrücken) for performing thermal shift assays on EclA-N-tag. The authors acknowledge excellent technical support from Dirk Hauck (HIPS, Saarbrücken). We further acknowledge the participation of the Protein-Glycan Interaction Resource of the CFG and the National Center for Functional Glycomics at Beth Israel Deaconess Medical Center, Harvard Medical School (supporting

grant R24 GM137763). Dr Jamie Heimburg-Molinaro and Dr Richard D. Cummings from the Consortium for Functional Glycomics and Dr Nadya Shilova from Semiotik LLC are gratefully acknowledged for performing and analyzing glycan arrays.

Author contribution—M.F., A.S., A.T., J.K., and S.W. writing–review & editing; M.F., G.B., J.K., and A.T. writing–original draft; G.B., A.S., J.K., and S.W. visualization; G.B., J.K., A.T., and S.W. validation; G.B., J.K., A.T., and S.W. supervision; G.B., A.S., J.G., J.K., and S.W. investigation; G.B., A.S., J.G., J.K., A.T., and S.W. formal analysis; G.B., J.G., J.K., and A.S. data curation; J.K., A.T., A.S., G.B., and S.W. methodology; A.T., J.K., and R.S. resources; A.T. project administration; A.T. funding acquisition; A.T. conceptualization.

Funding and additional information—A.T. acknowledges funding from the TANDEM graduate school of Saarland University. Support from EU COST Action EureStop is kindly acknowledged.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CBM, carbohydrate-binding module; CFG, Consortium for Functional Glycomics; DLS, dynamic light scattering; EclA, Enterobacter cloacae lectin A; EclA-L, EclA-M14 A; EclA-C, EclA C terminus; EclA-N-tag, EclA N terminus; FITC, fluorescein isothiocyanate; MeSeFuc, methyl α-l-selenofucoside.

References

1. Miethke, M., Pieroni, M., Weber, T., Brönstrup, M., Hammann, P., Halby, L., *et al.* (2021) Towards the sustainable discovery and development of new antibiotics. *Nat. Rev. Chem.* **5**, 726–749
2. Miller, W. R., and Arias, C. A. (2024) ESKAPE pathogens: antimicrobial resistance, epidemiology, clinical impact and therapeutics. *Nat. Rev. Microbiol.* **22**, 598–616
3. Wagner, S., Sommer, R., Hinsberger, S., Lu, C., Hartmann, R. W., Empting, M., *et al.* (2016) Novel Strategies for the treatment of Pseudomonas aeruginosa infections. *J. Med. Chem.* **59**, 5929–5969
4. Calvert, M. B., Jumde, V. R., and Titz, A. (2018) Pathoblockers or antivirulence drugs as a new option for the treatment of bacterial infections. *Beilstein J. Org. Chem.* **14**, 2607–2617
5. Dickey, S. W., Cheung, G. Y. C., and Otto, M. (2017) Different drugs for bad bugs: antivirulence strategies in the age of antibiotic resistance. *Nat. Rev. Drug Discov.* **16**, 457–471
6. Poole, K. (2011) Pseudomonas aeruginosa: resistance to the max. *Front. Microbiol.* **2**, 65
7. Sharon, N. (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta* **1760**, 527–537
8. Lis, H., and Sharon, N. (1998) Lectins: carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.* **98**, 637–674
9. Imberty, A., and Varrot, A. (2008) Microbial recognition of human cell surface glycoconjugates. *Curr. Opin. Struct. Biol.* **18**, 567–576
10. Zahorska, E., Rosato, F., Stober, K., Kuhaudomlarp, S., Meiers, J., Hauck, D., *et al.* (2023) Neutralizing the impact of the virulence factor LecA from Pseudomonas aeruginosa on human cells with new glycomimetic inhibitors. *Angew. Chem. Int. Ed. Engl.* **62**, e202215535
