Research Article
doi.org/10.1002/cbic.202100563
Chemistry Europe
European Chemical Societies Publishing

Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA

Eike Siebs,[a, b, c] Elena Shanina,[d, e] Sakonwan Kuhaudomlarp,[f, g] Priscila da Silva Figueiredo Celestino Gomes,[h] Cloé Fortin,[a] Peter H. Seeberger,[d, e] Didier Rognan,[h] Christoph Rademacher,[d, e, i, j] Anne Imberty,[f] and Alexander Titz*[a, b, c]

Pseudomonas aeruginosa is an opportunistic ESKAPE pathogen that produces two lectins, LecA and LecB, as part of its large arsenal of virulence factors. Both carbohydrate-binding proteins are central to the initial and later persistent infection processes, i.e. bacterial adhesion and biofilm formation. The biofilm matrix is a major resistance determinant and protects the bacteria against external threats such as the host immune system or antibiotic treatment. Therefore, the development of drugs against the *P. aeruginosa* biofilm is of particular interest to restore efficacy of antimicrobials. Carbohydrate-based inhibitors for LecA and LecB were previously shown to efficiently reduce biofilm formations. Here, we report a new approach for inhibiting LecA with synthetic molecules bridging the established carbohydrate-binding site and a central cavity located between two LecA protomers of the lectin tetramer. Inspired by *in silico* design, we synthesized various galactosidic LecA inhibitors with aromatic moieties targeting this central pocket. These compounds reached low micromolar affinities, validated in different biophysical assays. Finally, X-ray diffraction analysis revealed the interactions of this compound class with LecA. This new mode of action paves the way to a novel route towards inhibition of *P. aeruginosa* biofilms.

[a] *E. Siebs, C. Fortin, A. Titz*
Chemical Biology of Carbohydrates (CBCH)
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS)
Helmholtz Centre for Infection Research, 66123 Saarbrücken (Germany)
E-mail: alexander.titz@helmholtz-hzi.de
[b] *E. Siebs, A. Titz*
Department of Chemistry, Saarland University
66123 Saarbrücken (Germany)
[c] *E. Siebs, A. Titz*
Deutsches Zentrum für Infektionsforschung (DZIF)
Standort Hannover-Braunschweig (Germany)
[d] *E. Shanina, P. H. Seeberger, C. Rademacher*
Department of Biomolecular Systems
Max Planck Institute of Colloids and Interfaces, 14424 Potsdam (Germany)
[e] *E. Shanina, P. H. Seeberger, C. Rademacher*
Institute of Chemistry and Biochemistry
Department of Biology, Chemistry and Pharmacy
Freie Universität Berlin, 14195 Berlin (Germany)
[f] *S. Kuhaudomlarp, A. Imberty*
Université Grenoble Alpes, CNRS, CERMAV
38000 Grenoble (France)
[g] *S. Kuhaudomlarp*
Department of Biochemistry and Centre for Excellence in Protein and Enzyme Technology, Faculty of Science
Mahidol University, Bangkok (Thailand)
[h] *P. da Silva Figueiredo Celestino Gomes, D. Rognan*
Laboratoire d'Innovation Thérapeutique
UMR 7200 CNRS-Université de Strasbourg, Strasbourg, 67400 Illkirch (France)
[i] *C. Rademacher*
Department of Pharmaceutical Sciences
University of Vienna, Althanstrasse 14, 1090 Vienna (Austria)
[j] *C. Rademacher*
Department of Microbiology, Immunology and Genetics
University of Vienna, Max F. Perutz Labs, Biocenter 5, 1030 Vienna (Austria)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100563

Introduction

Pseudomonas aeruginosa belongs to the ESKAPE pathogens and is listed by the World Health Organization as the most critical bacterial pathogen. This rod-shaped, Gram-negative ubiquitous bacterium lives under moist conditions and thrives in hospitals, colonizing patients under ventilation or bearing catheters and cystic fibrosis patients.[1] For colonizing the host, many bacteria exploit cell surface carbohydrates, the glycocalyx,[2] through their carbohydrate-binding proteins (lectins), adhesins or capsid proteins.[3] In case of *P. aeruginosa* two soluble virulence factors are involved: the D-galactose-binding LecA and D-mannose-/L-fucose-binding LecB.[4] These two extracellular, calcium(II)-dependent and homotetrameric lectins are essential for establishing the biofilm matrix, a protective environment against environmental stress and antibiotic treatment and, thus, a major hurdle for therapy.[5–7]

To overcome this resistance mechanism, alternative strategies are currently being developed to dismantle the pathogen and restore antibiotic efficacy *via* inhibition of the lectins LecA and LecB.[8–10] Complementary to the numerous approaches for multivalent inhibition using native carbohydrates,[11,12] we have developed glycomimetics targeting LecA and LecB. In case of LecB, we have developed small molecule glycomimetics starting from a mannoside[13–15] *via* simple C-glycosides[16] into orally available anti-biofilm lead compounds[17,18] which are currently under further investigation. Potent monovalent glycomimetics have been proven difficult to obtain for galactophilic LecA,[19–21] thus requiring divalent ligands that display two galactose residues to simultaneously bind to two adjacent binding sites in the LecA tetramer and yield low nanomolar LecA inhibitors.[22,23] In addition, we have recently reported conceptionally new approaches for targeting LecA, i.e. covalent lectin inhibitors[24]

and non-carbohydrate glycomimetics mimicking the binding pattern of carbohydrates.[25]

For LecA, inhibitor design generally started from the monosaccharide galactose (Figure 1A)[10] which is the binding epitope from its natural ligand in human, glycosphingolipid Gb3, a host cell surface receptor mediating the engulfment of *P. aeruginosa* into host cells.[26] Galactose forms an extensive hydrogen bonding network with LecA using all hydroxy groups. OH3 and OH4 are further coordinating to the calcium ion of LecA.[27] OH6 is pointing into a small cavity formed by a loop from His50 to Gln53 and Val101 where it is coordinated through an extended hydrogen-bonding network. Introduction of β-linked aromatic aglycons led to a fivefold affinity increase compared to aliphatic analogues, which is explained by CH-π-interaction of the aryl aglycon with the side chain of His50.[28]

In the present work, we localized a cavity between two monomers of LecA (Figure 1B), designated as central pocket, and developed synthetic galactosides carrying aromatic moieties directed towards this cavity.

Results and Discussion

Identification of the central pocket in LecA

23 crystallographic structures of LecA from *P. aeruginosa* (Uniprot reference Q05097) were retrieved from the Protein Data Bank[30] (PDB) and analyzed for the presence of the central pocket (Figure 1, Table S1). A central pocket of LecA can be identified between two adjacent monomers in the homotetrameric LecA structure. The entrance of the pocket is clearly visible (Figure 1B) and the upper part of the cavity including its entrance, formed by residues Gln40, Lys41, Asp47, Arg48 and Glu49, is highly polar. In contrast, the interior is more hydrophobic due to the presence of two tryptophan residues, Trp33 and Trp42. The central pocket's cavity was consistently detected in 25 out of the 31 dimers in the available crystallographic structures (Table S2).

