Glycosides

Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectins

Sakonwan Kuhaudomlarp[+], Eike Siebs[+], Elena Shanina, Jérémie Topin, Ines Joachim,
Priscila da Silva Figueiredo Celestino Gomes, Annabelle Varrot, Didier Rognan,
Christoph Rademacher, Anne Imberty,* and Alexander Titz*

Abstract: Because of the antimicrobial resistance crisis, lectins are considered novel drug targets. Pseudomonas aeruginosa utilizes LecA and LecB in the infection process. Inhibition of both lectins with carbohydrate-derived molecules can reduce biofilm formation to restore antimicrobial susceptibility. Here, we focused on non-carbohydrate inhibitors for LecA to explore new avenues for lectin inhibition. From a screening cascade we obtained one experimentally confirmed hit, a catechol, belonging to the well-known PAINS compounds. Rigorous analyses validated electron-deficient catechols as millimolar LecA inhibitors. The first co-crystal structure of a non-carbohydrate inhibitor in complex with a bacterial lectin clearly demonstrates the catechol mimicking the binding of natural glycosides with LecA. Importantly, catechol 3 is the first non-carbohydrate lectin ligand that binds bacterial and mammalian calcium(II)-binding lectins, giving rise to this fundamentally new class of glycomimetics.

Introduction

Glycoconjugates on the cell surface of host tissues serve as recognition patterns for microbial carbohydrate-binding proteins, lectins, in the initial steps of an infection with a pathogen.[1] The co-evolution between host glycoconjugates and pathogen receptors resulted in the ability of lectins to decipher the three-dimensional structure of complex branched oligosaccharides, referred to as "glycocode".[2] Among pathogens, *Pseudomonas aeruginosa* is an opportunistic Gram-negative bacterium inter alia responsible for pneumonia in immunocompromised patients, those suffering from cystic fibrosis (CF), or patients under ventilation.[3] *P. aeruginosa* starts forming a biofilm on its colonization site to shield the bacterium from antimicrobial treatment and the immune response and thus hampers combating the bacteria. Host recognition, tissue adhesion and biofilm formation of *P. aeruginosa* are mediated by two lectins: D-galactose-specific and L-fucose/D-mannose-specific LecA and LecB, respectively.[4] Both lectins are virulence factors and have been considered as targets for the development of carbohydrate-based anti-infective compounds and for the delivery of antibiotics via carbohydrate targeting moieties.[5] LecA is of special interest because of its ability to damage cell membranes and to facilitate bacterial cell internalization.[6]

LecA is a homotetramer containing one calcium ion in the carbohydrate-binding site coordinating amino acids and galactosides (Figure 1 A,B).[7] The target of this lectin is likely the globotriaosylceramide glycolipid that presents the Galα1-4Gal epitope.[8] The relatively low affinity for galactose or galactobiose ($K_d = 50$ μM) is overcome by high avidity with divalent or multivalent ligands that can cross-link the neighbouring binding sites ($K_d = 1$–20 nM).[5a, 9] However, because of the presence of a hydrophobic surface nearby

[*] S. Kuhaudomlarp,[+] J. Topin, A. Varrot, A. Imberty
Université Grenoble Alpes, CNRS, CERMAV
38000 Grenoble (France)
E-mail: anne.imberty@cermav.cnrs.fr

E. Siebs,[+] I. Joachim, A. Titz
Chemical Biology of Carbohydrates (CBCH)
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS),
Helmholtz Centre for Infection Research
66123 Saarbrücken (Germany)
and
Department of Chemistry, Saarland University
66123 Saarbrücken (Germany)
and
Deutsches Zentrum für Infektionsforschung (DZIF)
Hannover-Braunschweig (Germany)
E-mail: alexander.titz@helmholtz-hzi.de

E. Shanina, C. Rademacher
Department of Biomolecular Systems, Max Planck Institute of
Colloids and Interfaces
14424 Potsdam (Germany)
and
Institute of Chemistry and Biochemistry

Department of Biology, Chemistry and Pharmacy
Freie Universität Berlin
14195 Berlin (Germany)

J. Topin
Institute of Chemistry-Nice, UMR 7272 CNRS
Université Côte d'Azur
06108 Nice (France)

P. da Silva Figueiredo Celestino Gomes, D. Rognan
Laboratoire d'Innovation Thérapeutique
UMR 7200 CNRS-Université de Strasbourg
67400 Illkirch (France)

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for
the author(s) of this article can be found under:
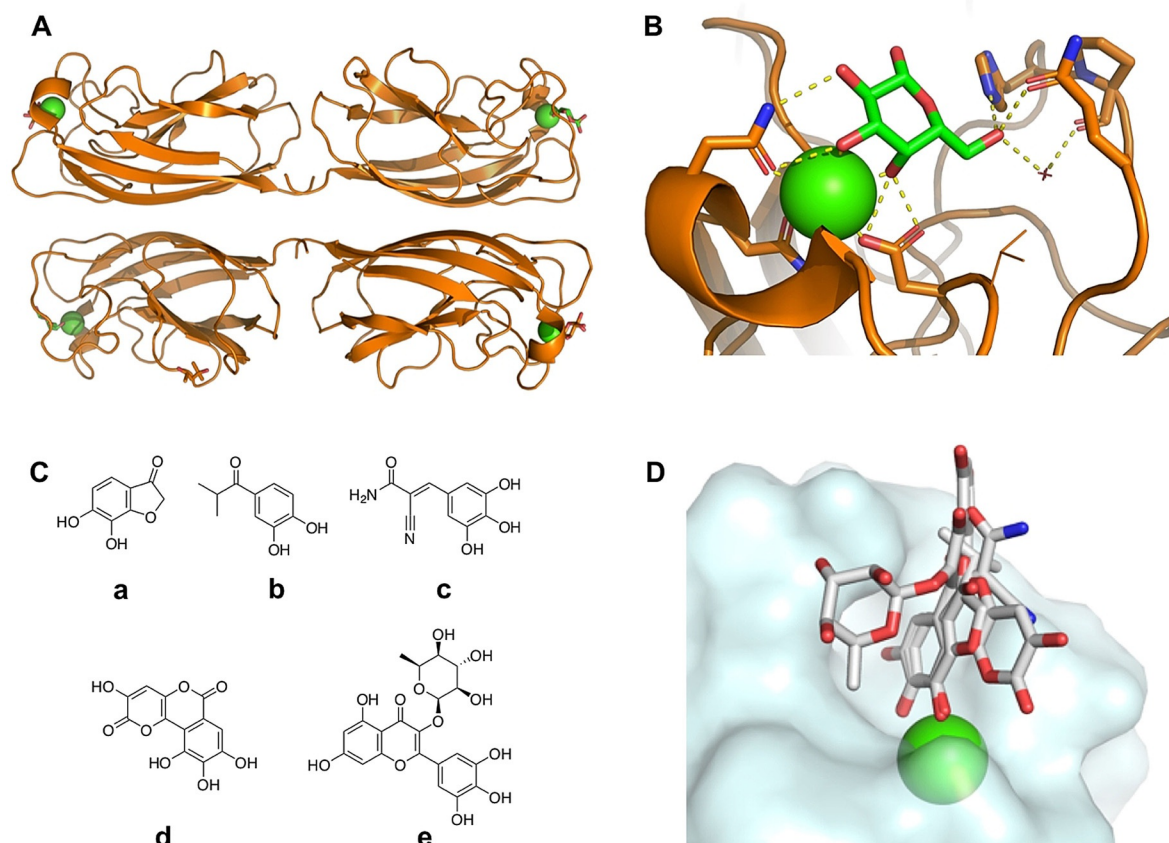https://doi.org/10.1002/anie.202013217.