11. Sommer, R., Wagner, S., Rox, K., Varrot, A., Hauck, D., Wamhoff, E.-C., *et al.* (2018) Glycomimetic, orally bioavailable LecB inhibitors block biofilm formation of Pseudomonas aeruginosa. *J. Am. Chem. Soc.* **140**, 2537–2545
12. Sommer, R., Rox, K., Wagner, S., Hauck, D., Henrikus, S. S., Newsad, S., *et al.* (2019) Anti-biofilm agents against Pseudomonas aeruginosa: a structure-activity relationship study of C-glycosidic LecB inhibitors. *J. Med. Chem.* **62**, 9201–9216
13. Kadam, R. U., Bergmann, M., Hurley, M., Garg, D., Cacciarini, M., Swiderska, M. A., *et al.* (2011) A glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of Pseudomonas aeruginosa biofilms. *Angew. Chem. Int. Ed. Engl.* **50**, 10631–10635
14. Reymond, J.-L., Bergmann, M., and Darbre, T. (2013) Glycopeptide dendrimers as Pseudomonas aeruginosa biofilm inhibitors. *Chem. Soc. Rev.* **42**, 4814–4822
15. Johansson, E. M. V., Crusz, S. A., Kolomiets, E., Buts, L., Kadam, R. U., Cacciarini, M., *et al.* (2008) Inhibition and dispersion of Pseudomonas aeruginosa biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB. *Chem. Biol.* **15**, 1249–1257
16. Sponsel, J., Guo, Y., Hamzam, L., Lavanant, A. C., Pérez-Riverón, A., Partiot, E., *et al.* (2023) Pseudomonas aeruginosa LecB suppresses immune responses by inhibiting transendothelial migration. *EMBO Rep.* **24**, e55971
17. Boukerb, A. M., Rousset, A., Galanos, N., Méar, J.-B., Thepaut, M., Grandjean, T., *et al.* (2014) Anti-adhesive properties of glycoclusters against Pseudomonas aeruginosa lung infection. *J. Med. Chem.* **57**, 10275–10289
18. Chemani, C., Imberty, A., de Bentzmann, S., Pierre, M., Wimmerová, M., Guery, B. P., *et al.* (2009) Role of LecA and LecB lectins in Pseudomonas aeruginosa-induced lung injury and effect of carbohydrate ligands. *Infect. Immun.* **77**, 2065–2075
19. Leusmann, S., Ménová, P., Shanin, E., Titz, A., and Rademacher, C. (2023) Glycomimetics for the inhibition and modulation of lectins. *Chem. Soc. Rev.* **52**, 3663–3740
20. Bucior, I., Abbott, J., Song, Y., Matthay, M. A., and 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–L363
21. Hauber, H.-P., Schulz, M., Pforte, A., Mack, D., Zabel, P., and Schumacher, U. (2008) Inhalation with fucose and galactose for treatment of Pseudomonas aeruginosa in cystic fibrosis patients. *Int. J. Med. Sci.* **5**, 371–376
22. Beshr, G., Sikandar, A., Jemiller, E.-M., Klymiuk, N., Hauck, D., Wagner, S., *et al.* (2017) Photorhabdus luminescens lectin A (PllA) - a new probe for detecting α-galactoside-terminating glycoconjugates. *J. Biol. Chem.* **292**, 19935–19951
23. Liu, S., Chen, L., Wang, L., Zhou, B., Ye, D., Zheng, X., *et al.* (2022) Cluster differences in antibiotic resistance, biofilm formation, mobility, and virulence of clinical Enterobacter cloacae complex. *Front. Microbiol.* **13**, 814831
24. Mezzatesta, M. L., Gona, F., and Stefani, S. (2012) Enterobacter cloacae complex: clinical impact and emerging antibiotic resistance. *Future Microbiol.* **7**, 887–902
25. Garinet, S., Fihman, V., Jacquier, H., Corvec, S., Le Monnier, A., Guillard, T., *et al.* (2018) Elective distribution of resistance to beta-lactams among Enterobacter cloacae genetic clusters. *J. Infect.* **77**, 178–182
26. Bousquet, A., van der Mee-Marquet, N., Dubost, C., Bigaillon, C., Larréché, S., Bugier, S., *et al.* (2017) Outbreak of CTX-M-15–producing Enterobacter cloacae associated with therapeutic beds and syphons in an intensive care unit. *Am. J. Infect. Control* **45**, 1160–1164
