Bacteria Lectin Recognition Towards Fucose Binding Motifs Highlights the Impact of Presenting Mucin Core Glycopeptides

Sandra Behren#,[a, b] Jin Yu#,[b, c]   Christian Pett,[a, b] Manuel Schorlemer,[a, b] Viktoria Heine, [d] Thomas Fischöder, [d] Lothar Elling, [d] Ulrika Westerlind*[a, b]

---

[a]   MSc. Sandra Behren, Dr. Christian Pett, Dr. Manuel Schorlemer, Prof. Ulrika Westerlind

Department of Chemistry, Umeå University

90187 Umeå, Sweden

E-mail: ulrika.westerlind@umu.se

[b]   MSc. Sandra Behren, Dr. Jin Yu, Dr. Manuel Schorlemer, Prof. Ulrika Westerlind

Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V.

44227, Dortmund, Germany

[c]   Dr. Jin Yu

Glycosciences Laboratory,

Imperial College London,

W12 0NN UK

[d]   Dr. Viktoria Heine, Dr. Thomas Fischöder, Prof. Lothar Elling

Laboratory for Biomaterials,

Institute of Biotechnology and Helmholtz-Institute for Biomedical Engineering,

RWTH Aachen University, Pauwelsstraße 20,

52074 Aachen, Germany

E-mail: l.elling@biotec.rwth-aachen.de

# These authors contributed equally to the manuscript.

Supporting information for this article is given via a link at the end of the document.

Abstract: Mucin glycoproteins are essential components of the mucosal protective barrier, which constantly senses and clears the host from pathogens. Throughout evolution, bacteria and virus have developed strategies to modulate and penetrate the mucosal barrier and cause virulence by interacting with the glycans of membrane-bound mucins at the epithelial cell-surface. These interactions may promote bacteria cell-adhesion, biofilm formation, protein toxin delivery, or cause an inflammatory environment. *O*-fucosylated glycan epitopes are commonly found on mucin glycoproteins, and are key ligands of many bacterial and viral lectins (glycan binding proteins). Herein we describe a chemoenzymatic synthesis strategy to efficiently prepare an extensive library of fucosylated mucin core tandem repeats glycopeptides to elucidate the fine fucose-binding specificities of the *Pseudomonas aeruginosa* lectin LecB and the *Clostridium difficile* toxin A. Therefore, glycan core structures were decorated with terminal Lewis and H-antigens, which play critical roles in infection biology. The fucosylated mucin glycopeptides were applied in microarray binding studies to explore the importance of the glycan and peptide backbone presentation of these terminal antigens in binding interactions with the two bacterial lectins.

Introduction

Mucin glycoproteins are central players in the host-defense machinery directed against invading pathogens.[1] In addition to the dense glycan shield formed by membrane-bound mucin glycoproteins on epithelial cell surfaces, secreted mucins cover the epithelial tissues as the major constituents of the mucus, which protects the epithelial tissues from invading pathogens. Thus, the majority of potential infections is prevented by mucus clearance, but bacteria and viruses have co-evolved with the human host and developed strategies to promote immune escape and virulence.[2] Pathogenic bacteria make use of the mucin carbohydrate ligands to adhere to the host cell surface. Additionally, by using specific glycosidases, bacteria may manipulate the glycan structures of the host and thereby trigger inflammation, promote biofilm formation or build-up their own glycan shield to prevent immune cell recognition.[1] Mucin glycans on the host epithelial cells are also targets of bacterial protein toxins that promote cell adhesion to allow intracellular protein toxin delivery.[2a]

Fucose residues decorate terminal positions of mucin carbohydrate ligands and other glycoconjugates, which are typically presented in the blood group A-, B- and H-antigens or on Lewis epitopes. These fucosylated structures are critical players in diverse bacteria and virus interactions. For example, blood group O-individuals are presenting the H-antigen structure, which increases the susceptibility for severe cholera infection caused by *Vibrio cholorae* and gastroenteritis caused by the Norwalk virus.[3] Patients suffering from cystic fibrosis (CF) or chronic obstructive pulmonary disease (COPD) have, in addition to a high mucin secretion, an increased presentation of fucosylated glycans on the lung mucins.[4] These fucosylated structures are targets for lectins of many bacterial pathogens such as the soluble lectin LecB from *Pseudomonas aeruginosa (PA)*[5]

and the of *Clostridium difficile* toxin A (TcdA) that can facilitate toxin delivery causing gastrointestinal symptoms by interacting with fucosylated host glycans.[6] *Pseudomonas aeruginosa* is a Gram-negative opportunistic bacterium that often infects patients suffering from chronic obstructive pulmonary disease (COPD) and cystic fibrosis (CF). The lectin LecB is secreted from *Pseudomonas aeruginosa* and causes critical virulence due to its involvement in bacterial biofilm formation.[7] LecB is a tetrameric adhesin that specifically recognizes and binds to L-fucosides, which are included in Lewis a (Le^a) antigen structures and have in previous studies shown to be preferred LecB ligands.[8] Glycoconjugates and glycomimetic structures interfering with LecB binding are interesting candidates in the development of new anti-microbial therapies. Thereby, multivalent presentation of short glycan epitopes on peptides, dendrimers or polymers has been found to efficiently mimic natural glycan presentation and increase lectin avidity.[9] In a previous work, multivalent presented α-L-fucose monosaccharides were applied for potential LecB inhibition.[10] Also, glycodendrimers presenting the terminal Le^a epitope were explored regarding their ability to inhibit LecB.[10b] Titz and co-workers developed glycomimetic *C*-glycosides of amides and sulfonamides to elucidate LecB structure-activity relationships.[11] *Clostridium difficile* is a Gram-positive opportunistic bacterium that often causes recurrent mild to severe gastrointestinal infections in immune compromised patients.[12] Toxin A (TcdA) and toxin B (TcdB) are secreted from *C. difficile* and are two multi-domain toxins that inactivate critical host GTPases, including Rac, Cdc42 and the Rho

A protein family, by monoglucosylation, thus inducing cell death.[13] Therapeutic treatments to combat *C. difficile* infections are limited and often rely on strong antibiotics or therapeutic antibody administration. Consequently, better knowledge of TcdA binding interactions would facilitate the development of glycomimetic anti-adhesives. Inhibition of TcdA-mediated cell toxicity was recently demonstrated with multivalent bovine serum albumin (BSA) neo-glycoproteins bearing the Lewis antigens Le^y and Le^x.[6b] A mucin-type fusion protein carrying the Gal-α-1,3 epitope was also found to interact with *C. difficile* TcdA and showed inhibition by rabbit erythrocyte hemagglutination.[14] In the current work, we aimed to explore the molecular fine-specificities behind the TcdA interactions with fucosylated mucin tandem repeat glycopeptides.