Figure 1. A) Crystal structure of the homotetrameric LecA in complex with galactose (PDB code: 1OKO). B) View of calcium-ion dependent galactose coordination in the carbohydrate-binding site of LecA. C/D) Shown are docking poses of five catechol-containing compounds from the 60 best virtual screening hits. Compound b is identical to **21** in Table 2.

the galactose-binding site, monovalent galactosides with aromatic aglycons demonstrated high affinity for the lectin in the low micromolar range.[10] Recently, we reported an epoxide-carrying galactoside that binds covalently to a cysteine residue present in the carbohydrate-binding site of LecA.[11]

Targeting LecA is a promising way to counteract *P. aeruginosa* infections. The chemical properties of the binding pocket of LecA were used as an inspiration for the synthesis of many compounds, all maintaining the presence of a galactose residue for the primary calcium-dependent binding site.[5a,e,12] The development of non-carbohydrate mimics would be an alternative strategy, opening a large chemical space for novel inhibitors. Indeed, since carbohydrate-protein interactions are mostly governed by a complex arrangement of hydrogen bonds and hydrophobic contacts, the success of non-carbohydrate analogues is challenging. Non-carbohydrate glycomimetics were obtained for the model plant lectin Concanavalin A[13] and virtual screening resulted in ligands with higher affinities than the natural ligand mannose.[14] Apart from this work on a model lectin, only C-type lectins, that is, calcium-dependent animal lectins with roles in inflammation and immunity, have been targeted. Several compounds for the selectins have been obtained.[12,15] Glyco-

mimetics based on a shikimic acid scaffold have shown good inhibitory activity against the C-type lectin DC-SIGN.[16] The presence of a calcium ion is therefore a promising factor for a glycomimetic strategy against LecA.

Virtual screening is a powerful technique to identify potential ligands from large chemical databases.[17] Therefore, this approach was applied to screen the National Cancer Institute (NCI) Diversity IV database for identifying non-carbohydrate ligands directed to the galactose binding site of LecA. The best virtual screening hits were experimentally tested in a competitive binding assay based on fluorescence polarization,[18] identifying catechol-containing compounds as binder. Further, catechols were confirmed in orthogonal biophysical methods: a thermal shift assay (TSA), ligand- and protein-observed ¹⁹F (PrOF) and ¹H-¹⁵N TROSY NMR spectroscopy, surface plasmon resonance (SPR) and X-ray crystallography was employed to elucidate their binding modes to LecA. This work demonstrates the first example of a novel group of non-carbohydrate glycomimetic compounds targeting a bacterial lectin, which can be subsequently developed into a novel class of LecA inhibitors.

Results and Discussion

Virtual Screening of Ligands for LecA from NCI and Validation of Catechol-Containing Hits in a Competitive Binding Assay

A docking protocol was designed using Glide[19] and the 1.5 Å resolution crystal structure of LecA in complex with a galactoside (PDB code 3ZYH).[10a] After internal validation with ligands of known structures and affinities, the E-model score function was selected for ranking of orientations together with MM-GBSA[20] for ranking of energies (Figure S1). In total 1597 molecules from the National Cancer Institute (NCI) diversity set IV were docked into the binding site of LecA. Only those poses presenting a minimum of four contacts in the site, that is, hydrogen bonds and coordination to the calcium ion, were retained and ranked based on MM-GBSA rescoring function which resulted in the top 60 poses (Table S1). Those poses with at least two contacts to the calcium ion were manually clustered according to their similarity in binding mode into 15 groups and ranked as a function of their lower energy pose. We identified the 10 top clusters shown in Figure S2 that illustrate the encountered variety of chemical space. Interestingly, clusters 1, 5 and 6 corresponded to analogs of nucleotides, binding to LecA through 5-membered furanose rings. Other molecules bound to the calcium ion through glycerol side chains (cluster 3), sulfate groups (cluster 4) or a carboxylate (cluster 9). Of special interest was cluster 2 with five molecules (Figure 1C) that have a catechol motif to coordinate to the calcium ion with the two vicinal oxygen atoms in a very conserved geometry (Figure 1D). Finally, the lowest energy hit, 4-isobutyryl catechol (compound b in Figure 1C, **21** in Table 2), allows coordination to the calcium ion and an appropriate balance of hydrogen bonds and hydrophobic contacts that mimics the binding of galactose (see Figure S3).
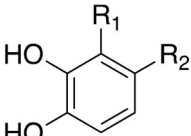
Initially, 40 compounds derived from virtual screening were selected based on structural diversity (Table S1), purchased and screened in the established competitive binding assay based on fluorescence polarization at a single concentration of 3.3 mM. Among the 40 compounds tested, 4-isobutyryl catechol (b in Figure 1C, compound **21** in Table 2) was identified as weak inhibitor (approx. 10% inhibition at 3.3 mM). It was the only molecule able to displace the fluorescent galactoside ligand in this assay. Therefore, both virtual and experimental screening suggest the catechol as a new scaffold for LecA.
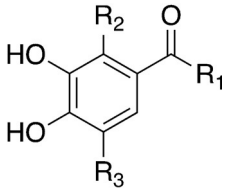
Experimental Screening of a Catechol Library

Despite the fact that some catechols are approved drugs, for example, the β_1-adrenergic agonist dobutamine or the antibiotic cefiderocol, this class of compounds is well known as frequent hitter and pan-assay interference compounds (PAINS).[21] Catechols are prone to oxidation into highly reactive quinones, which are reactive towards nucleophiles and may lead to an unspecific interaction with proteins.[22] However, the concept of a general exclusion of PAINS has been recently questioned based on extensive analysis of publicly available assay data[23] and a case-to-case consideration using orthogonal assay activity has been stated necessary.

The structure–activity relationship of catechols was followed up by purchasing 29 further commercial catechols with different substitution patterns and biophysical analysis using various orthogonal methods (Table 1, Table 2). Unsubstituted

Table 1: Catechols and LecA binding by fluorescence polarization (FP), thermal shift assay (TSA), and surface plasmon resonance (SPR).

	CAS no.	R₁	R₂	FP	FP with 2ME	TSA	SPR	PrOF NMR
				Inhibition [%] at 4 mM	Inhibition [%] at 4 mM	ΔT [K]	Efficiency [%] at 1 mM	CSP of W42 [ppm]
1	120-80-9	H	H	inconcl.	none	−4.5	5.2	n.s.
2	4018-65-9	Cl	H	inconcl.	none	−	2.4	n.s.
3	67984-81-0	CN	H	23 ± 2	23 ± 2	+1.8	20.6	0.045
4	363-52-0	F	H	inconcl.	none	−	1.3	n.s.
5	488-17-5	Me	H	inconcl.	none	none	8	0.055
6	14235-77-9	CH₂OH	H	inconcl.	11 ± 4	none	13.8	0.035
7	2144-08-3	OH	CHO	inconcl.	inconcl.	−3.3	17.2	n.s.
8	2138-22-9	H	Cl	inconcl.	none	−	2.5	n.s.
9	17345-61-8	H	CN	4 ± 1	7 ± 2	+0.5	6.1	0.030
10	367-32-8	H	F	inconcl.	none	−	2.6	n.s.
11	452-86-8	H	Me	inconcl.	none	−11.0	6.9	n.s.
12	3316-09-4	H	NO₂	37 ± 1	33 ± 2	autofluor.	56.1	0.065
13	98-29-3	H	*tert*-butyl	inconcl.	none	−15.4	7.9	0.130
14	331-39-5	H	E-CHCHCO₂H	inconcl.	none	−	7.3	n.s.
15	3843-74-1	H	E-CHCHCO₂Me	inconcl.	none	−	7.8	0.045
16	133550-30-8	H	E-CHC(CN)CONHBn	inconcl.*	inconcl.*	−	15.8 at 0.5 mM	n.s.