27. Girlich, D., Ouzani, S., Emeraud, C., Gauthier, L., Bonnin, R. A., Le Sache, N., *et al.* (2021) Uncovering the novel Enterobacter cloacae complex species responsible for septic shock deaths in newborns: a cohort study. *Lancet Microbe* **2**, e536–e544
28. Barnes, A. I., Ortiz, C., Paraje, M. G., Balanzino, L. E., and Albesa, I. (1997) Purification and characterization of a cytotoxin from *Enterobacter cloacae*. *Can. J. Microbiol.* **43**, 729–733
29. Krzymińska, S., Mokracka, J., Koczura, R., and Kaznowski, A. (2009) Cytotoxic activity of *Enterobacter cloacae* human isolates. *FEMS Immunol. Med. Microbiol.* **56**, 248–252
30. Soria-Bustos, J., Ares, M. A., Gómez-Aldapa, C. A., González-y-Merchand, J. A., Girón, J. A., and De la Cruz, M. A. (2020) Two type VI secretion systems of Enterobacter cloacae are required for bacterial competition, cell adherence, and intestinal colonization. *Front. Microbiol.* **11**, 560488

31. Ramirez, D., and Giron, M. (2024) *Enterobacter Infections*, StatPearls Publishing
32. Davin-Regli, A., and PagÃ¨s, J.-M. (2015) Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment. *Front. Microbiol.* **6**, 392
33. Ren, Y., Ren, Y., Zhou, Z., Guo, X., Li, Y., Feng, L., *et al.* (2010) Complete genome sequence of *Enterobacter cloacae* subsp. *cloacae* type strain ATCC 13047. *J. Bacteriol.* **192**, 2463–2464
34. Pantel, L., Guérin, F., Serri, M., Gravey, F., Houard, J., Maurent, K., *et al.* (2022) Exploring cluster-dependent antibacterial activities and resistance pathways of NOSO-502 and colistin against Enterobacter cloacae complex species. *Antimicrob. Agents Chemother.* **66**, e0077622
35. Bhar, S., Edelmann, M. J., and Jones, M. K. (2021) Characterization and proteomic analysis of outer membrane vesicles from a commensal microbe, Enterobacter cloacae. *J. Proteomics* **231**, 103994
36. Frutos-Grilo, E., Kreling, V., Hensel, A., and Campoy, S. (2023) Host-pathogen interaction: Enterobacter cloacae exerts different adhesion and invasion capacities against different host cell types. *PLoS One* **18**, e0289334
37. Joachim, I., Rikker, S., Hauck, D., Ponader, D., Boden, S., Sommer, R., *et al.* (2016) Development and optimization of a competitive binding assay for the galactophilic low affinity lectin LecA from Pseudomonas aeruginosa. *Org. Biomol. Chem.* **14**, 7933–7948
38. Hauck, D., Joachim, I., Frommeyer, B., Varrot, A., Philipp, B., Möller, H. M., *et al.* (2013) Discovery of two classes of potent glycomimetic inhibitors of Pseudomonas aeruginosa LecB with distinct binding modes. *ACS Chem. Biol.* **8**, 1775–1784
39. Sommer, R., Wagner, S., Varrot, A., Nycholat, C. M., Khaledi, A., Haussler, S., *et al.* (2016) The virulence factor LecB varies in clinical isolates: consequences for ligand binding and drug discovery. *Chem. Sci.* **7**, 4990–5001
40. Beshr, G., Sommer, R., Hauck, D., Siebert, D. C. B., Hofmann, A., Imberty, A., *et al.* (2016) Development of a competitive binding assay for the Burkholderia cenocepacia lectin BC2L-A and structure activity relationship of natural and synthetic inhibitors. *Med. Chem. Commun.* **7**, 519–530
41. Dingjan, T., Gillon, É., Imberty, A., Pérez, S., Titz, A., Ramsland, P. A., *et al.* (2018) Virtual screening against carbohydrate-binding proteins: Evaluation and Application to bacterial Burkholderia ambifaria lectin. *J. Chem. Inf. Model.* **58**, 1976–1989