Besides their importance in lectin binding interactions, the presentation of the unique terminal mucin glycan epitopes by the underlying core structures and the peptide backbone is potentially essential for fine binding specificities and biological functions of these lectins. However, this knowledge is often neglected in studies of lectin interactions, even if these structures may define the glycan orientation, structural rigidity or possible limitations for ligand recognitio

In this work, we employed a library of fucosylated structures displayed on glycan cores of mucin 1 (MUC1) and mucin 5B (MUC5B) tandem repeat peptides to elucidate the roles of different terminal fucose motifs, and of glycan presentation on different glycosylation sites of the peptide
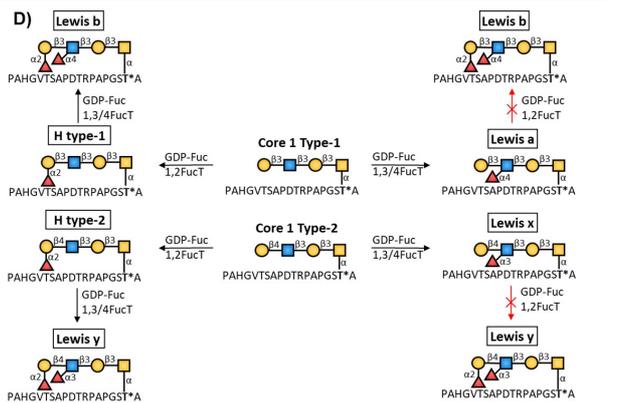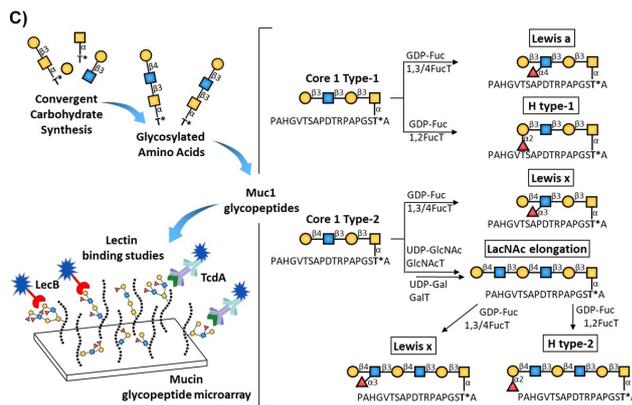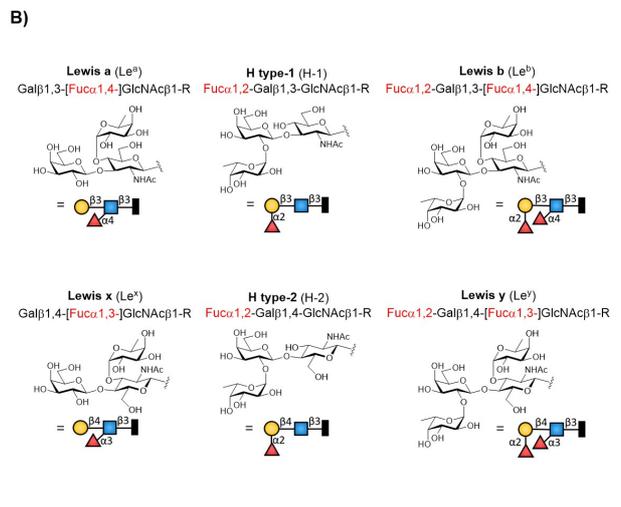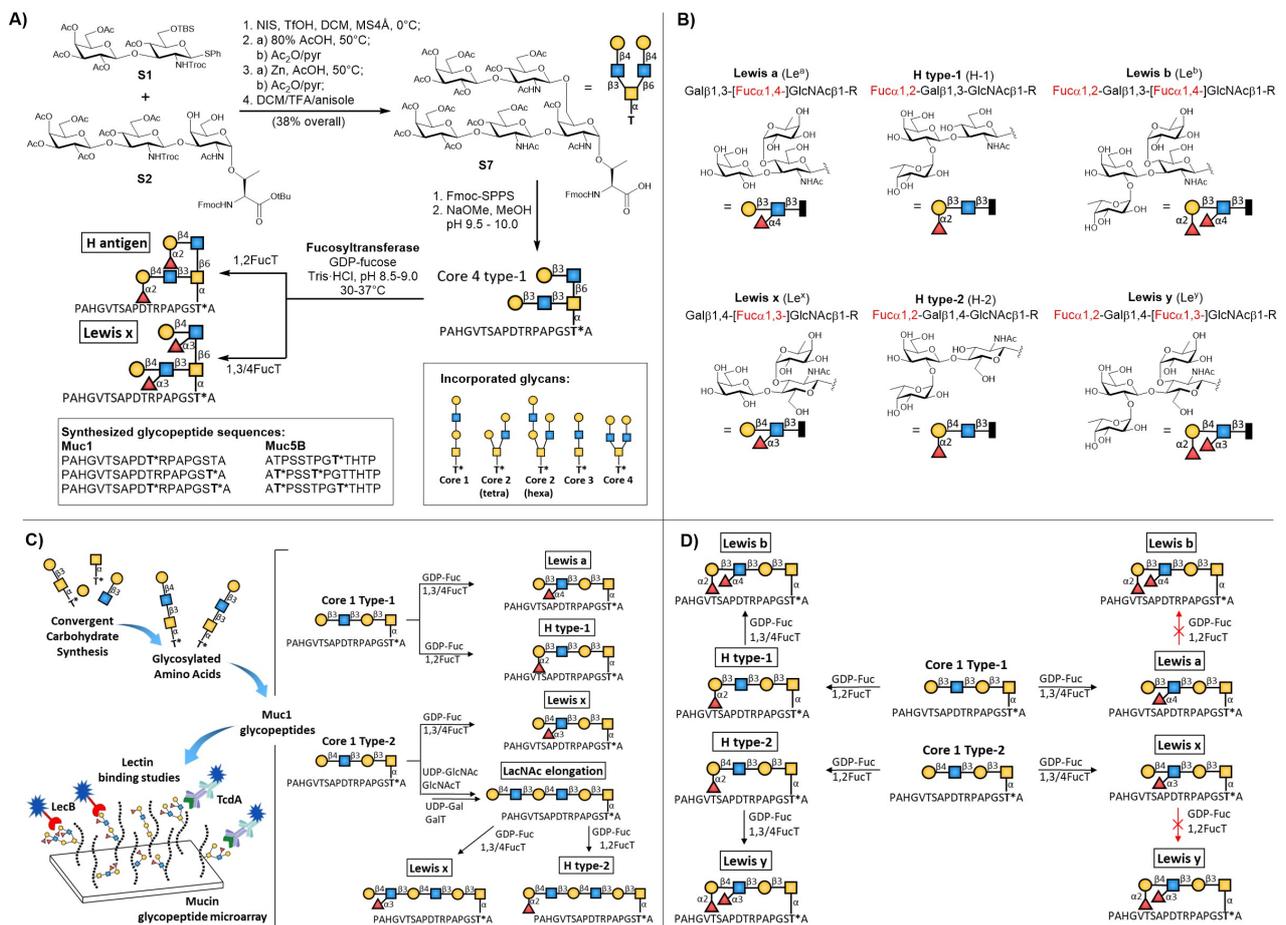
Figure 1. A) Synthesis overview of prepared fucosylated MUC1 and MUC5B *O*-glycopeptides. An example is given for synthesis of the mucin core 4 type-1 glycosylated amino acid building block followed by Fmoc-SPPS glycopeptide synthesis and enzymatic modification. B) Schematic representation of prepared fucosylated terminal glycan antigens. C) Example of enzymatic transformations (fucosylations and additional LacNAc elongation) made for generation of a fucosyl glycopeptide microarray library. D) Synthesis of Lewis b and y modified glycopeptides.