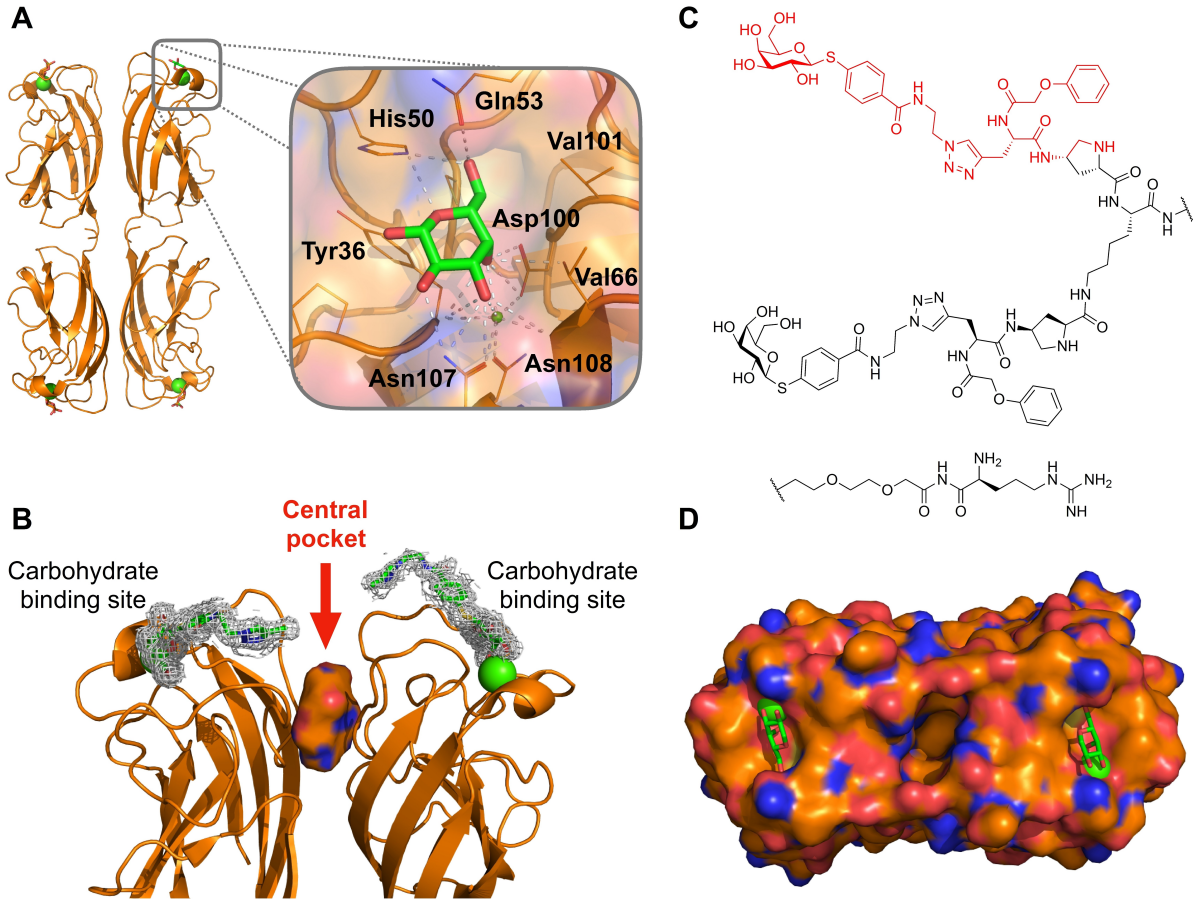


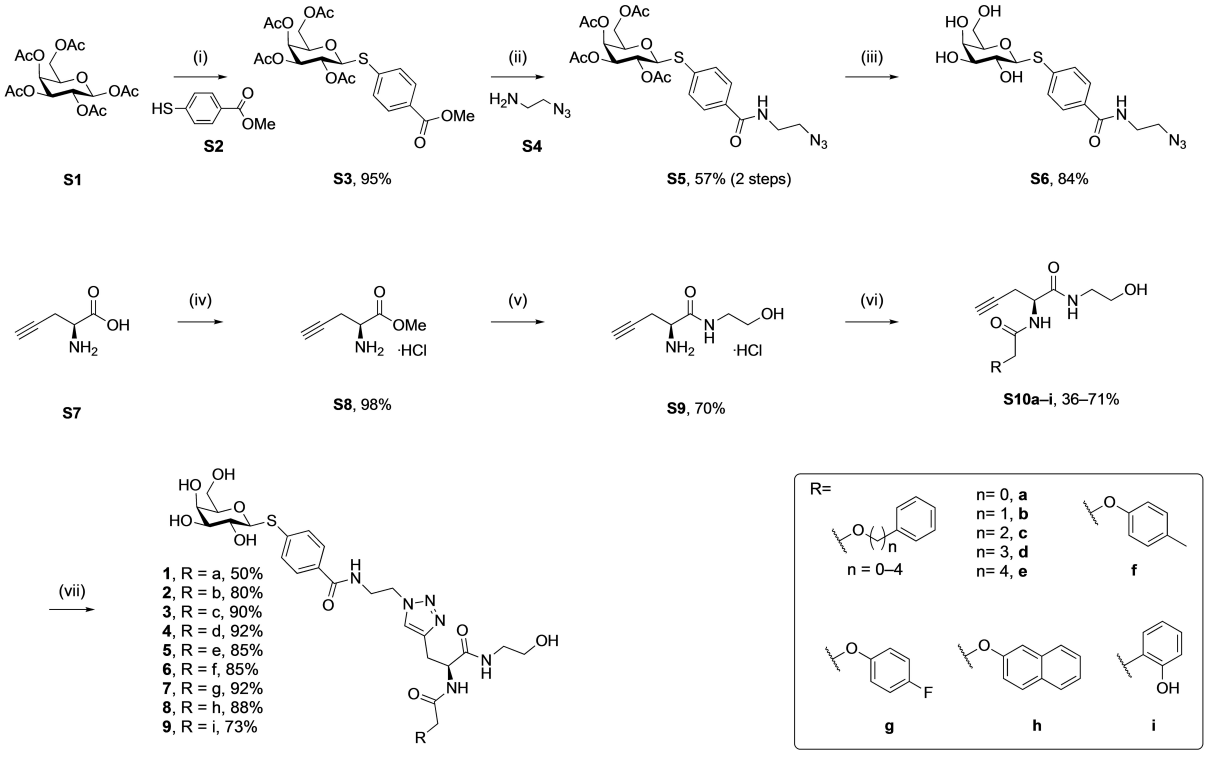
Figure 1. Rationale for targeting the central pocket in LecA: (A) Co-crystal structure showing the tetramer of LecA in complex with galactose (PDB: 1OKO). (B) Side view of one of the two adjacent dimers of LecA with the identified cavity (solid surface) between the two monomers defined as central pocket. The divalent ligand (PDB: 4CP9) reported by Winssinger *et al.*[29] is represented as sticks in electron density and is pointing towards the central pocket. (C) Structure of the divalent LecA ligand.[29] and the structural motif studied in this work highlighted in red. (D) Top view of the surface of one LecA dimer showing the cavity between the two monomers (PDB: 4LKE). The entrance of this cavity is polar due to the presence of Gln40, Lys41, Asp47, Arg48 and Glu49 and the interior is hydrophobic due to residues Trp33 and Trp42. Calcium ions in the carbohydrate binding sites are shown as green spheres.