42. Kadam, R. U., Bergmann, M., Garg, D., Gabrieli, G., Stocker, A., Darbre, T., *et al.* (2013) Structure-based optimization of the terminal tripeptide in glycopeptide dendrimer inhibitors of Pseudomonas aeruginosa biofilms targeting LecA. *Chem. Eur. J.* **19**, 17054–17063
43. 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
44. Holm, L. (2020) DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140
45. Cecioni, S., Imberty, A., and Vidal, S. (2015) Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. *Chem. Rev.* **115**, 525–561
46. Sicard, D., Cecioni, S., Iazykov, M., Chevolot, Y., Matthews, S. E., Praly, J.-P., *et al.* (2011) AFM investigation of Pseudomonas aeruginosa lectin LecA (PA-IL) filaments induced by multivalent glycoclusters. *Chem. Commun. Camb.* **47**, 9483–9485
47. Cooling, L. (2015) Blood groups in infection and host susceptibility. *Clin. Microbiol. Rev.* **28**, 801–870
48. Heggelund, J. E., Varrot, A., Imberty, A., and Krengel, U. (2017) Histo-blood group antigens as mediators of infections. *Curr. Opin. Struct. Biol.* **44**, 190–200
49. Boral, L. I., Staubach, Z. G., de Leeuw, R., Macivor, D. C., Kryscio, R., and Bada, H. S. (2013) Comparison of outcomes of group O vs non-group O premature neonates receiving group O RBC transfusions. *Am. J. Clin. Pathol.* **140**, 780–786
50. Jajosky, R. P., Wu, S.-C., Zheng, L., Jajosky, A. N., Jajosky, P. G., Josephson, C. D., *et al.* (2023) ABO blood group antigens and differential glycan expression: perspective on the evolution of common human enzyme deficiencies. *iScience* **26**, 105798
51. Houser, J., Komarek, J., Kostlanova, N., Cioci, G., Varrot, A., Kerr, S. C., *et al.* (2013) A soluble fucose-specific lectin from Aspergillus fumigatus conidia–structure, specificity and possible role in fungal pathogenicity. *PLoS One* **8**, e83077
52. Sommer, R., Makshakova, O. N., Wohlschlager, T., Hutin, S., Marsh, M., Titz, A., *et al.* (2018) Crystal structures of fungal tectonin in complex with O-methylated glycans suggest key role in innate immune defense. *Structure* **26**, 391–402.e4
53. Fornstedt, N., and Porath, J. (1975) Characterization studies on a new lectin found in seeds of vicia-ervilia. *Febs Lett.* **57**, 187–191
54. Edelheit, O., Hanukoglu, A., and Hanukoglu, I. (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9**, 61
55. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., *et al.* (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17033–17038
56. Olivera-Ardid, S., Khasbiullina, N., Nokel, A., Formanovsky, A., Popova, I., Tyrtysh, T., *et al.* (2019) Printed glycan array: a sensitive technique for the analysis of the repertoire of circulating anti-carbohydrate antibodies in small animals. *J. Vis. Exp.* https://doi.org/10.3791/57662
57. Kabsch, W. (2010) Xds. *Acta Crystallogr. D.* **66**, 125–132
58. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190
59. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., *et al.* (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. Biol. Crystallogr.* **65**, 582–601
60. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of coot. *Acta Crystallogr. D* **66**, 486–501
61. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D.* **66**, 213–221
62. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D.* **66**, 12–21
63. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **51**, 2778–2786
64. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
65. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191