backbone in bacterial lectin recognition events. Therefore, a library of 63 synthetic α-1,2-, α-1,3- and α-1,4-fucosylated mono- and bivalent MUC1 and MUC5B glycopeptides was prepared. The glycopeptides carried differently fucosylated LacNAc core-1 to core-4 structures on distinct mucin peptide tandem repeat glycosylation sites. The obtained fucose glycopeptides presenting different Lewis- and H-antigen structures on the glycan cores were immobilized on NHS-activated microarray slides and applied to elucidate binding interactions of two bacterial lectins, LecB of *P. aeruginosa* and the *C. difficile* toxin A.

Results and Discussion

To study the interactions of the fucose binding lectins LecB and TcdA, LacNAc (type-1 and type-2) elongated mucin core 1-4 glycosylated amino acids were prepared and incorporated into to the human mucin MUC1 and MUC5B peptide tandem repeat sequences, PAHGVTSAPD**T***RPAPGS**T***A and A**T***PSS**T***PG**T***THTP (**T*** = modified glycosylation sites), by Fmoc-solid-phase peptide synthesis (Fmoc-SPPS) (Fig. 1A, Supplementary Fig. S2). The core 1 type-1 and -2, core 2 type-1 and -2 tetrasaccharide and hexasaccharide, core 3 type-1 and -2, and core 4 type-2 threonine building blocks were prepared according to previously reported strategies.[15] The core 4 type-1 building block was synthesized analogously to the reported core 2 type-1 synthesis.[15a] The type-1 core 4

glycosylated amino acid building block **S3** was assembled in a yield of 59% by a [3+2] glycosylation using the LacNAc type-1 disaccharide donor **S1** and the type-1 core 3 glycosylated amino acid acceptor **S2** (Fig. 1A, Supplementary Fig. S1).[16] The *tert*-butyldimethylsilyl protecting group (TBS) was then removed under acidic conditions using 80% acetic acid (AcOH), followed by acetylation with acetic anhydride in pyridine to obtain compound **S4**. The *N*-Troc groups were removed by reductive elimination using zinc dust in AcOH,[17] followed by acetylation to obtain the corresponding acetamide **S5**. Finally, the amino acid *tert*-butyl ester was cleaved using trifluoroacetic acid (TFA)[18] and anisole[19] leading to the formation of the desired type-1 core 4 amino acid **S6**. Starting from the type-1 core 3 acceptor **S2**, the type-1 core 4 glycosylated amino acid **S6** was synthesized in four steps in a yield of 38%. Next, the obtained glycosylated threonine building blocks were introduced into the above mentioned peptide sequences by applying our reported Fmoc-SPPS protocol for glycopeptide synthesis (Fig. 1A, Supplementary Fig. S2).[15, 20] The obtained glycopeptides were enzymatically modified with LacNAc and/or Lewis a (Le^a), Lewis x (Le^x), and/or H type (H) core glycans (Figure 1B, Supplementary Fig. S3-S5). *Helicobacter pylori* β-1,3-*O*-*N*-acetylglucosaminyltransferase (β3GlcNAcT)[21] and a fusion protein of human β-1,4-*O*-galactosyltransferase (His₆-Propeptide-catβ4GalT-1, β4GalT)[22] were applied to enzymatically extend selected glycopeptides with additional LacNAc units followed by additional fucosylation using

Helicobacter pylori α-1,3/4-*O*-fucosyltransferase (Hpα1,3/4FucT) resulting in Le^x and Le^a antigens,[23] or *H. mustelae* α-1,2-*O*-fucosyltransferase (Hmα1,2FucT) resulting in H type-1 or H type-2 antigens[24] (Figure 1C, Supplementary Fig. S3, S4). By combining the two fucosyltransferases on the terminal type-1 or type-2 LacNAc units, bi-fucosylated Lewis b (Le^b) and Lewis y (Le^y) motifs were generated to complete the glycopeptide library. To obtain Le^b and Le^y glycans, the order of the applied fucosyltransferases was crucial for the synthesis outcome (Fig 1D, Supplementary Fig. S6). Selected glycopeptides were in initial attempts α-1,3/4-fucosylated followed by extension with α-1,2-fucose. Using this approach, the desired bi-fucosylated products were not obtained in satisfying yields, since α-1,3/4-fucosylated glycans were found to be poor α-1,2-*O*-fucosyltransferase-substrates. Consequently, the order of fucosylation was reversed and the α-1,2- was followed by the α-1,3/4-fucosylation. Applying this synthesis strategy, the desired Le^b and Le^y mucin glycopeptides were successfully generated. In conclusion, we were able to generate an extensive library of *O*-fucosyl MUC1 and MUC5B glycopeptides with well-defined and closely related glycans that provides an unique platform to explore fine specificities of fucose-binding proteins (Supplementary Table S1).