Design and *in silico* analysis by molecular dynamics simulation

As a starting point for ligand design, we were inspired by a bivalent peptide-based LecA inhibitor ($K_D = 82$ nM, Figure 1C) previously reported by Winssinger *et al.*[29] As often observed for multivalent ligands, the crystal structure of the complex showed electron density only for part of the molecule, i.e. the galactose residue, the aromatic aglycon and the attached triazole ring which is in proximity to the entrance of the central pocket (Figure 1B). The authors reported the crystallographically invisible phenoxy acetyl substituent adjacent to the triazole to be important for LecA binding. We hypothesized that this phenoxyacetyl moiety may bind close to or into the central pocket. To test this hypothesis, we truncated the original divalent molecule and retained the galactose residue, its aglycon and the attached phenoxyacetyl residue (shown in red in Figure 1C) and designed derivatives carrying various substituents to target the central pocket *in silico*.

Compound **1** (Scheme 1) was then docked to LecA, and in the resulting poses, the phenoxy acetate part reaches towards the central pocket. Subsequently, molecular dynamics simulation was performed on one selected pose obtained from docking. Visual inspection of the 240 ns MD trajectory revealed significant fluctuation between the phenoxyacetate side chain and LecA with repeated contacts of the pharmacophore and the central pocket (Figure 2). In general, the galactose moiety remained firmly bound inside the carbohydrate binding site with strong hydrogen bonds to residues His50, Gln53, Asp100 and Asn103, and hydrophobic contacts to Tyr36, Cys62 and Val101. The terminal phenyl ring was entering and leaving the central pocket repeatedly, possibly due to a suboptimal spacer length to position the phenyl group firmly within the pocket and a lack of interactions with the polar entrance of the cavity. Less frequent hydrogen bonds were also observed between the linker and Gln40 and Asp47. Last, hydrophobic contacts could be detected between the linker and Pro38, the phenoxy moiety and Trp42 (chains A and D), and the triazole ring and Pro51.

Based on these *in silico* considerations, we designed and synthesized derivatives **1**–**5** with stepwise increased linker length between the phenyl group and its ether oxygen to increase the possibility for hydrophobic contacts inside the central pocket. Furthermore, we modified the phenyl group in compounds **6**–**8** to assess the effect of increased lipophilicity and synthesized compound **9** carrying a hydroxy group to probe for hydrogen bonding with the cavity's polar entrance (see Scheme 1 for structures).



Scheme 1. Synthesis of LecA inhibitors targeting the central pocket. The different side chains were introduced in the penultimate step *via* amide coupling with alkyne **S9**. Final assembly was achieved by coupling azide **S6** with alkynes **S10a–i** in a copper(I)-catalyzed cycloaddition. Reagents and conditions: (i) $BF_3{\cdot}OEt_2$, **S2**, CH_2Cl_2, 0–25 °C, 18 h; (ii) 1. LiI, pyridine, 25 °C, 3 d; 2. **S4**, HOBt, EDC, DMF, 25 °C, 24 h; (iii) NaOMe, MeOH, 25 °C, 1.5 h; (iv) $SOCl_2$, MeOH, 0–25 °C, 18 h; (v) ethanolamine, 25 °C, 18 h; (vi) various acetic acids, EDC, HOBt, DIPEA, DMF, 25 °C, 18 h; (vii) **S6**, $CuSO_4$ in H_2O, sodium ascorbate in H_2O, DMF, 25 °C, 2 h.

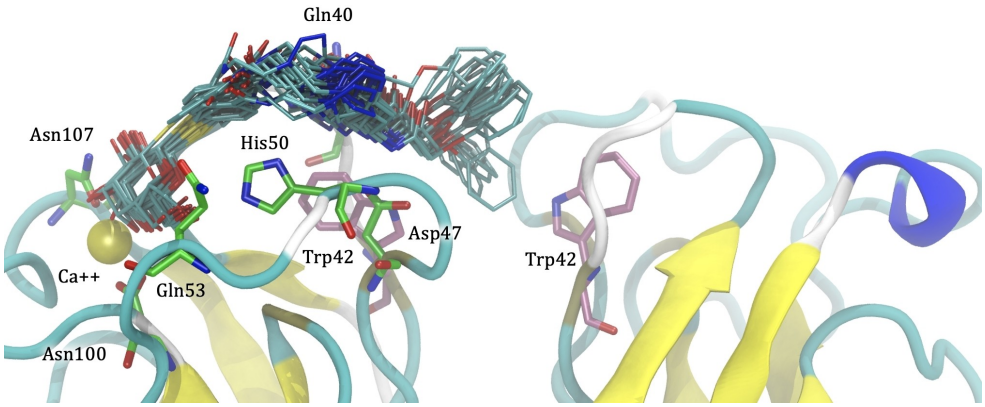


Figure 2. Molecular dynamics simulation of **1** in complex with LecA. Several snapshots of ligand **1** from the 240 ns molecular dynamic trajectory are depicted as cyan sticks. The phenoxy group of **1** is partially entering the central pocket of LecA and interacts with the Trp42 of both monomers (depicted as pink sticks). The galactosyl residue firmly coordinates the calcium ion (yellow sphere) in the carbohydrate binding site, and its β-phenyl aglycon binds to His50. Residues involved in hydrogen bonds to the galactosyl moiety are depicted by green sticks.

Synthesis

Galactose pentaacetate **S1** was reacted with thiol **S2** in presence of a Lewis acid to give β-thioglycoside **S3** in 95% yield (Scheme 1). Then, the methyl ester in **S3** was cleaved under S_N2 conditions using LiI to leave the acetates unchanged. The corresponding acid intermediate was then activated with EDC/HOBt and coupled with amine **S4** to give amide **S5** (57%, 2 steps). The latter was subsequently deprotected under Zemplén conditions to yield azide **S6**.

L-Propargylglycin **S7** was first transformed into ester **S8** using thionylchloride in MeOH and then stirred in neat ethanolamine to give amide **S9**. In the next step, the differently substituted acetic acids bearing the central pocket-targeting pharmacophores were introduced by amide coupling with alkynyl amine **S9** into amides **S10a–i** (36–71%). These alkynyl amides were then linked to azide **S6** in a copper(I)-catalyzed azide-alkyne cycloaddition reaction resulting in the final 9 triazoles **1**–**9** with varying side chains in high yields (50–92%).

Biophysical evaluation

All nine synthesized LecA ligands were then tested for LecA inhibition in a competitive binding assay based on fluorescence polarization[25] and azide **S6** was included as a control devoid of the side chain targeting the central pocket. In addition, the two positive controls methyl α-D-galactoside (Me-α-D-Gal) and *para*-nitrophenyl β-D-galactoside (pNPG) were included (Figure 3A). Unsubstituted azide **S6** ($IC_{50} = 50.3 \pm 5.4$ μM) showed a fourfold higher activity than Me-α-D-Gal ($IC_{50} = 196 \pm 7.8$ μM). Phenoxy acetate **1** ($IC_{50} = 49 \pm 4.5$ μM), the closest derivative of the original divalent ligand, was as active as unsubstituted **S6**. Compounds **6** and **7** carrying electron donating or withdrawing groups in *para* position of the phenyl ring, as well as naphthyl **8** and phenol **9** did not show increased potency compared to **1** in this assay. Derivatives **3**, **4** and **5** with increased spacer length by two to four methylene groups showed highest inhibition of LecA with IC_{50}s between 39 and 43 μM.