Protein-observed ^{19}F (PrOF) NMR spectroscopy was performed with 5FW-LecA. FP averages and std. dev. from 3 experiments, TSA averages from two experiments, * = tested at 25% DMSO, 2ME = 2-mercaptoethanol, CSP = chemical shift perturbation, *n.s.* = not significant, std. dev. < 0.03 ppm.

Table 2: Acylated catechols and LecA binding by fluorescence polarization (FP), thermal shift assay (TSA), and surface plasmon resonance (SPR).

	CAS no.	R₁	R₂	R₃	FP Inhibition [%] at 4 mM	FP with 2ME Inhibition [%] at 4 mM	TSA ΔT [K]	SPR Efficiency [%] at 1 mM	PrOF NMR CSP of W42 [ppm]
17	1197-09-7	Me	H	H	11±3	10±3	+0.5	14.4	0.035
18	99-50-3	OH	H	H	none	none	−	5.0	n.s.
19	99-40-1	CH₂Cl	H	H	inconcl.	none	−	22.8	0.095
20	3943-89-3	OEt	H	H	none	none	−	9.0	n.s.
21	5466-89-7	ⁱPr	H	H	10±2	11±2	+0.35	9.0	0.08
22	62-13-5	(CH₂)₂NHMe*HCl	H	H	inconcl.	insoluble	−	9.6	0.065
23	10425-11-3	Ph	H	H	39±5	39±3	+0.36	50.9	0.110
24	134612-84-3	pCl-C₆H₄-	H	H	29±3*	22±2*	−	insoluble	0.185
25	61445-50-9	2,4-(HO)₂-C₆H₃-	H	H	14±5	13±2	inconcl.	61.8	0.105
26	52479-85-3	3,4,5-(HO)₃-C₆H₂-	OH	H	inconcl.*	inconcl.*	−	142.6	0.125
27	1143-72-2	Ph	OH	H	inconcl.	inconcl.	−	13.2 at 0.5 mM	0.045
28	31127-54-5	pHO-C₆H₄-	OH	H	inconcl.	inconcl.*	−	38.5	0.035
29	5995-86-3	OH	H	OH	inconcl.	none	−	24.0	n.s.
30	3856-05-1	OCH₂ⁱPr	H	OH	8±2	8±1	−	16.4	0.055

Protein-observed ¹⁹F (PrOF) NMR spectroscopy was performed with 5FW-LecA. FP averages and std. dev. from 3 experiments, TSA averages from two experiments, * = tested at 25% DMSO, 2ME = 2-mercaptoethanol, CSP = chemical shift perturbation, n.s. = not significant, std. dev. < 0.03 ppm.

catechol **1** and various derivatives carrying electron-donating or -withdrawing substituents in different positions were obtained (**2**–**30**) among which are fourteen acylated catechols (**17**–**30**, including the hit 4-isobutyryl catechol **21**).

All compounds were tested in a modified competitive binding assay with a Cy5 dye conjugated to a galactoside **S7** (synthesis of **S7** see Scheme S1, binding of **S7** to LecA see Figure S4). This change in fluorescent dye was necessary to circumvent spectral overlap of some catechols with the previously used fluorescein conjugate.[18] After 1 h incubation, various catechols bearing electron-withdrawing substituents showed a weak inhibition of LecA. Surprisingly, the polarization measured for some of the tested compounds has changed dramatically after extended incubation (16 h). While some catechols showed a constant inhibition over time, in particular catechols bearing electron-donating substituents such as tert-butyl or methyl groups showed the strongest increase in inhibition.

To avoid the false-positive detection of an apparent inhibition due to an unspecific reactivity of potentially formed quinones, 2-mercaptoethanol was added to the competitive binding assay as an agent to trap those reactive species at equimolar concentration to the highest catechol concentration. Interestingly, the previously active electron-rich catechols were inactive under these conditions (Figure S5). In contrast, LecA inhibition of catechols carrying electron withdrawing substituents, such as nitrile **3**, nitro **12**, and ketone **21**, has not changed in presence of 2-mercaptoethanol and over time (Table 1, Table 2).

At a concentration of 4 mM, ten out of the 30 catechols tested showed a quantifiable and 2-mercaptoethanol-independent inhibition of LecA in this competitive binding assay

(Table 1, Table 2, Figure 2A). Benzoylcatechol **23** was the most active compound with an inhibition of 39%. Docking of **23** into the LecA binding site suggests that the benzoyl ring can interact with Gln53 (Figure S3A). In the competitive binding assay, this compound's activity is followed by nitro-catechol **12** and cyanocatechol **3** with inhibition of 33% and 23%, respectively. All three compounds were two- to fourfold more active than the initial in silico hit, isobutyryl catechol **21** (11% inhibition). Even though some of the catechols, such as 4-parachlorobenzoyl catechol (**24**), were only soluble at the test concentration of 4 mM in presence of 25% DMSO as a cosolvent, LecA bound successfully to the Cy5 galactoside **S7** and **24** showed 22% inhibition of LecA. These results support the notion that electron-deficient catechols can either allosterically inhibit the LecA-galactoside interaction or competitively bind to its carbohydrate-binding site.

Secondary Assays Verifying Binding of Catechols to LecA.

First, a subset of electron-rich and electron-deficient catechols was tested in a thermal shift assay to verify their interaction with LecA. In case of a ligand binding event, the thermal stability of a protein increases due to stabilizing interactions in the ligated state.[24] Therefore, the melting point of LecA was measured in absence of ligand (T_m 89.4°C), in presence of a galactoside (T_m 91.1°C) or thirteen catechols (Figure S6).

A general trend could be observed where the electron-deficient catechols, that is, nitriles **3** and **9**, acetyl **17**, isobutyryl **21** and benzoyl **23**, led to an increase of the melting point of LecA. In particular, 4-benzoyl catechol **23** and 3-
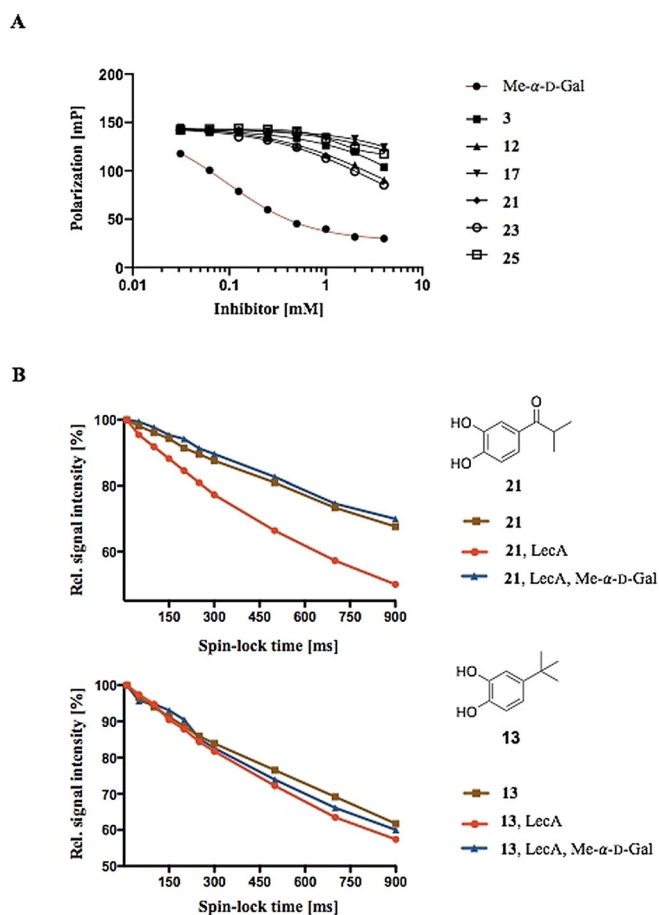
A