Next, the glycopeptide library was spotted on NHS-activated hydrogel slides (Nexterion® slide H, Schott). To determine the binding preferences of LecB and TcdA, the glycopeptide microarrays were incubated with dilution series of TcdA (27 nM – 3.5 µM) and LecB-biotin lectins (31 nM – 16 µM), followed by incubation with a mouse anti-TcdA mAb and a Cy5-labeled anti-mouse antibody, and Cy5-labeled streptavidin, respectively, for fluorescent detection. Microarray analysis (see Fig 2-5, Supplementary Fig. S7-S11) showed that the *Pseudomonas aeruginosa* lectin LecB and *Clostridium difficile* Toxin A both exhibited unique fine specificities that strongly depended on the different fucose motifs, presenting peptide backbone, underlying core structures, LacNAc-extension as well as placement of the glycosylation sites. Whereas TcdA selectively recognized α-1,3-fucosylated Muc1 and Muc5B core structures, LecB exhibited a broader selectivity towards all fucosylated

glycopeptides. Calculated apparent K_D values for binding of LecB and TcdA towards the fucosylated MUC1 and MUC5B peptides are reported in Tables 1-3, and supplementary Tables S2 and S3. The microarray IDs of the printed glycopeptides are used henceforth in the discussion.

The *Pseudomonas aeruginosa* lectin LecB bound to all fucosylated MUC1 and MUC5B glycopeptides in a high nanomolar to low micromolar range ($K_{D,Muc1}$ Surf = 0.16 – 2.97 µM and $K_{D,Muc5B}$ Surf = 0.39 – 2.91 µM) (Table 1, Supplementary table S2, Fig. 2, 3, Supplementary Fig. S7-S9). The microarray analysis showed that α-1,4- and type-2 α-1,2-fucosylated MUC1 glycopeptides were better binders than the respective type-1 α-1,2- and α-1,3-fucosylated glycopeptides with Le^x glycans being the weakest binders (Fig. 2A, Fig 3A). This glycopeptide binding pattern was consistent with findings from previous glycan recognition studies of LecB.[9a-c] Here, LecB bound Le^a glycans 1.2-3.1-fold and 3.3-7.0-fold better than the respective H-antigen type and Le^x glycopeptides. Whereas the Le^a antigen can create an additional hydrogen bond to the protein backbone due to the favorable steric location of the GlcNAc *O*-6 position, the GlcNAc *N*-acetyl group of the Le^x glycan located in the same position would lead to sterical hindrance and the Le^x glycan must adapt into a less favorable conformation upon binding to LecB.[9d,e]

In agreement with previous studies, H-antigen type-2 structures were better LecB-binders than the type-1 derivatives.[9a, c] Here, the affinity was 1.3-1.9-fold increased for the H-type-2 glycans with the PDTR core 3 peptides being the only exception (Fig. 2B). Additionally, LecB bound to mucin core structures with different affinities (Fig. 3A): We found that H-antigen peptides glycosylated in the GSTA region were recognized with an increasing affinity order: core 3 < core 1 < core 4 ≈ core 2 tetrasaccharide ≤ core 2 hexasaccharide (Table 1). The α1,2-fucosylated PDTR peptides showed a similar recognition pattern. The linear H-antigen core 1 and 3 glycans showed comparable binding to LecB, while the branched core 2 and 4 structures with two arms carrying one fucose residue each showed a higher
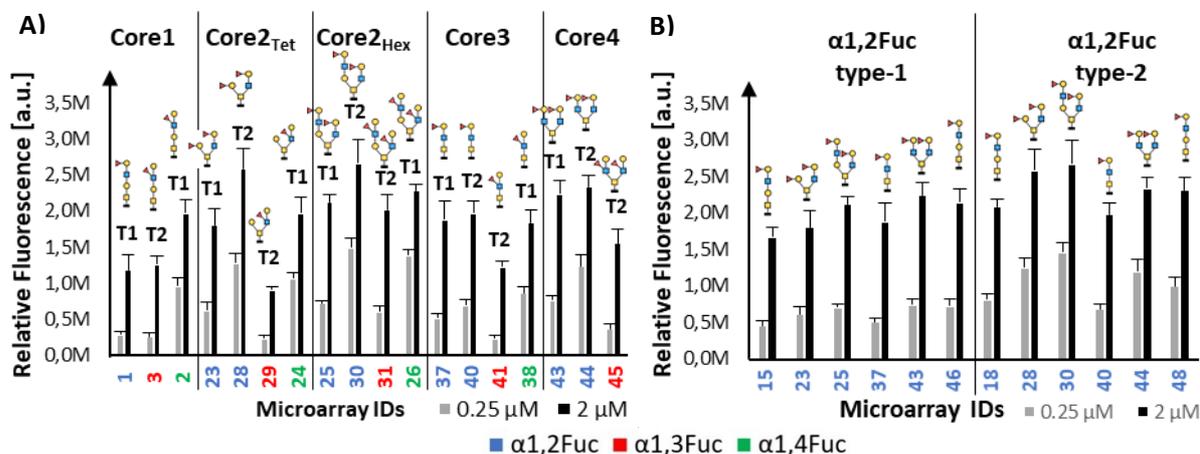
Figure 2. A) Binding of LecB here shown at two selected concentrations at 0.25 µM and 2 µM to different H-type (blue), Le^x (red) and Le^a (green) modified Muc1 core structures; B) Comparison of LecB-binding affinity towards H type-1 and type-2 Muc1 glycopeptides (blue); Notation for carbohydrates see Figure 1.

binding avidity. These terminal epitopes situated on the branched core structures are oriented in opposite directions and are thus presenting a spatial arrangement that may favor intra- or intermolecular multivalent lectin binding.

The Le[a] glycans showed a similar LecB-binding pattern as the corresponding α-1,2-fucosylated glycopeptides with LecB binding to the core structures in the increasing affinity order core 3 < core 1 ≤ core 2 tetrasaccharide < core 2 hexasaccharide (Fig 2A, 3A, Table 1). In contrast to the case in the α-1,2-fucosylated modified glycopeptides, the core 2 tetrasaccharide in Le[a] contains only one fucose residue and therefore shows a similar binding affinity as the core 1 structures. This also explains that the fucosylated core 2 hexasaccharide was a better binder than the respective tetrasaccharide derivative because it contains an additional α-1,4-fucosylated LacNAc unit, which can participate in multivalent interactions leading to an overall higher avidity.

Peptides glycosylated with the Le[x] antigen in the PDTR or GSTA region were recognized by LecB in the following increasing affinity order: core 3 < core 2 tetrasaccharide < core 1 tetrasaccharide < core 4 < core 2 hexasaccharide (Fig. 2A, 3A, table 1). The branched core 2 hexasaccharide and the core 4 glycopeptides contain two fucose residues each, which leads to an enhanced binding compared with the core 2 tetrasaccharide, core 1 and core 3 glycopeptides. The increased affinity of the core 1 glycopeptides compared with the core 2 tetrasaccharide, which both contains only one fucose residue may be explained by less steric hindrance of the bound Le[x] conformation in a linear core 1 structure.