To further evaluate the affinity of the ligands towards LecA in a direct binding experiment, we performed surface plasmon resonance (SPR) analysis in which LecA was immobilized and compounds **1**–**9** were injected. Dissociation constants at steady state were calculated from multiple cycle analysis after injection of the tested compounds in 0–200 μM range (Figure 3C). The kinetic rate constants k_{on} and k_{off} could not be reliably determined due to fast association and dissociation of the compounds and LecA. All compounds showed affinities towards LecA with little variation between derivatives in the μM range (6.0 ± 0.2 to 10.3 ± 1.3 μM, Table 1) which is comparable to previously reported aromatic galactosides. The similar affinities observed for compounds carrying the additional aryloxy pharmacophores and that of **S6** suggest that the central pocket targeting pharmacophores in **1**–**9** either do not reach the intended site or the entropic penalties upon binding are compromising attractive interactions.

Isothermal titration calorimetry was used to obtain thermodynamic data on the LecA – inhibitor interaction with both binding partners in solution. Supporting the competitive binding assay and SPR, all tested compounds showed affinities in the low micromolar range in ITC (Table 1). Surprisingly, the unsubstituted ligand **S6** was the most active compound with a K_d of 4.8 μM. The triazole derivatives had K_d values between 6.3 and 9.4 μM, and binding enthalpies between −38.7 and −50.3 kJ/mol. Compound **2**, with an increased linker length of 1 methylene group compared to **1**, showed the highest affinity (K_d (**6**) = 6.3 μM) and a maximal enthalpy gain (ΔH (**2**) = -50.3 kJ mol^{-1}), followed by compound **8** with the longest linker of 4 methylene groups (K_d (**5**) = 6.7 μM, ΔH (**5**) = -46.3 kJ mol^{-1}). When comparing compounds **3**, **4** and **5**, that carry 2, 3 or 4 methylene groups respectively, a steady increase in binding enthalpy is observed which was counterbalanced by an unfavorable trend in entropy. Substitution in *para*/*ortho* position of the phenyl ring (**6**, **7**, **9**) or its replacement (**8**) did not lead to significant changes in binding affinity, except for

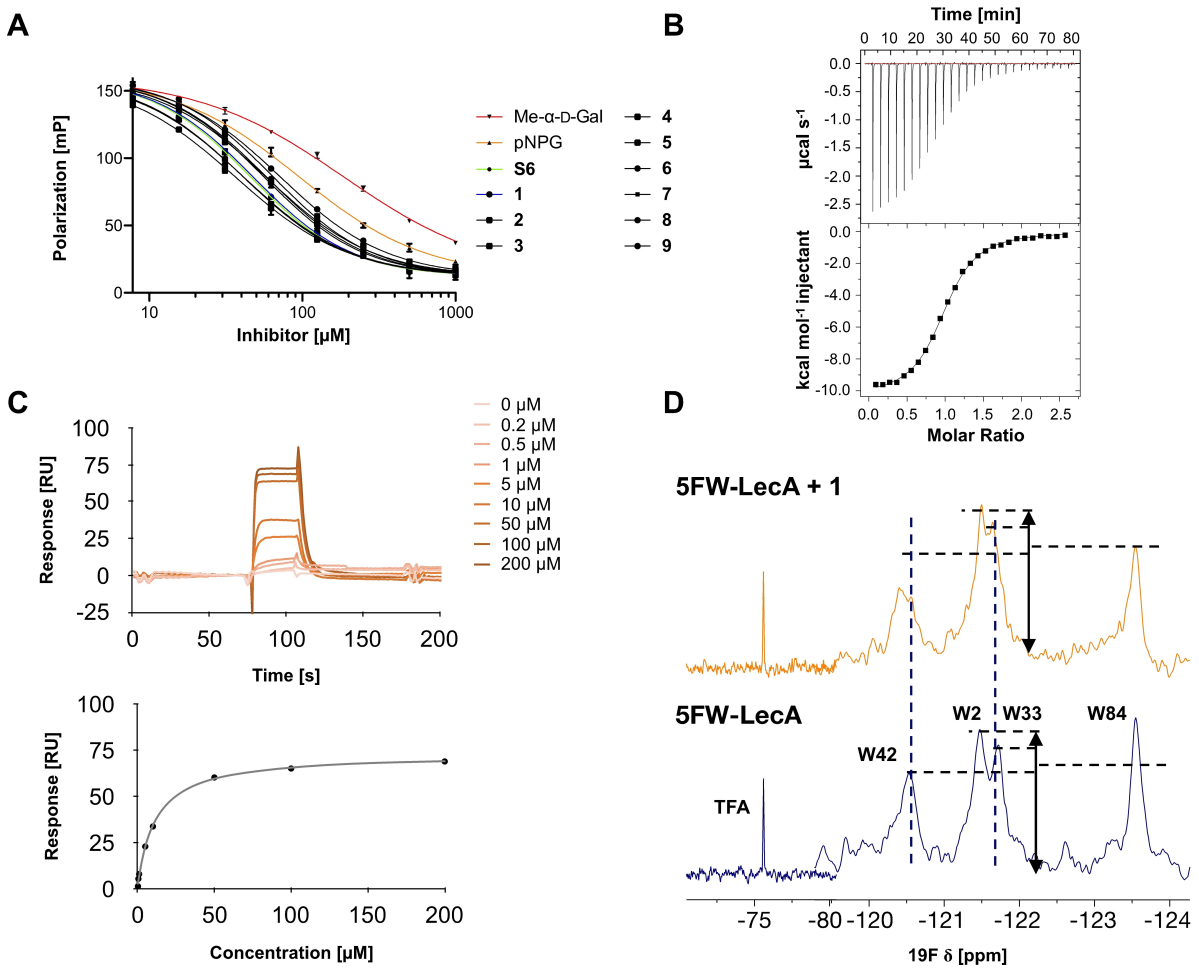


Figure 3. Analysis of synthetic inhibitor **1** interacting with LecA using (**A**) a competitive binding assay based on fluorescence polarization for all nine compounds including azide **S6** and controls methyl α-D-galactoside and *para*-nitrophenyl β-D-galactopyranoside, (**B**) isothermal titration calorimetry sensorgram (top panel) obtained by titration of **1** to LecA with integration of peaks and fit (bottom panel), (**C**) surface plasmon resonance using multi-cycle kinetic studies (data shown for **1**), top: sensorgram, bottom: affinity analysis, and (**D**) ^{19}F-protein-observed fluorine (PrOF) NMR demonstrates the impact of binding of **1** on the NMR resonances of Trp.

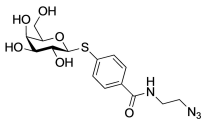
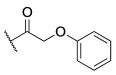
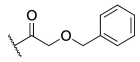
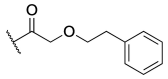
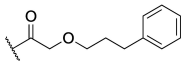
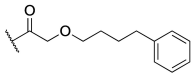
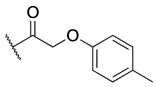
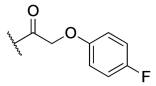
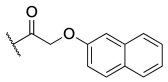
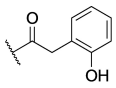
tolyl **6** that was the most active (K_d (**6**) = 6.8 μM), with highest binding enthalpy (ΔH (**6**) = −45.1 kJ mol^{-1}) at the cost of the highest entropic penalty ($-T\Delta S$(**6**) = 14.6 kJ mol^{-1}) among this subset of compounds **6**–**9**.

Protein observed ^{19}F (PrOF) NMR allows the detection of weak binding interactions with ligands. However, it requires the introduction of fluorine nuclei as sensitive NMR probes into the protein using ^{19}F-labelled amino acids or its precursors. Given the presence of tryptophans inside of the central pocket, we aimed to use PrOF NMR for detection of binders to this pocket.