B

Figure 2. Analysis of catechol binding to LecA using: A) a competitive binding assay based on fluorescence polarization with catechols **3**, **12**, **17**, **21**, **23** and **25** and methyl α-D-galactoside as positive control; B) T1ρ relaxation NMR spectroscopy of isobutyryl catechol **21** and *tert*-butyl catechol **13**. Signal intensity decay of methyl protons plotted for **21** and **13** vs. spin-lock times.

cyano catechol **3** had the largest effect on the melting point of LecA with an increase of $+0.4\,°C$ and $+1.8\,°C$, respectively (Tables 1, 2, Figure S6). In contrast, the electron-rich catechols (methyl **11** and *tert*-butyl **13**) and the unsubstituted catechol **1** destabilized LecA and thus, lowered its melting temperature. A strong destabilization was observed in case of 4-*tert*-butyl catechol **13** ($-15.4\,°C$) and 4-methyl catechol **11** ($-11.0\,°C$).

As an orthogonal binding assay ligand-observed NMR spectroscopy experiments monitoring the $T_{1\rho}$ relaxation of the electron-deficient catechol, virtual screening hit isobutyryl **21**, and the electron-rich catechol 4-*tert*-butyl catechol **13** were measured. The $T_{1\rho}$ relaxation rates of a small molecule depend on its correlation time and hence can be used to detect binding to higher molecular weight receptors.[25] Therefore, $T_{1\rho}$ relaxation of isobutyryl catechol **21** or *tert*-butyl catechol **13** were measured in presence and absence of LecA, and both compounds showed an increase in relaxation with LecA (Figure 2B, Figure S7). However, the relaxation rate of isobutyryl catechol **21** was strongly increased in presence of LecA (red) and could be fully reversed to the free

state in presence of an excess of Me-α-D-Gal (blue). The effect of *tert*-butyl catechol **13** with LecA was much weaker and addition of Me-α-D-Gal had no significant effect. Since the binding of isobutyryl catechol **21** to LecA can be fully competed with a galactoside, this data suggests that electron-deficient catechols bind LecA in the carbohydrate-binding site. Moreover, this result is in agreement with our competitive binding assay with a fluorescent tracer, where isobutyryl catechol **21** inhibited LecA and the inhibition of *tert*-butyl catechol **13** was unspecific and absent in presence of a scavenging nucleophile.

Both assays, $T_{1\rho}$ relaxation NMR spectroscopy and the thermal shift assay, have further validated electron-deficient catechols as ligands of LecA and NMR gave strong hints for a direct galactose-competition as mode of action. Next, we quantified the catechols' binding affinity using surface plasmon resonance (SPR) as an orthogonal method and localized their binding site by protein-observed NMR spectroscopy and X-ray crystallography.

Quantification of direct catechol binding to LecA was performed using SPR analysis. Binding analyses were performed by injecting various catechols at two concentrations (0.2 and 1 mM, except for **16** and **27** which were injected at 0.2 and 0.5 mM). We identified 14 catechol hits by SPR, six of which were in agreement with other biophysical assays (nitrile **3**, hydroxymethyl **6**, nitro **12**, acetyl **17**, benzoyl **23** and dihydroxybenzoyl **25**, Figure 3 A, Figure S8). Subsequently, multicycle kinetic analyses were performed on these six prioritized hits to obtain the corresponding dissociation constants (K_d) (Figure 3 B, Figure S9). These data showed that hydroxymethyl **6**, acetyl **17** and dihydroxybenzoyl **25** have very weak affinity towards LecA (extrapolated K_d of >100 M) (Figure S9), whereas compounds nitrile **3**, nitro **12** and benzoyl **23** catechols have promising K_d values in the low millimolar range (Figure 3). The lowest K_d value was obtained for nitro **12** ($K_d = 0.56 \pm 0.34$ mM) followed by nitrile **3** ($K_d = 1.11 \pm 0.07$ mM) and the most active compound in the competitive binding assay, benzoyl **23**, was somewhat less active ($K_d = 3.46 \pm 0.41$ mM). Finally, the initial virtual screening hit isobutyryl **21** had a K_d of approx. 145 M corresponding to the results from the competitive binding assay. For weaker binding hits, the efficiency was used for comparison of the affinity, among which **26** showed the highest efficiency (142.6 %), which is suspiciously high for a small molecule. We hypothesized that this unexpected high binding response observed on SPR is due to a potential oxidation and unspecific interaction with the protein.

Protein-Observed NMR Reveals Catechols Targeting the Carbohydrate-Binding Site of LecA