The influence of bivalent glycan presentation on the MUC1 peptide backbone for lectin recognition was further analyzed. Avidity effects with stronger binding of LecB toward bivalent over monovalent MUC1 glycopeptides were observed (Fig. 3A): The bivalent H antigen glycopeptides **46** (K_D = 0.50 µM) and **48** (K_D = 0.35 µM) were better binders than their corresponding monovalent PDTR and GSTA peptides **1**; and **15** and **18**, respectively (Table 1). Again, LecB showed

Table 1. K_D surf values for mono- and bivalent[b] H-type, Lewis x and Lewis a MUC1 glycopeptides determined by incubation of LecB at 8 different concentrations at 31 nM – 16 µM.

	H type			Lewis x			Lewis a	
ID	Core	K_D[µM]	ID	Core	K_D[µM]	ID	Core	K_D[µM]
	PDTR			PDTR				
4	C2T1[Tet]	0.55	3	C1T2	1.19	8	C2T1[Hex]	0.24
7	C2T1[Hex]	0.63	6	C2T2[Tet]	1.72	5	C2T1[Tet]	0.31
9	C3T1	1.02	14	C3T2	2.97	2	C1T1	0.36
1	C1T1	1.10				10	C3T1	0.42
	GSTA			GSTA				
25	C2T1[Hex]	0.53	31	C2T2[Hex]	0.91	26	C2T1[Hex]	0.19
23	C2T1[Tet]	0.54	45	C4T2	1.17	24	C2T1[Tet]	0.25
43	C4T1	0.55	29	C2T2[Tet]	1.72	16	C1T1	0.26
15	C1T1	0.64	41	C3T2	1.76	38	C3T1	0.31
37	C3T1	0.82		Bivalent[b]			Bivalent[b]	
30	C2T2[Hex]	0.28	51	C2T2[Hex]	0.54	47	C1T1	0.16
28	C2T2[Tet]	0.37	53	C4T2	0.75			
44	C4T2	0.37	49	C1T2	0.79			
18	C1T2	0.48	50	C2T2[Tet]	1.12			
40	C3T2	0.60	52	C3T2	1.26			
	Bivalent[b]							
46	C1T1	0.5						
48	C1T2	0.35						

[a] Relative Affinity represents the ratio of binding affinities in comparison to the best binder of the respective glycosylation site.

[b] Bivalent means that two glycosylation sites of the peptide backbone are modified.
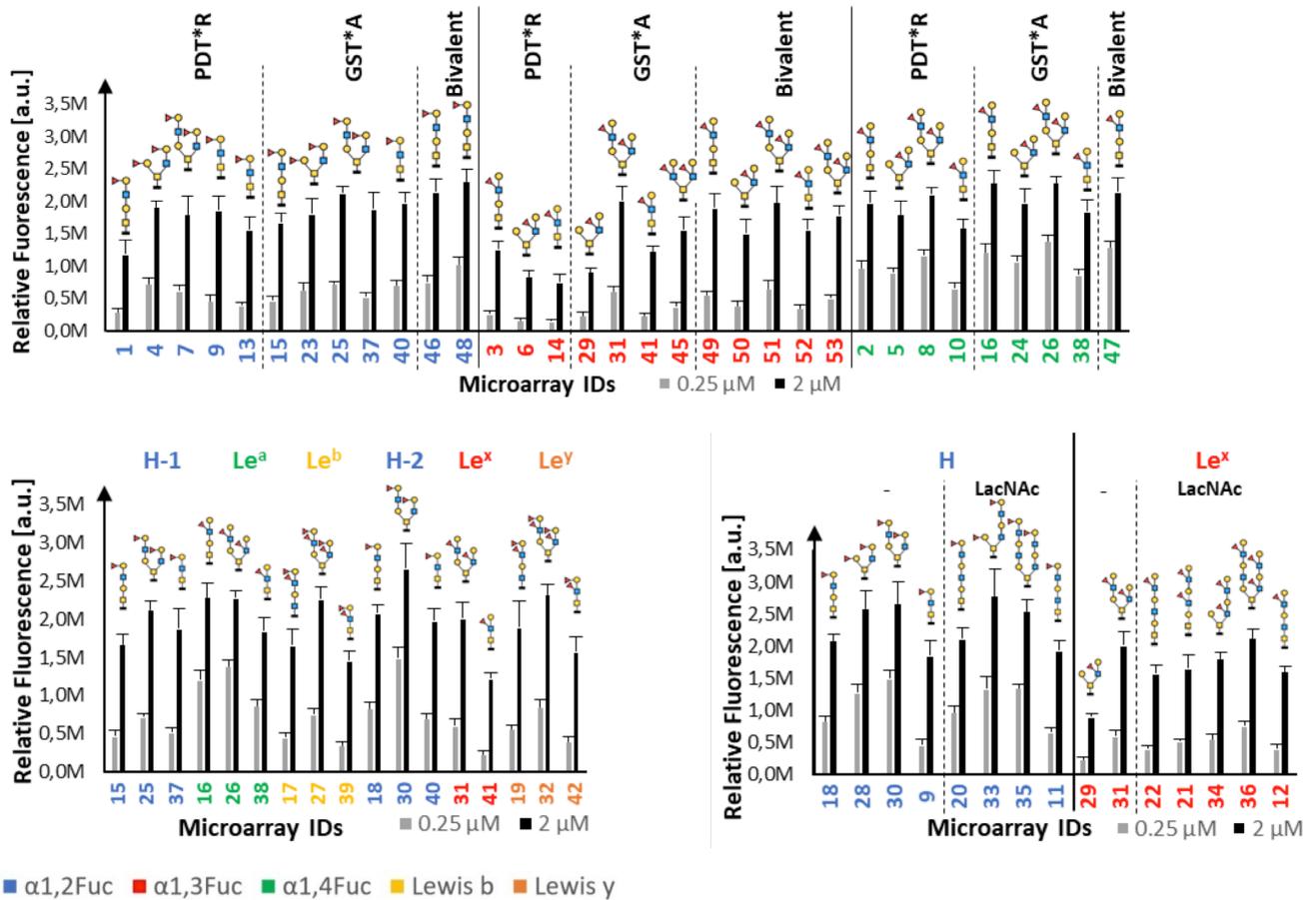
Figure 3. A) Comparison of LecB-binding towards monovalent H-type (blue), Le^x (red) and Le^a (green) Muc1 peptides glycosylated in the PDTR or GSTA region, and the corresponding bivalent glycopeptides; B) Influence of LacNAc elongation on LecB-binding towards H-type and Le^x Muc1 glycopeptides; C) Comparison of LecB-binding between Le^b (yellow) and Le^y (orange); and the respective mono-fucosylated glycopeptides.

preferred binding to the H type-2 glycan over the corresponding type-1 structure. Consistently, the bivalent Le^a peptide **47** (K_D = 0.16 µM) was a better binder than the corresponding monovalent glycopeptides **2** and **16**. Also, the bivalent Le^x glycopeptides **49** (K_D = 0.79 µM), **50** (K_D = 1.12 µM), **51** (K_D = 0.54 µM), **52** (K_D = 1.26 µM) and **53** (K_D = 0.75 µM) showed enhanced binding compared to the respective monovalent glycans (Fig. 3A, Table 1).