To this end, all four tryptophans of LecA were simultaneously metabolically labeled at position 5 on the indole rings with fluorine and assignment was previously done by site-directed mutagenesis.[31] Compounds **1**, **2** or **8** were added to the labeled protein and PrOF NMR spectra were recorded and compared to those of the protein in absence of ligands (see Figure 3D). Trp42 is located in the carbohydrate binding site and its resonance was therefore affected by all tested ligands as observed by line broadening and chemical shift perturbation ($CSP_{(1)}$ = 0.13 ppm, $CSP_{(2)}$ = 0.14 ppm and $CSP_{(8)}$ = 0.14 ppm respectively). The signals for Trp2 and Trp84 were not affected in any case while the signal intensity of the tryptophan located in the central pocket, Trp33, was shifted in presence of **1** ($CSP_{(1)}$ = 0.08 ppm), **2** ($CSP_{(2)}$ = 0.06 ppm) and **8** ($CSP_{(8)}$ = 0.07 ppm) (Figure S2), supporting a contact of the arylether moieties with the central pocket of LecA.

To analyze the interaction of ligands with the central pocket at atomic resolution, we crystallized and analyzed LecA in complex with **1**. The resulting structure was determined at 1.53 Å resolution in space group $P2_12_12_1$ (Figure 1A, Table S3). The asymmetric unit consisted of a homotetramer of LecA, in agreement with previously reported LecA structures. One Ca^{2+}-ion and one molecule of **1** could be located in each monomer, although the completeness of electron density of the ligand beyond the triazole ring varies among the four sites (Figure 4A). The galactoside moiety of **1** interacts with LecA in the same manner as previously reported, with hydrogen bonds to residues His50, Gln53, Asp100, Thr104, Asn107 and Asn108, and

Table 1. Analysis of LecA inhibitors **S6** and **1–9** in direct binding (ITC, SPR) and competitive binding (FP) biophysical assays.

Structure	Name	ITC K_D [μM]	ΔH [kJ/mol]	−T ΔS [kJ/mol]	n	FP IC_{50} [μM]	SPR K_D [μM]
	**S6**	4.8 ± 0.3	−48.9 ± 1.6	18.6 ± 0.5	0.9 ± 0.0	50.3 ± 5.4	6.7 ± 0.3
	**1**	9.4 ± 4.3	−46.2 ± 3.5	17.3 ± 1.5	1.0 ± 0.1	49.0 ± 4.5	10.3 ± 1.3
	**2**	6.3 ± 0.3	−50.3 ± 1.1	20.7 ± 0.3	1.0 ± 0.1	60.7 ± 2.5	6.0 ± 0.2
	**3**	8.3 ± 0.3	−41.0 ± 1.1	12.0 ± 0.4	1.0 ± 0.1	43.3 ± 3.2	6.9 ± 0.4
	**4**	8.6 ± 2.2	−43.9 ± 2.6	14.9 ± 0.8	1.0 ± 0.0	41.7 ± 3.3	7.2 ± 0.2
	**5**	6.7 ± 2.4	−46.3 ± 3.0	16.6 ± 1.0	0.8 ± 0.1	39.3 ± 2.6	6.7 ± 0.3
	**6**	6.8 ± 0.9	−45.1 ± 1.7	15.6 ± 0.6	0.8 ± 0.1	60.7 ± 6.8	7.7 ± 0.4
	**7**	8.6 ± 0.6	−43.5 ± 2.0	14.6 ± 0.7	1.0 ± 0.2	58.3 ± 4.0	7.8 ± 0.4
	**8**	7.6 ± 1.7	−38.7 ± 2.8	9.4 ± 0.8	1.1 ± 0.1	67.7 ± 4.5	6.4 ± 0.2
	**9**	8.8 ± 0.7	−41.4 ± 0.6	12.2 ± 0.1	0.9 ± 0.1	75.3 ± 3.4	9.8 ± 0.3

FP-data for controls: IC_{50} Me-α-D-Gal = 196 ± 7.8 μM; IC_{50} pNPG = 103 ± 6.1 μM.

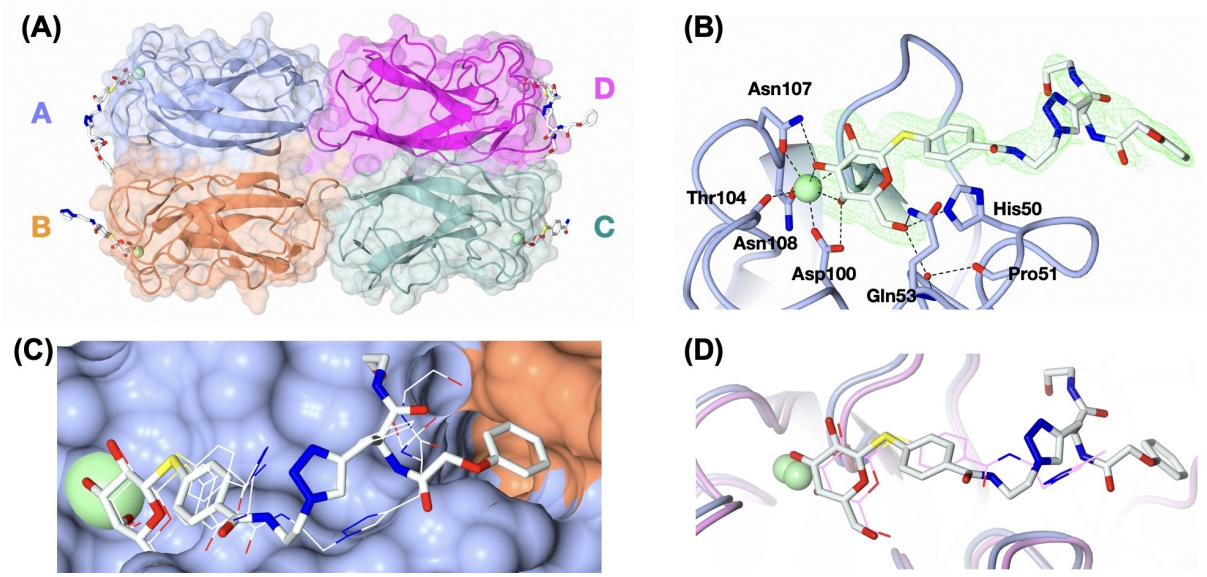


Figure 4. Crystal structure of LecA in complex with **1**. (A) The overall structure with **1** bound in all four monomers. (B) Binding pose of ligand **1** in site A with electron density displayed for the ligand. (C) Superposition of ligand conformations from all 4 binding sites. Ligand in site A is shown as sticks, whilst the others are shown as lines. (D) superposition of LecA structure in complex with **1** (protomer A) and structure in complex with peptide-based divalent ligand (PDB: 4CP9) reported by Winssinger *et al.*[29] **1** is shown in white, thick stick and the visible part of the peptide-based ligand reported by Winssinger *et al.*[29] in pink line.

coordination with the calcium ion.[27] The aglycon forms CH-π interaction with His50 (Figure 4B) in the same way as previously reported for aromatic galactosides and LecA.[21,32] However, superposition of the ligand in all four sites showed that the position of the phenyl ring varied slightly (Figure 4C).