Protein-observed ^{19}F (PrOF) NMR spectroscopy is a sensitive technique to spot weak binders in a low millimolar affinity range. Here, we used previously established PrOF NMR with 5-fluorotryptophan-labeled (5FW)-LecA[26] to determine the binding region of catechols. First, we measured 5FW-LecA in presence of *para*-nitrophenyl β-D-galactoside (pNPGal) as well-known ligand for the carbohydrate-binding
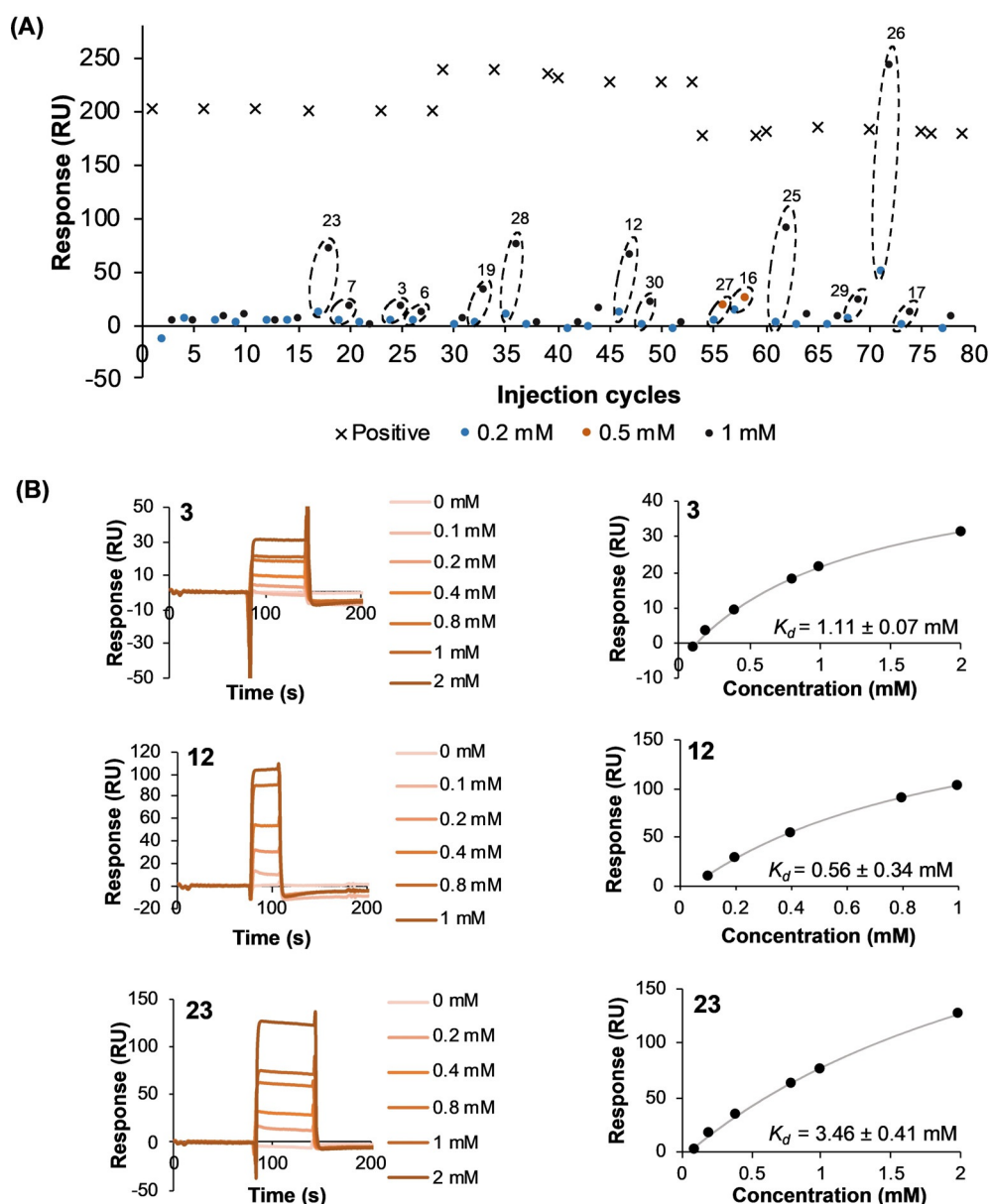
Figure 3. SPR analysis of catechol direct binding to LecA. A) The binding responses at steady state on the sensorgrams, when 0.2, 0.5 or 1 mM of catechols were injected (blue, orange and black dots, respectively), were plotted against injection cycles of compounds. Dose-response increases more than twofold are indicated by dotted circles. The catechol entries corresponding to the dose-response binding are indicated. Positive control (0.1 mM pNPGal represented by crosses) was injected to monitor the activity of immobilized LecA and enable data normalization. B) Multi-cycle kinetic analyses of the prioritized hits nitrile **3**, nitro **12**, benzoyl **23** by SPR. Left: Sensorgrams, Right: affinity analyses based on the data from the sensorgrams.

site of LecA. As expected, the resonance of the 5-fluoro-tryptophan located in the carbohydrate-binding site, W42, showed a clear chemical shift perturbation (CSP) towards a ligand-bound form in presence of pNPGal indicating that PrOF NMR can be used to spot the binding site of ligands (Figure 4).

Next, we validated 30 catechols in PrOF NMR (Table S2) and observed the perturbation of W42 in presence of 18 catechols that target the carbohydrate-binding site (Figure S10). Notably, the chemical shift perturbation of W42 in presence of the electron-deficient benzoyl catechol **23** was comparable to pNPGal (Figure S11) and nitrile **3** also induced

a chemical shift perturbation of the W42 resonance (Figure 4).

We further analyzed uniformly ^{15}N-labelled LecA in ^{1}H,^{15}N TROSY NMR experiments in presence and absence of two galactosides or four catechols (Figure 4). The fingerprint of the resulting chemical shift perturbations was highly similar for D-Gal and pNPGal compared to the three tested catechols **3**, **23** and **25** (Figure 4).

Taken together, compounds nitrile **3** and benzoyl **23** show a comparable binding of LecA in protein-observed NMR spectroscopy as the positive control pNPGal allowing us to
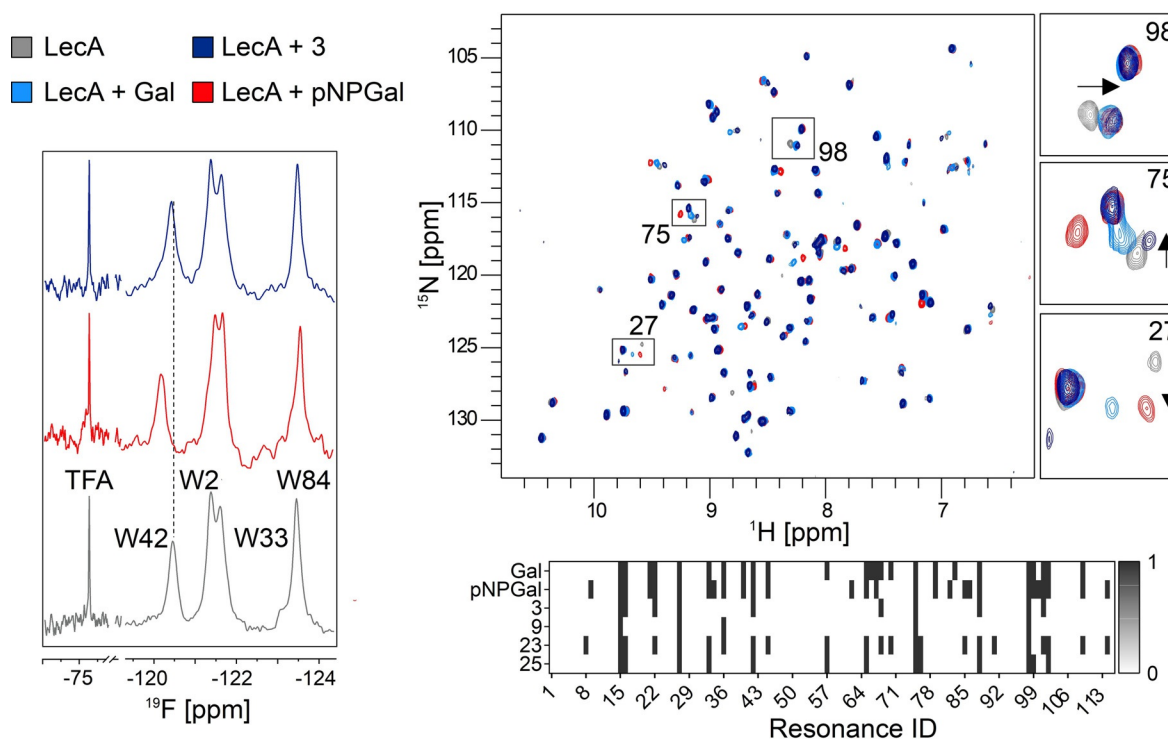
Figure 4. (left) Protein-observed ^{19}F (PrOF) NMR spectroscopy of nitrile **3** and pNPGal with 5FW-LecA; (right) $^1H,^{15}N$ TROSY-HSQC NMR spectra (*upper panel*) of uniformly ^{15}N-labelled LecA in absence and presence of ligands D-Gal, pNPGal or nitrile **3** and a plot with a 1:0 barcode (*bottom panel*) shows CSPs and changes in peak intensity of arbitrarily numbered resonances in LecA observed for ligands D-Gal, pNPGal, nitriles **3**, **9**, benzoyl **23** and dihydroxybenzoyl **25**. This data indicated catechols **3**, **9**, **23**, **25** bind LecA in a region similar to D-Gal and pNPGal suggesting binding to the carbohydrate-binding site of LecA.

hypothesize that both compounds target the carbohydrate-binding site of LecA.