Next, the recognition of LecB towards a selection of fucosylated LacNAc elongated MUC1 core structures was evaluated (Fig. 3B). It was previously reported, that the increased length of the oligosaccharide carrying Le^a and Le^x structures can enhance the affinity of LecB.[6b] Consistently, LacNAc elongation on different mucin core structures was here found to increase the LecB-binding up to 3.3-fold (Table 2A). Our results indicated that extension with LacNAc branches of Le^x structures, which may supply one additional fucose residue, allow bidentate or multivalent binding leading to overall stronger LecB binding interactions. In case of the H antigen peptides, increased binding was observed to LacNAc extended core 1 and core 3 glycans, but not for the branched LacNAc core 2 structures **35** and **33**. Our data stand in contrast to a previous study on N-glycans where a di-LacNAc bi-antennary H type-2 glycan structure showed an enhanced

affinity compared to the mono-LacNAc analog due to a favored sterical fit towards the LecB binding pockets.[25] The mucin type O-glycans are structurally different from mannose containing N-glycans and it can be assumed that the spatial orientations of the mucin core branches in relation to the LecB binding pockets promote different binding modes that consequently lead to divergent LecB binding preferences. Additionally, only the terminal galactose residues are α-1,2-fucosylated and therefore, no affinity gain is obtained through increased fucose presentation on the mucin core structures. These findings indicate that extension of the core 1 and core 3 branches leads to a more optimal ligand presentation for the LecB binding sites either by a multivalent or bidentate interaction. In accordance with our findings for the shorter core glycopeptides, LacNAc-modified H-antigen peptides were also better binders than the corresponding Le^x glycopeptides. Also, LecB bound again better to the α1,3-Fuc LacNAc core 2 type-2 hexasaccharide peptide **36**, which carries an additional fucose moiety, than to the respective tetrasaccharide peptide **34**..

Finally, we explored the differences in LecB-binding between Le^b and Le^y core structures with regard to their respective Le^a, Le^x and H type mono-fucosylated analogs. The increased fucose content in the mucin core Le^b

Table 2. A) K_D surf values of LecB-binding towards LacNAc elongated and un-elongated Muc1 glycopeptides; B) K_D surface values of LecB-binding towards of Le^b, Le^y modified glycans, and their respective mono-fucosylated glycopeptides.

A)				B)		
ID	Core Structure	K_D [μM]		ID	Core Structure	
	H- type				Lewis y	
18	C1T2	0.48		18	C1T2+ H	0.48
20	C1T2LacNAc	0.32		19	C1T2+Le^y	0.72
28	C2T2^{Tet}	0.37		30	C2T2^{Hex}+H	0.28
33	C2T2LacNAc	0.34		31	C2T2^{Hex}+Le^x	0.91
30	C2T2^{Hex}	0.28		32	C2T2^{Hex}+Le^y	0.51
35	C2T2^{Hex}LacNAc	0.29		40	C3T2+H	0.60
9	C3T1	1.02		41	C3T2+Le^x	1.76
11	C3T1LacNAc	0.54		42	C3T2+Le^y	1.18
	Lewis x				Lewis b	
22	C1T2LacNAc	0.90		15	C1T1+H	0.64
21	C1T2LacNAc	0.79		16	C1T1+Le^a	0.26
29	C2T2^{Tet}	1.72		17	C1T1+Le^b	0.82
34	C2T2^{Tet}LacNAc	0.76		25	C2T1^{Hex}+H	0.53
31	C2T2^{Hex}	0.91		26	C2T1^{Hex}+Le^a	0.19
36	C2T2^{Hex}LacNAc	0.49		27	C2T1^{Hex}+Le^b	0.54
14	C3T2	2.97		37	C3T1+H	0.82
12	C3T1LacNAc	0.91		33	C3T1+Le^a	0.38

structures could not enhance the LecB binding affinity due to avidity effects (Fig. 3C).[8a, c] Instead, these structures were even weaker binders than both the α1,2-fucosylated H-antigen and Le^a modified glycopeptides (Table 2B) These results indicate that a second fucose residue on the same glycan may sterically hinder the binding interaction of LecB. On the other hand the Le^y core structures were better binders than the respective Le^x glycans, but weaker binders than the H-type glycopeptides. The binding specificities of LecB were further evaluated for selected MUC5B glycopeptides. In accordance with the MUC1 data, the H-type glycopeptides were better binders than the corresponding Le^x glycans (Supplementary Table S2, Fig. S10). Furthermore, we observed that LacNAc elongation on Muc5B glycopeptides increased the binding strength 1.9-fold for H-type glycopeptides, and 1.5-fold for Le^x glycans. In line with the binding data obtained for fucosylated MUC1 glycopeptides, the LecB bound to Le^x core structures in the increasing affinity order: core 3 < core 1 < core 2 hexasaccharide. Unexpectedly and in contrast to the MUC1 data, bivalent glycan presentation on the MUC5B peptide backbone did not have a major impact on LecB-binding. Here, the monovalent peptide **55** was even a slightly better binder than the corresponding bivalent peptides **60** and **63**. In conclusion, these data suggest that the presentation of the ligand(s) on

the peptide backbone as well as on distinct glycosylation sites is important for LecB binding, in particular to gain optimal effects from bidentate and multivalent binding interactions.

A second bacterial lectin was evaluated, the *C. difficile* toxin A (TcdA). This lectin was in previous reports reported to recognize the Galili epitope, but also the fucosylated Le^y, Le^x, Silayl-Le^x recognition to our fucosylated mucin core glycopeptide library showed complete selectivity towards α1,3-fucosylated glycopeptides. Le^x and Le^y modified glycopeptides bound to TcdA in a high nanomolar to low micromolar range (Surf. K_D = 0.28 - 2.46μM), whereas H-type, Le^a and Le^b modified glycopeptides were not recognized at all (Supplementary Fig. S10-S11). Analogously to LecB, we evaluated the binding differences towards MUC1 and MUC5B glycopeptides with regard to different fucosylation types, core structures and glycosylation sites. We found that TcdA showed differences in binding strength towards different MUC1 core structures: Le^x peptides glycosylated in the PDTR and GSTA regions were recognized in the increasing affinity order: C2T2Tet < C1T2 < C3T2, and C2T2Tet < C4T2 < C2T2Hex ≤ C3T2, (Fig. 4A, 4B) indicating that branched structures, although presenting two Le^x units instead of one, are less favored than the linear core 1 and core 3 glycans.