The visible triazole rings in site A, B and D are surprisingly found in different positions and do not form productive contacts with the protein or water molecules (Figure 4B, C and Figure S4), which is in contrast to the previously reported divalent ligand with the similar phenyl aglycon and triazole design (Figure 1C),[29] where the triazole ring interacted with the side chains of His50, Tyr36 and Asp47 through a water bridge. Superposition of LecA in complex with **1** and that in complex with the peptide-based divalent ligand reported by Winssinger *et al.*[29] shows that the triazole ring oriented differently among the two ligands (Figure 4D). However, the conformation of our ligand **1** in protomer A is supporting our working hypothesis, where the terminal phenyl ring reaches the targeted central pocket. This data for the shortest derivative **1** also shows that an extension of the linker as implemented in molecules **2**–**6** could enable a better interaction with this pocket, which is consistently supported by our SPR and FP data for the extended molecules **3**–**5**.

Conclusions

We identified a central pocket localized between two monomers of the bacterial lectin LecA as a new target site for additional pharmacophores introduced into galactose-based inhibitors. Inhibitors were designed to bind simultaneously to the carbohydrate binding site and the central pocket. The nine synthesized inhibitors were analyzed for their interaction with LecA in the complementary biophysical assays such as ITC, SPR and a competitive binding assay. All compounds bound to LecA in the low micromolar range. Protein NMR data supported the interaction between the phenoxyacetate moiety with the central pocket as deduced from the CSP on the central pocket localized Trp33. Finally, the crystal structure of **1** in complex with LecA revealed that the additional aromatic moiety is oriented towards the central pocket. This is further supported by the increased affinity and binding enthalpy for the extended derivative **2** and the slight increase in binding affinity for the longer spacer containing compounds **3**–**5** suggests their interaction with the central pocket.

One fundamental question is raised by the fact that such a galactose-based inhibitor targeting the central pocket may block the single central pocket between two C2-symmetry related galactose binding sites and prevent efficient binding of a second copy of the ligand bound to the other galactose site. This question may be solved by future bivalent galactosides carrying an appropriately positioned single pharmacophore addressing the central pocket.

Experimental Section

***In silico* crystal structure analysis**: Twenty three crystallographic structures were retrieved from the Protein Data Bank[30] (PDB) (Table S1) and checked for the presence of the cavity termed here as central pocket. Cavity detection was performed using VolSite.[33] VolSite detects the shape and pharmacophore properties e.g. H-Bond donor/acceptor, hydrophobic and aromatic points of protein cavities and gives details about its buriedness, volume and a drug-ability score. The central pocket cavity was detected in 25 out of the 31 dimers that are available as crystallographic structures on PDB (Table S2).

Docking of 1 and molecular dynamics simulation: We used two crystal structures for docking. The structure of Winssinger's divalent LecA ligand (PDB code: 4cp9) co-crystallized with LecA contains parts of the ligand **1** and was used to guide the docking of the whole ligand **1**. It was further redocked on a LecA dimer structure with the central pocket in a more open conformation (PDB code: 4lke). Hydrogens were added to the protein structures using PROTOSS.[34] A 3D model of **1** carrying the phenoxy acetate side chain was prepared using CORINA.[35–37] Compound **1** was docked with the placed fragments method from Surflex.[38]

The selected docked conformation for the complex was parameterized according to the General Amber Force Field 2 (GAFF2)[39] using Ambertools 16.[40] RESP charges were fitted from the electrostatic potential and optimized at the HF/6-31G* level of theory by using Gaussian 09.[41] The protein was prepared according to the ff14SB forcefield,[42] using *tleap*. Ions (K^+, Cl^- and Ca^{2+}) were used to neutralize the complex. PBradii was set to mbondi3. The system was solvated with TIP3P[43] by using water molecules in an octahedral box, under periodic boundary conditions, with a distance of 12.0 Å between the protein and each face of the box. The systems were minimized in three steps by using the steepest descent, followed by a conjugate gradient after 2000 cycles: (i) position restraints for all heavy atoms of the complex (weight of 10 kcal/mol-Å^2) (ii) position restraints for the atoms of the backbone of LecA, (weight of 10 kcal/mol-Å^2); (iii) weak position restraints for the backbone atoms (weight of 5 kcal/mol-Å^2). After minimization, the system equilibration was performed on the solvent/ions in three steps: (i) the system was gradually heated from 30 to 100 K for 50 ps under the NVT ensemble with weak position restraints for the heavy atoms of LecA, by using a weight of 5 kcal/mol-Å^2; (ii) heating from 100 to 298.15 K for 250 ps under the NPT ensemble with weak position restraints for the heavy atoms of the protein using a weight of 5 kcal/mol-Å^2; (iii) MD under the NPT ensemble at constant temperature of 298.15 K for 1 ns with very weak position restraints for the protein heavy atoms, using a weight of 0.1 kcal/mol-Å^2.

The temperature was set to at 300 K using the Langevin dynamics with a collision frequency of 2 ps^{-1}. The pressure was kept constant using the Monte Carlo barostat with a relaxation time of 2 ps. SHAKE[44] was used to control bonds length involving hydrogen atoms. The production runs, under the NPT ensemble, consisted in one replica of ~240 ns and three independent replicas of 30 ns each, to verify the influence of sampling on ligand **1** conformations.

Visualization and hydrogen bond calculations for the trajectories were done using VMD.[45] Hydrophobic contacts were calculated using the cpptraj module available with AmberTools with the native contacts routine using a distance cutoff of 5 Å and a filter selection on non-polar amino acids. The short trajectories and the long MD yelled similar results regarding the mentioned analyzes. Images were rendered using VMD and UCSF-Chimera.[46]

Recombinant expression and purification of LecA: LecA expression and purification was performed as previously described.[20] The expression strain *E. coli* BL21(DE3) carrying the pET25pa1l[47] plasmid

was grown in 4 L LB containing ampicillin (100 μg/mL) at 37 °C and 180 rpm until an $OD_{600} \approx 0.5$ was reached. For induction of protein expression IPTG (0.25 mM) was added, and the culture was grown for additional 4 h at 30 °C. Bacterial cells were harvested by centrifugation (3000 g, 10 min, 4 °C) and the cell pellet was resuspended in TBS/Ca^{2+}-buffer (1 mM $CaCl_2$, 150 mM NaCl, 20 mM TRIS, 2.5 mM KCl, pH = 7.4) and PMSF (1 mM) and lysozyme (0.4 mg/mL) were added. Cell lysis was performed by 5 cycles in a homogenizer (M-110P, Microfluidics, USA), debris was removed by centrifugation (60 min, 10000 g, 4 °C), and LecA from the supernatant was purified on a galactosylated Sepharose CL-6B column using an Äkta start chromatography device. After washing with TBS/Ca^{2+}-buffer, bound LecA was eluted with 100 mM D-galactose in buffer and then dialyzed against fresh TBS/Ca^{2+}-buffer every day for 7 days to remove galactose. Protein concentration was determined by UV spectroscopy (ε = 27960 $M^{-1}\,cm^{-1}$, molecular weight 12893 $g\,mol^{-1}$, ExPASy ProtParam). LecA used for ITC was dialyzed against ddH_2O for 7 days and lyophilized afterwards prior to dissolving in the ITC buffer.