X-Ray Crystallography Reveals Glycomimetic Binding Mode of One Catechol Hit

To further investigate the interaction between the catechol hits and LecA at atomic resolution, we co-crystallized LecA with the six hits that were identified among all biophysical assays, nitrile **3**, hydroxymethyl **6**, nitro **12**, acetyl **17**, benzoyl **23** and dihydroxybenzoyl **25**. We adapted a dry co-crystallization approach previously reported by Gelin et al., which involves deposition of the compound onto the crystallization slide and gentle evaporation of DMSO prior to dispensing the protein onto the dried compound.[27] This approach enables the deposition of high amounts of the dried compounds onto the slide, whilst avoiding complicated crystal manipulation and use of high DMSO concentrations. Protein crystals were obtained for co-crystallization of LecA with nitrile **3**, benzoyl **23** and dihydroxybenzoyl **25**. X-ray diffraction data were collected from crystals of LecA in complex with nitrile **3** (LecA-**3**) and in complex with dihydroxybenzoyl **25** (LecA-**25**) whereas the crystals of LecA with benzoyl **23** diffracted poorly (resolution > 4 Å) and were not further investigated. Both complex structures were successfully solved at 1.84 Å in space group $I2$ (for LecA-**3**) and $P2_122_1$ (for LecA-**25**) (Table S3).

The overall structure of LecA-**25** is very similar to LecA-galactose (global RMSD = 0.29 Å). Intriguingly, the electron density corresponding to compound **25** was found remotely from the galactose binding site (Figure S12A), which was unexpected considering its ability to outcompete the galactose-derived fluorescent probe. We hypothesized that the interaction of **25** with the remote site on LecA maybe a crystallization artefact since the molecule locates at the interface of the protein crystal contacts and forms only minimal interactions with the protein (Figure S12B).

The LecA-**3** structure is also very similar to the LecA-galactose complex (PDB: 1OKO) with a global RMSD of 0.8 Å (superposition of equivalent Cα atoms). Here, however, catechol **3** occupies the previously reported galactose binding site in monomers A and C of the LecA tetramer (Figure 5A and B). The two vicinal hydroxy groups of **3** coordinate the Ca^{2+} ion and form hydrogen-bonded interactions with the side chains of Asp100 and Asn107 and the backbone oxygen of Tyr36 (Figure 5D), the interaction of which imitates the roles of OH3 and OH4 of galactose in the LecA-galactose interaction (Figure 5C). A conserved water molecule (WAT1) forms hydrogen bond bridges between the nitrogen atom of the nitrile in catechol **3** and the oxygen atom the carbonyl group of Pro51 (Figure 5D). The nitrogen atom thus assumes the role of OH6 in the LecA-galactose interaction, where hydrogen bond bridges through WAT1 were identified between the OH6 group and Pro51 (Figure 5C). Shifting the nitrile group to the neighboring carbon atom, that is, to
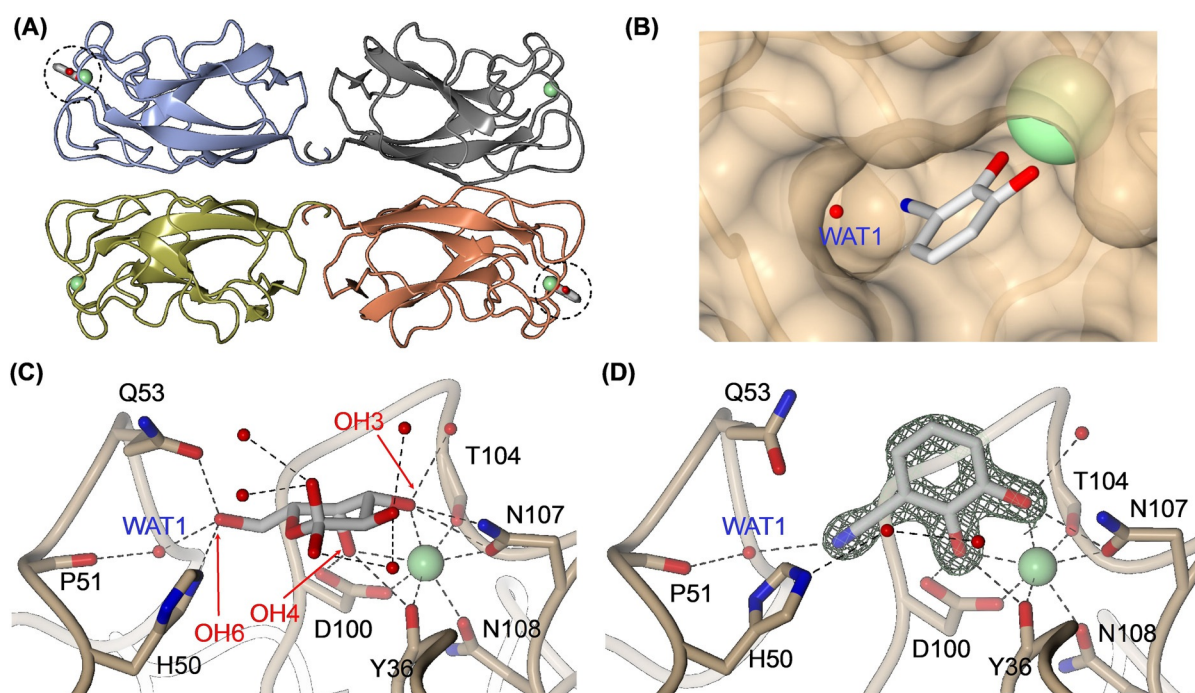
Figure 5. A) Overall structure of LecA-**3** complex (PDB code: 6YO3). 2 molecules of compound **3** are shown as sticks and the Ca²⁺ ions as green spheres. Dotted circles indicate the binding sites of **3**. Each LecA monomer is depicted as cartoon in different colors. B) Surface display of compound **3** at the binding site. C) Interaction of galactose with LecA (PDB code: 1OKO). D) Interaction of **3** with LecA, electron density is displayed at 1σ.

position R2 in compound **9** instead of position R1 in compound **3**, resulted in approximately a threefold-reduction of the SPR efficiency and inhibition in the FP assay (Table 1), which further substantiates the essence of the positioning and the role of the nitrile group in the LecA-**3** interaction.

To investigate if catechols interact with other Ca²⁺-binding lectins and could be used as a general motif for glycomimetics, we performed ¹H-¹⁵N HSQC NMR with ¹⁵N-labeled Langerin CRD (Figure 6A). We observed that **3** promoted chemical shift perturbations (CSPs) of amino acids located in the carbohydrate-binding site of Langerin (Figure 6B), but also affected amino acids that are located at a certain distance. The magnitude of CSPs observed for **3** (tested at 1 mM due to solubility restrictions) was generally weaker than mannose which was tested at 30 mM. Given the allosteric nature of some calcium binding lectins such as Langerin,[16a, 28] changes in the resonances located far away from the carbohydrate-binding site can be correlated to changes in the binding site. Potentially, **3** can be located in the binding site, or in a remote site that triggers the allosteric network.

To demonstrate that **3** targets the carbohydrate-binding site of Langerin, we performed ¹H STD competition NMR experiments (Figure 6C). For this, we recorded STD spectra of **3** without or with Langerin ECD and upon addition of mannose as the competitor. As result, we observed a 65% reduction in peak intensities of resonances from compound **3** in STD NMR spectra in presence of mannose indicating its partial competition from the carbohydrate-binding. This is in agreement with ¹H-¹⁵N HSQC data, where **3** has been observed to alter resonances from both the carbohydrate-

and secondary binding sites in Langerin. Thus, these data reveal catechol **3** as a general glycomimetic motif for a bacterial and a mammalian calcium-binding lectin.