Additionally, TcdA preferably bound to MUC1 peptides glycosylated in the GSTA over the PDTR region (Fig. 4B). We also found that bivalent peptides occupying two glycosylation sites were better TcdA binders than the respective monovalent glycopeptides. These findings can again be related to the multivalent binding effects.

Interestingly, LacNAc elongation decreased binding of TcdA towards MUC1 peptides (Fig. 4C). For example, the shorter core glycopeptides **14**, **29** and **31** were better binders than their LacNAc elongated analogues **12**, **34** and **36**. Further, TcdA bound to Le^y structures **32** and **42** with a similar strength as to the Le^x glycopeptides **31** and **41** (Fig. 4D). These findings indicate that the shorter glycan structures might have a better sterical fit to the TcdA binding sites and that their higher rigidity might lead to a beneficial entropy effect.

Finally, we evaluated TcdA regarding its binding specificities towards fucosylated MUC5B glycopeptides. Again, TcdA did not recognize H-type MUC5B glycopeptides and only bound to the α-1,3-fucosylated Le^x glycans (Supplementary Table S3, Fig. S12, Fig S13). In agreement with the MUC1 glycopeptides, LacNAc elongation decreased the TcdA binding affinity compared with the respective shorter core glycopeptides. The bivalent Le^x peptides **60** and **63** were better binders than the monovalent glycopeptide **55**. The placement of the second glycosylation site was again not important for lectin binding. In contrast to the MUC1 data, core specificity of for Le^x bivalent glycopeptides was observed in the increasing affinity order: C1T2 < C3T2 < C2T2Hex.
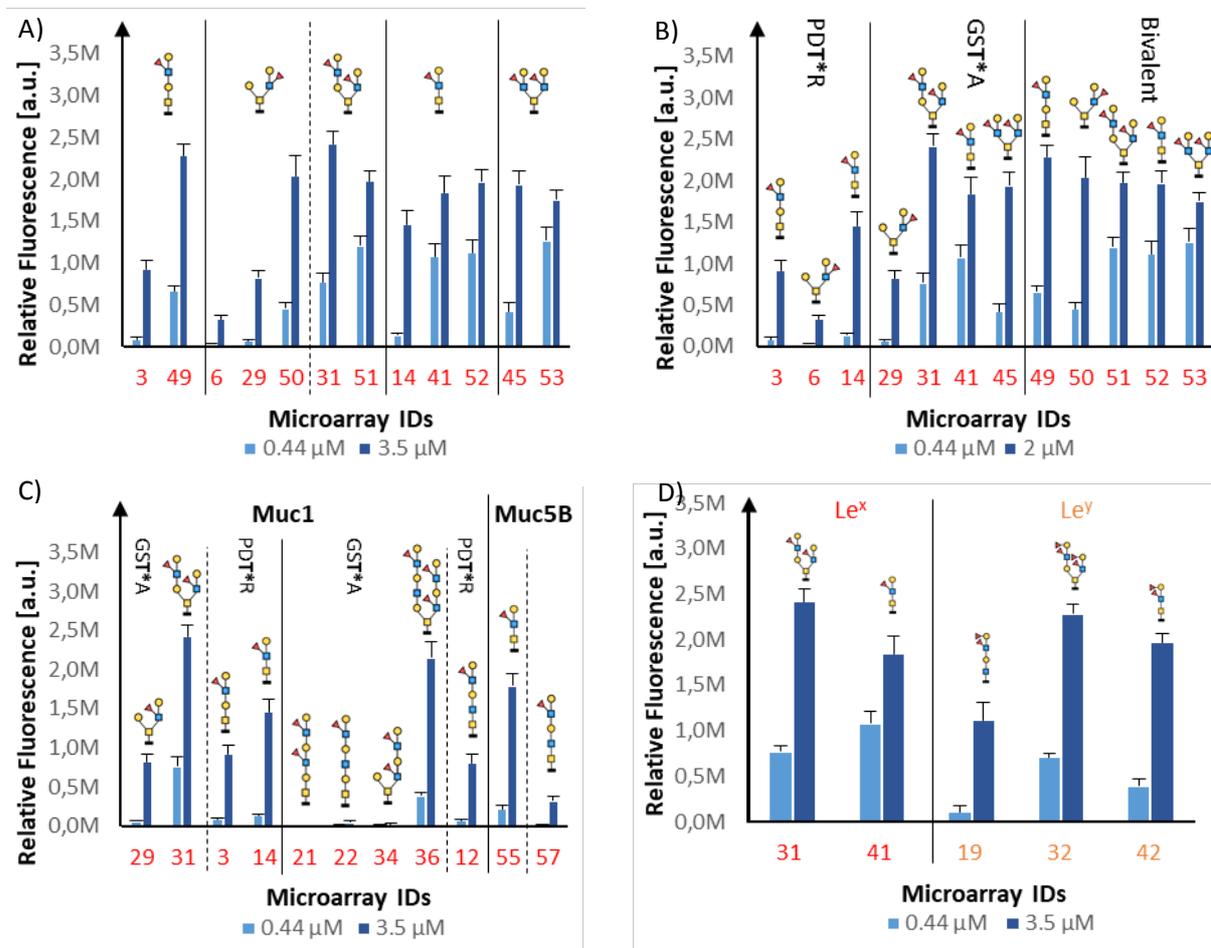
Figure 4. A) TcdA-binding towards different Le^x (red) MUC1 core structures; B) Comparison of TcdA-binding towards monovalent Le^x MUC1 peptides glycosylated in the PDTR or GSTA region, and the corresponding bivalent glycopeptides; C) Influence of LacNAc elongation on TcdA-binding towards Le^x Muc1 glycopeptides; D) Comparison of TcdA-binding between and Le^y (orange) glycans; and the corresponding Le^x glycopeptides.