Fluorescence polarization assay: This assay was performed in analogy to Joachim *et al.*[20] using SulfoCy5Gal, an adapted Cy5 dye.[25] 10 μL of a solution containing SulfoCy5Gal (20 nM) and LecA (40 μM) in TBS/Ca^{2+}-buffer were added to 10 μL of inhibitor in the same buffer at concentrations from 2000–15.6 μM and containing 2% DMSO in a black 384-well plate (cat no 781900, Greiner Bio-One, Germany) in technical triplicates. Methyl α-D-galactopyranoside and 4-nitrophenyl β-D-galactopyranoside were included as positive controls. The plate was sealed (EASYseal, cat no 676001, Greiner Bio-One), centrifuged (1500 × g, 1 min, 25 °C) and incubated in a dark chamber under shaking conditions for 16 h at r.t. The foil was removed and the fluorescence intensity was measured with a PheraStar FS microplate reader (BMG Labtech GmbH, Germany) at ex. 590 nm and em. 675 nm. Data were analyzed using MARS Data Analysis Software (BMG Labtech GmbH, Germany) after subtracting blank values (LecA in TBS/Ca^{2+}-buffer with 1% DMSO) from the samples. Fluorescence polarization was calculated and the data were fitted according to the four-parameter variable slope model. This experiment was independently repeated three times and data were averaged and visualized using GraphPad PRISM version 5.

Isothermal titration calorimetry (ITC): Galactoside (1.0 or 1.5 mM) in TBS/Ca^{2+}-buffer was titrated into a stirred (700 rpm) LecA solution (75–232 μM) dissolved in the same buffer using a MicroCal iTC_{200} (Malvern, United Kingdom) instrument at 25 °C. The reference power was set to 5 μcal s^{-1}, the filter period to 5 s, and 20–39 injections (0.5–2 μL per injection) with an injection duration of 1 s were performed per experiment with a spacing of 240 s between each injection. For some titrations (in case of compounds **2** (1x), **4** (2x) and **8** (2x)), the syringe was refilled after the first titration ended, and the experiment was continued with the same sample cell contents to reach saturation, and the resulting files were merged using the MicroCal Concat ITC software. The first injection of each experiment was discarded, and the data were analyzed with the MicroCal Origin software using the one-site binding model. Individual titrations are depicted in Figure S1.

Protein-observed ^{19}F (PrOF) NMR: Labelled 5FW-LecA for PrOF NMR studies was produced recombinantly as reported previously.[31] NMR experiments were conducted on a Bruker AscendTM700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe in 3 mm tubes (Norell S-3-800-7). PrOF NMR was recorded using 200 μM 5FW-LecA in 20 mM Tris-HCl pH 7.8 with 150 mM NaCl, 10% D_2O and 100 μM TFA at 310 K. Changes in the spectra in presence of DMSO and upon addition of 3 mM ligand were processed and referenced to trifluoroacetic acid (TFA) as internal reference at −75.6 ppm. We considered only changes in chemical shift perturbation (CSP) upon ligand addition being two-fold greater than standard deviation (2x std. dev., 0.03 ppm) of the fluorine resonance.

Surface plasmon resonance (SPR): All experiments were performed on a BIACORE X100 at 25 °C. For activation and immobilization of LecA, buffer A (10 mM phosphate buffer + 2.7 mM KCl, 137 mM NaCl, 0.05% Tween 20 + 100 μM $CaCl_2$) was used as a running buffer. A CM5 BIACORE chip was activated by 3 injections of 1:1 NHS/EDC mixture (contact time = 540 s, flow rate = 10 μL/min) on channel 1 and 2, followed by injections of LecA (100 μg/mL in 10 mM sodium acetate buffer pH 4.5) on channel 2 only until the final binding response reached 2939 RU. Compounds were dissolved in DMSO to the final concentration of 100 mM and subsequently diluted to 5 mM in buffer A. The compounds were then prepared to the required analytic concentrations (0.2–200 μM) in buffer A supplemented with 5% DMSO. For multi-cycle kinetic analysis of the compounds, buffer A supplemented with 5% DMSO was used as a running buffer. The compounds were analyzed by injections of multiple concentrations on the immobilized LecA (contact time = 30 s, dissociation time = 60 s, flow rate 30 μL/min) and affinity analysis was performed using BIACORE evaluation software, which plots the binding response at the steady state against the analyzed concentrations and fits a non-linear curve to obtain K_d values. Individual experiments are depicted in Figure S3.

X-ray crystallography: Lyophilized LecA was dissolved to saturation in PBS at pH 7.4, containing 100 μM $CaCl_2$. 1 μL of the protein solution was mixed with 1 μL of the reservoir solution (20% PEG6000, 1 M LiCl, 100 mM sodium acetate pH 4.2, 5% DMSO containing 1 mM of compound **1** and the mixture was deposited on a siliconized glass circle cover slide (22 mm, Hampton research). Crystallization was performed by hanging drop vapor diffusion at 19 °C. Protein crystals were cryo-protected in 20% PEG6000, 1 M LiCl, 100 mM sodium acetate pH 4.2 supplemented with 20% ethylene glycol and flash frozen in liquid nitrogen. Data collection was conducted at SOLEIL PROXIMA 1 beamline (Saint Aubin, France). The recorded data were indexed, integrated and scaled at SOLEIL using XDS,[48] and merged using AIMLESS.[49] The structure was solved by molecular replacement in PHASER[50] using 1OKO as a searching template. The model was improved by manual re-building in COOT[51] and refinement in REFMAC5.[52] The model of compound **1** was manually built in AceDRG[53] and manually placed in the protein model. The final model was validated with MolProbity,[54] PDB-redo[55] and wwPDB validation server (http://validate-rcsb-1.wwpdb.org). Structural figures were prepared using CCP4MG.[56] Data merging, phasing and model re-building and refinement were performed through CCP4i2[57] graphical interface.

Supporting Information

The supporting information contains the synthesis of LecA inhibitors and ^1H- and ^{13}C-NMR spectra of new compounds, the tables of X-ray structures used for the analysis and description of the carbohydrate binding site and the central pocket, ITC and PrOF NMR raw data, the sensorgrams and affinity analyses of the SPR data as well as the X-ray structure of compound **1** with focus of each binding center of the LecA tetramer.

Acknowledgements

The authors acknowledge the financial support of the French-German ANR/DFG project (ANR-AAPG-2017) funded by the Agence Nationale de la Recherche (grant no. ANR-17-CE11-0048) and Deutsche Forschungsgemeinschaft (grant no. Ti756/5-1 and RA1944/

7-1) They also thank Glyco@Alps (ANR-15-IDEX-02) for support and synchrotron SOLEIL (Saint Aubin, France) for access and technical support at beamline PROXIMA1 and the Max-Planck Society for support. The authors also thank Emilie Gillon (CERMAV) for providing recombinant LecA for SPR and X-ray crystallographic experiments. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: carbohydrates · glycoconjugates · glycomimetics · lectin · LecA