Conclusion

In this work, a non-carbohydrate glycomimetic has been identified for a relevant anti-infective drug target. We have established catechols as mimicry of the carbohydrate ligand, binding to the protein via the crucial calcium-ion present in the carbohydrate-binding site of LecA, a key lectin involved in infection mechanisms of the ESKAPE pathogen *P. aeruginosa*. Therefore, the catechols have good potential beyond LecA and may also serve as glycomimetics for the large family of animal C-type lectins and other bacterial lectins relying on calcium-mediated glycan recognition. Target specificity required for a selective drug can be achieved by further optimizing a central catechol motif addressing the carbohydrate-binding site.

Despite the widespread belief, catechols are not necessarily PAINS when care is taken in analysing their binding properties in orthogonal assays under appropriate experimental conditions. We showed that electron-deficient catechols are stable under the conditions tested and their interaction with LecA has been confirmed in several orthogonal assays. In contrast, electron-rich catechols proved to be unspecific binders and their observed effect on LecA could be reversed in presence of a competing nucleophile, 2-mercaptoethanol.
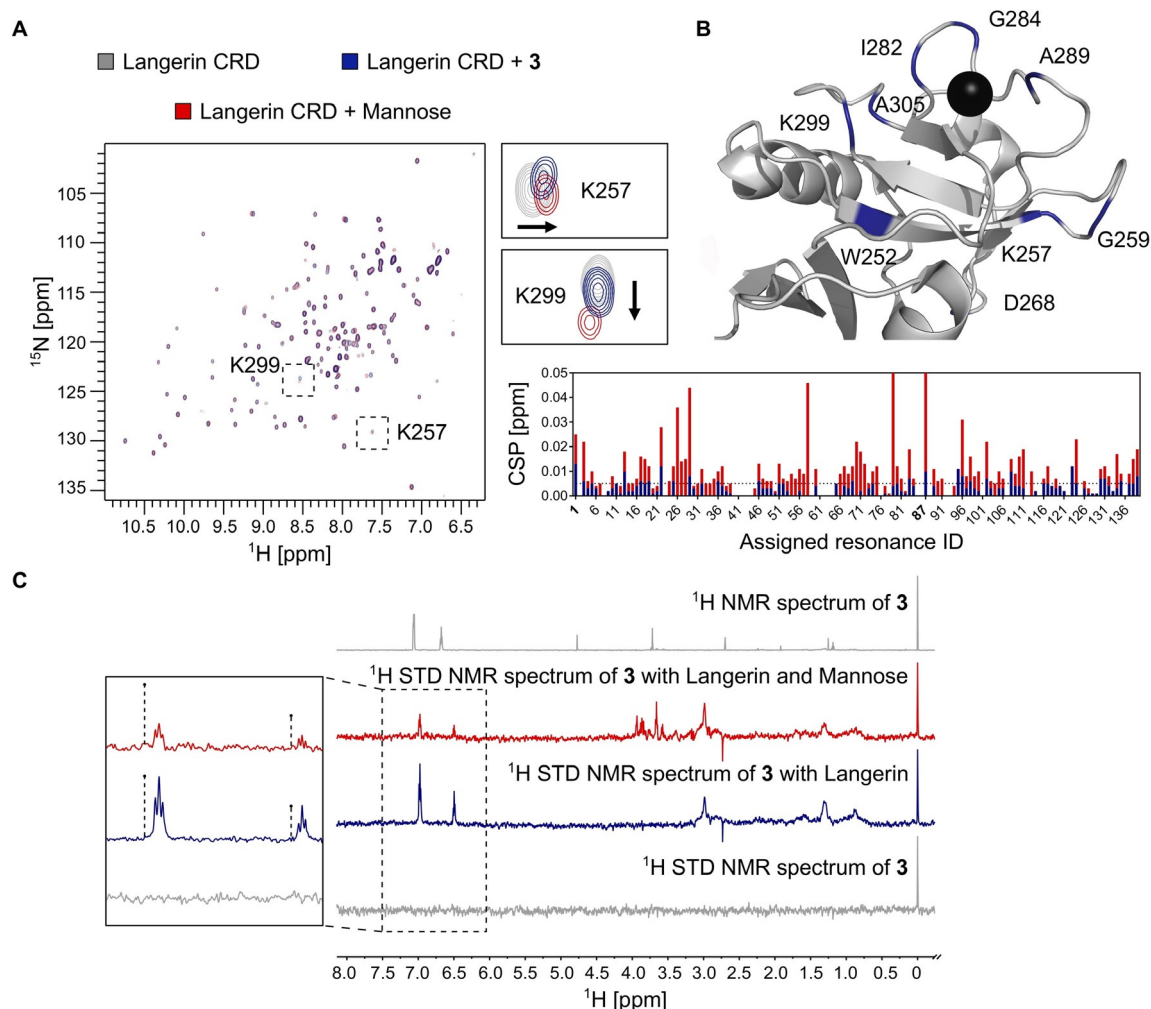
Figure 6. A) ^1H-^{15}N HSQC NMR spectra of uniformly ^{15}N-labelled human Langerin CRD in absence of ligands (*grey*) and in presence of mannose (*red*) or catechol **3** (*blue*). The two resonances K257 and K299 are magnified and shown as an example for the interaction of **3** (*blue*) with Langerin, which is in the similar manner to mannose (*red*). B) The perturbed resonances in the CSP plot (CSP > 0.005 p.p.m., for correlation of resonance ID and amino acid see Table S4) were used for mapping the binding site of **3** on a structure of Langerin CRD (PDB: 3P5D). K257 and K299 are highlighted in the CSP plot as 1 and 87, respectively. C) The competition ^1H STD NMR experiment with mannose was performed with **3** in presence and absence of Langerin ECD. Binding of **3** to the carbohydrate-binding site of Langerin has been confirmed as indicated by decreasing peak intensities in the STD NMR spectrum upon addition of mannose.

From an initial virtual screening approach for the bacterial lectin LecA, we have identified electron-deficient catechols as non-carbohydrate inhibitors for the lectin LecA. We validated our hits in various orthogonal biophysical assays. The binding of the initial catechol hit isobutyryl **21** derived from virtual screening was confirmed in a competitive binding assay based on fluorescence polarization. This assay and a thermal shift assay were used to demonstrate that only catechols with electron-withdrawing groups are robust inhibitors of LecA, while catechols with electron-donating substituents showed protein destabilization and unspecific inhibition which was absent in presence of a nucleophilic scavenger. This specificity of catechols was further confirmed in ligand-observed $T_{1\rho}$-relaxation NMR experiments for electron-deficient isobutyryl catechol **21** and electron-rich 4-*tert*-butylcatechol **13**. The binding affinity of catechols was determined by SPR and led to K_d-values in the millimolar range, for example, nitrile **3** with a K_d of 1.11 ± 0.07 mM and

nitro **12** with a K_d of 0.56 ± 0.34 mM. These affinities of the fragment hits (MW 135–155 g mol^{-1}) result in ligand efficiencies of 0.40 for both compounds, providing a good basis for future fragment growing. ^{19}F-protein observed NMR spectroscopy revealed that the catechols with electron-withdrawing substituents were binding to the carbohydrate-binding site of LecA. Finally, X-ray analysis of LecA in complex with nitrile **3** unambiguously confirmed catechols as galactose-mimicking ligands located in the carbohydrate-binding site. The structure revealed that catechols were coordinating to the calcium ion via their vicinal hydroxy groups, and the nitrile group underwent a hydrogen-bonding interaction with a crystal bound water molecule, highly homologous to the primary hydroxy group of galactose in complex with LecA.