[1] H. Fazzeli, R. Akbari, S. Moghim, T. Narimani, M. R. Arabestani, A. R. Ghoddousi, *J. Res. Med. Sci.* **2012**, *17*, 332–337.
[2] M. C. Plotkowski, M. Chevillard, D. Pierrot, D. Altemayer, J. M. Zahm, E. Colliot, G. Puchelle, *J. Clin. Invest.* **1991**, *87*, 2018–2028.
[3] A. Imberty, A. Varrot, *Curr. Opin. Struct. Biol.* **2008**, *18*, 567–576.
[4] N. Gilboa-Garber, *Methods Enzymol.* **1982**, *83*, 378–385.
[5] S. P. Diggle, R. E. Stacey, C. Dodd, M. Cámara, P. Williams, K. Winzer, *Environ. Microbiol.* **2006**, *8*, 1095–1104.
[6] D. Tielker, S. Hacker, R. Loris, M. Strathmann, J. Wingender, S. Wilhelm, F. Rosenau, K. E. Jaeger, *Microbiology* **2005**, *151*, 1313–1323.
[7] D. Davies, *Nat. Rev. Drug Discovery* **2003**, *2*, 114–122.
[8] J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Curr. Opin. Chem. Biol.* **2019**, *53*, 51–67.
[9] M. B. Calvert, V. R. Jumde, A. Titz, *Beilstein J. Org. Chem.* **2018**, *14*, 2607–2617.
[10] S. Wagner, R. Sommer, S. Hinsberger, C. Lu, R. W. Hartmann, M. Empting, A. Titz, *J. Med. Chem.* **2016**, *59*, 5929–5969.
[11] S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, *115*, 525–561.
[12] A. Bernardi, J. Jiménez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K. E. Jaeger, M. Lahmann, et al., *Chem. Soc. Rev.* **2013**, *42*, 4709–4727.
[13] D. Hauck, I. Joachim, B. Frommeyer, A. Varrot, B. Philipp, H. M. Möller, A. Imberty, T. E. Exner, A. Titz, *ACS Chem. Biol.* **2013**, *8*, 1775–1784.
[14] R. Sommer, D. Hauck, A. Varrot, S. Wagner, A. Audfray, A. Prestel, H. M. Möller, A. Imberty, A. Titz, *ChemistryOpen* **2015**, *4*, 756–767.
[15] A. Hofmann, R. Sommer, D. Hauck, J. Stifel, I. Göttker-Schnetmann, A. Titz, *Carbohydr. Res.* **2015**, *412*, 34–42.
[16] R. Sommer, T. E. Exner, A. Titz, *PLoS One* **2014**, *9*, 1–22.
[17] R. Sommer, K. Rox, S. Wagner, D. Hauck, S. S. Henrikus, S. Newsad, T. Arnold, T. Ryckmans, M. Brönstrup, A. Imberty, et al., *J. Med. Chem.* **2019**, *62*, 9201–9216.
[18] R. Sommer, S. Wagner, K. Rox, A. Varrot, D. Hauck, E.-C. Wamhoff, J. Schreiber, T. Ryckmans, T. Brunner, C. Rademacher, et al., *J. Am. Chem. Soc.* **2018**, *140*, 2537–2545.
[19] J. Rodrigue, G. Ganne, B. Blanchard, C. Saucier, D. Giguère, T. C. Shiao, A. Varrot, A. Imberty, R. Roy, *Org. Biomol. Chem.* **2013**, *11*, 6906.
[20] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933–7948.
[21] R. U. Kadam, D. Garg, J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J.-L. Reymond, *ACS Chem. Biol.* **2013**, *8*, 1925–1930.
[22] F. Pertici, N. J. de Mol, J. Kemmink, R. J. Pieters, *Chem. Eur. J.* **2013**, *19*, 16923–16927.
[23] E. Zahorska, S. Kuhaudomlarp, S. Minervini, S. Yousaf, M. Lepsik, T. Kinsinger, A. K. H. Hirsch, A. Imberty, A. Titz, *Chem. Commun.* **2020**, *56*, 8822–8825.
[24] S. Wagner, D. Hauck, M. Hoffmann, R. Sommer, I. Joachim, R. Müller, A. Imberty, A. Varrot, A. Titz, *Angew. Chem. Int. Ed.* **2017**, *56*, 16559–16564; *Angew. Chem.* **2017**, *129*, 16786–16791.
[25] S. Kuhaudomlarp, E. Siebs, E. Shanina, J. Topin, I. Joachim, P. da Silva Figueiredo Celestino Gomes, A. Varrot, D. Rognan, C. Rademacher, A. Imberty, et al., *Angew. Chem. Int. Ed.* **2021**, *60*, 8104–8114.
[26] T. Eierhoff, B. Bastian, R. Thuenauer, J. Madl, A. Audfray, S. Aigal, S. Juillot, G. E. Rydell, S. Müller, S. de Bentzmann, et al., *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12895–12900.
[27] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerová, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Lett.* **2003**, *555*, 297–301.
[28] R. U. Kadam, D. Garg, J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J.-L. Reymond, *ACS Chem. Biol.* **2013**, *8*, 1925–1930.
[29] A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winssinger, *Angew. Chem. Int. Ed.* **2014**, *53*, 8885–8889; *Angew. Chem.* **2014**, *126*, 9031–9035.
[30] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.
[31] E. Shanina, E. Siebs, H. Zhang, D. Varón Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2021**, *31*, 159–165.
[32] R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, et al., *Angew. Chem. Int. Ed.* **2011**, *50*, 10631–10635; *Angew. Chem.* **2011**, *123*, 10819–10823.
[33] J. Desaphy, K. Azdimousa, E. Kellenberger, D. Rognan, *J. Chem. Inf. Model.* **2012**, *52*, 2287–2299.
[34] S. Bietz, S. Urbaczek, B. Schulz, M. Rarey, *J. Cheminf.* **2014**, *6*, 1–12.
[35] "CORINA Classic – High-Quality 3D Molecular Models | MN-AM," can be found under https://www.mn-am.com/products/corina.
[36] C. H. Schwab, *Drug Discovery Today Technol.* **2010**, *7*, e245–e253.
[37] J. Sadowski, J. Gasteiger, G. Klebe, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000–1008.
[38] A. N. Jain, *J. Med. Chem.* **2003**, *46*, 499–511.
[39] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.
[40] D. A. Case, R. M. Betz, D. S. Cerutti, C. I. T. E., T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, et al., AMBER 2016, University of California, San Francisco **2016**.
[41] Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., Gaussian Inc. Wallingford CT, **2009**.
[42] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Genet.* **2006**, *65*, 712–725.
[43] W. L. Jorgensen, C. Jenson, *J. Comput. Chem.* **1998**, *19*, 1179–1186.
[44] V. Kräutler, W. F. Van Gunsteren, P. H. Hünenberger, *J. Comput. Chem.* **2001**, *22*, 501–508.
[45] W. Humphrey, D. Andrew, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33–38.
[46] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[47] B. Blanchard, A. Nurisso, E. Hollville, C. Tétaud, J. Wiels, M. Pokorná, M. Wimmerová, A. Varrot, A. Imberty, *J. Mol. Biol.* **2008**, *383*, 837–853.
[48] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.
[49] P. R. Evans, *Acta Crystallogr. Sect. D* **2011**, *67*, 282–292.
[50] A. J. McCoy, *Acta Crystallogr. Sect. D* **2006**, *63*, 32–41.
[51] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D* **2010**, *66*, 486–501.
[52] 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, *Acta Crystallogr. Sect. D* **2011**, *67*, 355–367.
[53] F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallogr. Sect. D* **2017**, *73*, 112–122.
[54] 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, *Acta Crystallogr. Sect. D* **2010**, *66*, 12–21.
[55] R. P. Joosten, F. Long, G. N. Murshudov, A. Perrakis, *IUCrJ* **2014**, *1*, 213–220.
[56] S. McNicholas, E. Potterton, K. S. Wilson, M. E. M. Noble, *Acta Crystallogr. Sect. D* **2011**, *67*, 386–394.
[57] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, et al., *Acta Crystallogr. Sect. D* **2018**, *74*, 68–84.

Manuscript received: October 28, 2021
Revised manuscript received: November 16, 2021
Accepted manuscript online: November 17, 2021
Version of record online: December 2, 2021