This work sets the basis for future optimization of the catechol scaffold to increase binding potency and drug-like properties of a possible antimicrobial agent. Therefore, this work sets a new starting point for the development of future

LecA inhibitors against biofilm-associated *P. aeruginosa* infections by developing new non-carbohydrate LecA glycomimetics and provides a structure-based rationale for the development of lectin-inhibitors in general.

Moreover, catechol **3** proved to target the carbohydrate-binding site of both Ca²⁺-binding bacterial (LecA) and mammalian (Langerin) lectins. Given that Ca²⁺-binding lectins have been considered "undruggable" for a long time,[29] we believe catechols challenge this paradigm as a novel class of non-carbohydrate glycomimetics. Further growing of this low affinity fragment hit, catechol **3**, into lead structures will provide high target affinity. Structure-based design of additional functional groups will provide the specificity towards the lectin of interest for the desired indication.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Keywords: carbohydrates · catechol · glycomimetic · lectin · PAINS

[1] a) A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.* **2004**, *6*, 221–228; b) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674.

[2] D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana, Jr., H. J. Gabius, *Biochim. Biophys. Acta Gen. Subj.* **2015**, *1850*, 186–235.

[3] S. de Bentzmann, P. Plesiat, *Environ. Microbiol.* **2011**, *13*, 1655–1665.

[4] a) S. P. Diggle, R. E. Stacey, C. Dodd, M. Camara, P. Williams, K. Winzer, *Environ. Microbiol.* **2006**, *8*, 1095–1104; b) D. Tielker, S. Hacker, R. Loris, M. Strathmann, J. Wingender, S. Wilhelm, F. Rosenau, K.-E. Jaeger, *Microbiology* **2005**, *151*, 1313–1323.

[5] a) S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, *115*, 525–561; b) J. Meiers, E. Zahorska, T. Röhrig, D. Hauck, S. Wagner, A. Titz, *J. Med. Chem.* **2020**, *63*, 11707–11724; c) R. Sommer, K. Rox, S. Wagner, D. Hauck, S. Henrikus, S. Newsad, T. Arnold, T.

Ryckmans, M. Bronstrup, A. Imberty, A. Varrot, R. W. Hartmann, A. Titz, *J. Med. Chem.* **2019**, *62*, 9201–9216; d) R. Sommer, S. Wagner, K. Rox, A. Varrot, D. Hauck, E.-C. Wamhoff, J. Schreiber, T. Ryckmans, T. Brunner, C. Rademacher, R. W. Hartmann, M. Brönstrup, A. Imberty, A. Titz, *J. Am. Chem. Soc.* **2018**, *140*, 2537–2545; e) S. Wagner, R. Sommer, S. Hinsberger, C. Lu, R. W. Hartmann, M. Empting, A. Titz, *J. Med. Chem.* **2016**, *59*, 5929–5969.

[6] T. Eierhoff, B. Bastian, R. Thuenauer, J. Madl, A. Audfray, S. Aigal, S. Juillot, G. E. Rydell, S. Müller, S. de Bentzmann, A. Imberty, C. Fleck, W. Römer, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12895–12900.

[7] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerová, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Lett.* **2003**, *555*, 297–301.

[8] 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.

[9] a) M. Bergmann, G. Michaud, R. Visini, X. Jin, E. Gillon, A. Stocker, A. Imberty, T. Darbre, J. L. Reymond, *Org. Biomol. Chem.* **2016**, *14*, 138–148; b) 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; c) 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.

[10] a) R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, M. Camara, A. Stocker, T. Darbre, J. L. Reymond, *Angew. Chem. Int. Ed.* **2011**, *50*, 10631–10635; *Angew. Chem.* **2011**, *123*, 10819–10823; b) 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; c) J. Rodrigue, G. Ganne, B. Blanchard, C. Saucier, D. Giguère, T. C. Shiao, A. Varrot, A. Imberty, R. Roy, *Org. Biomol. Chem.* **2013**, *11*, 6906–6918.

[11] S. Wagner, D. Hauck, M. Hofmann, I. Joachim, R. Sommer, R. Müller, A. Imberty, A. Varrot, A. Titz, *Angew. Chem. Int. Ed.* **2017**, *56*, 16559–16564; *Angew. Chem.* **2017**, *129*, 16786–16791.

[12] J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Curr. Opin. Chem. Biol.* **2019**, *53*, 51–67.

[13] M. Goel, D. Jain, K. J. Kaur, R. Kenoth, B. G. Maiya, M. J. Swamy, D. M. Salunke, *J. Biol. Chem.* **2001**, *276*, 39277–39281.

[14] K. T. Welch, T. A. Turner, C. E. Preast, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6573–6575.

[15] a) N. Kaila, B. E. Thomas IV, *Med. Res. Rev.* **2002**, *22*, 566–601; b) R. Kranich, A. S. Busemann, D. Bock, S. Schroeter-Maas, D. Beyer, B. Heinemann, M. Meyer, K. Schierhorn, R. Zahlten, G. Wolff, E. M. Aydt, *J. Med. Chem.* **2007**, *50*, 1101–1115.

[16] a) J. Aretz, H. Baukmann, E. Shanina, J. Hanske, R. Wawrzinek, V. A. Zapol'skii, P. H. Seeberger, D. E. Kaufmann, C. Rademacher, *Angew. Chem. Int. Ed.* **2017**, *56*, 7292–7296; *Angew. Chem.* **2017**, *129*, 7398–7402; b) M. J. Borrok, L. L. Kiessling, *J. Am. Chem. Soc.* **2007**, *129*, 12780–12785; c) K. C. Garber, K. Wangkanont, E. E. Carlson, L. L. Kiessling, *Chem. Commun.* **2010**, *46*, 6747–6749.

[17] O. Slater, M. Kontoyianni, *Expert Opin. Drug Discovery* **2019**, *14*, 619–637.

[18] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933–7948.

[19] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759.

[20] Z. Yu, M. P. Jacobson, R. A. Friesner, *J. Comput. Chem.* **2006**, *27*, 72–89.

[21] J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, *53*, 2719–2740.

[22] L. Sleno, R. F. Staack, E. Varesio, G. Hopfgartner, *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2301–2311.

[23] S. J. Capuzzi, E. N. Muratov, A. Tropsha, *J. Chem. Inf. Model.* **2017**, *57*, 417–427.

[24] M. C. Lo, A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, *Anal. Biochem.* **2004**, *332*, 153–159.

[25] M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, G. Siegal, *Nat. Rev. Drug Discovery* **2008**, *7*, 738–745.

[26] E. Shanina, E. Siebs, H. Zhang, D. V. Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2020**, *31*, 159–165.

[27] M. Gelin, V. Delfosse, F. Allemand, F. Hoh, Y. Sallaz-Damaz, M. Pirocchi, W. Bourguet, J.-L. Ferrer, G. Labesse, J.-F. Guichou, *Acta Crystallogr. Sect. D* **2015**, *71*, 1777–1787.

[28] J. Hanske, S. Aleksic, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *J. Am. Chem. Soc.* **2016**, *138*, 12176–12186.

[29] a) A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730; b) A. P. Russ, S. Lampel, *Drug Discovery Today* **2005**, *10*, 1607–1610.