Conclusion

In summary, a library of fucosylated mucin core 1-4 MUC1 and MUC5B tandem repeat glycopeptides was generated to study the fine binding specificities of the fucose-recognizing bacterial lectins LecB from *Pseudomonas aeruginosa* and toxin A from *Clostridium difficile.* Selected glycopeptides were extended with additional LacNAc units using the Helicobacter pylori β-1,3-*O*-*N*-acetylglucosaminyltransferase (β3GlcNAcT) and a fusion protein of human β-1,4-*O*-galactosyltransferase (His$_6$-Propeptide-catβ4GalT-1, β4GalT). Subsequently, the different fucose motifs, including the Le^a, Le^x and H-type as well as bi-fucosylated Le^b and Le^y antigens, were enzymatically prepared using *Helicobacter pylori* α-1,3/4-*O*-fucosyltransferase (Hpα1,3/4FucT) and/or *H. mustelae* α-1,2-*O*-fucosyltransferase (Hmα1,2FucT). Thereby, the order of the applied fucosyltransferases was crucial to prepare the Le^b and Le^y determinants. The obtained fucosylated mucin glycopeptide library was printed on NHS-activated microarrays, which was applied to determine the binding preferences of LecB and TcdA. Whereas TcdA exclusively bound to α-1,3-fucosylated Muc1 and Muc5B core structures consisting of Le^x and Le^y epitopes, LecB exhibited a broader selectivity toward all presented fucosylated glycopeptides. Additionally, both lectins exhibited unique fine specificities that strongly depended on the different fucose motifs, presenting peptide backbone, underlying core structures, LacNAc-extension as well as placement of the glycosylation sites on the MUC1 and MUC5B glycopeptides. These findings highlight the importance of the evaluated structural glycopeptide properties in lectin binding interactions, which defines the glycan orientation, structural rigidity or possible limitations for ligand recognition.

[1]   a) G. C. Hansson, *Annu. Rev. Biochem* **2020** *89,* 769-793; b) J. F. Sicard, G. Le Bihan, P. Vogeleer, M. Jacques, and J, Harel, *Chem. Commun.* **1990**, 270-272; c) G. H. Veeneman, S. H. van Leeuwen and J. H. van Boom, *Tet. Lett.* **1990**, *31*, 1331-1334.

[2]   a) J. Poole, C. J. Day, M. von Itzstein, J. C. Paton, and M. P. Jennings, *Nat. Rev. Microbiol.* **2018**, *16*, 440-452; b) Y. Watanabe, T. A. Bowden, I. A. Wilson and M. Crispin, *Biochim. Biophys. Acta Gen. Subj.* **2019**, *1863*, 1480-1497.

[3]   a) L. Lindesmith, C. Moe, S. Marionneau, N. Ruvoen, X. Jiang, L. Lindblad, P. Stewart, J. LePendu and R. Baric, Nat. Med. **2003**, *9*, 548-553; b) S. A. Berger, N. A. Young, S. C. Edberg, Eur. J. Clin. Microbiol. Infect. Dis*.* **1989,** *8,* 681-689; c) J. E. Heggelund, E. Haugen, B. Lygren, A. Mackenzie, S. Holmner, F. Vasile, J. J. Reina, A. Bernardi and U. Krengel, *Biochem. Biophys. Res. Commun.* **2012,** *418,* 731-735; d) J. B. Heim, V. Hodnik, J. E. Heggelund, G. Anderluh and U. Krengel. *Scientific Reports* **2019,** *9,* 12243-12256; e) A. M. Wands, J. Cervin, H. Huang, Y. Zhang, G. Youn, C. A. Brautigam, M. Matson Dzebo, P. Björklund, V. Wallenius, D. K. Bright, C. S. Bennett, P. Wittung-Stafshede, N. S. Sampson, U. Yrlid, and J. J. Kohler, *ACS Infect Dis.* **2018,** *4,* 758-770.

[4]   G. Lamblin, S. Degroote, J. M. Perini, P. Delmotte, A. Scharfman, M. Davril, J.-M. Guidice, N. Houdret, V. Dumur, A. Klein and P. Roussel *Glycoconj. J.* **2001**, *18*, 661-684.

[5]   a) E. P. Mitchell, C. Houles, D. Sudakevitz, M. Wimmerova, C. Gautier, S. Perez, A. M. Wu, N. Gilboa-Garber and A. Imberty, *Nature Struct. Biol.* **2002** *9,* 918-921; b) E. P. Mitchell, C. Sabin, L. Šnajdrová, M. Pokorná, S. Perret, C. Gautier, C. Hofr, N. Gilboa-Garber, J Koca,M. Wimmerova and A. Imberty, *Proteins: Struct. Funct. Bioinfo.* **2005**, *58,* 735-748.

[6]   a) K. D. Tucker and T. D. Wilkins, *Infect. Immun.* **1991**, *59*, 73-78; b) V. Heine, S. Boesveld, H. Pelantová, V. Kren, C. Trautwein, P.Strnad and L. Elling, *Bioconj. Chem.* **2019**, *30,* 2373-2383; c) C.-Y. Yeh, C.-N. Lin, C.-F. Chang, C.-H. Lin, H.-T. Lien, J.-Y. Chen, and J.-S. Chia, *Infect. Immun.* **2008**, *76*, 1170-1178.

[7]   D. Tielker, S. Hacker, R. Loris, M. Strathmann, J. Wingender, S. Wilhelm, F. Rosenau, and K.-E. Jaeger, *Microbiology* **2005**, *151*, 1313-1323.

[8]   a) K. Marotte, C. Sabin, C. Préville, M. Moumé-Pymbock, M. Wimmerová, E. P. Mitchell, A. Imberty and R. Roy, *ChemMedChem* **2007**, *2*, 1328-1338; b) A. M. Wu, J. H. Wu, T. Singh, J. H. Liu, M. S. Tsai and N. Gilboa-Garber, *Biochimie* **2006**, *88*, 1479-1492; c) A. M. Boukerb, A. Decor, S. Ribun, R. Tabaroni, A. Rousset, L. Commin, S. Buff, A. Doléans-Jordheim, S. Vidal, A. Varrot, A. Imberty and B. Cournoyer, *Frontiers in Microbiology* **2016**, *7*, 811; d) S. Perret, C. Sabin, C. Dumon, M. Pokorná, C. Gautier, O. Galanina, S. Ilia, N. Bovin, M. Nicaise, M. Desmadril, N. Gilboa-Garber, M. Wimmerová, E. P. Mitchell and A. Imberty, *The Biochemical journal* **2005**, *389*, 325-332; e) E. Mitchell, C. Houles, D. Sudakevitz, M. Wimmerova, C. Gautier, S. Pérez, A. M. Wu, N. Gilboa-Garber and A. Imberty, *Nature Structural Biology* **2002**, *9*, 918-921.

[9]   a) R. J. Pieters, *Org. Biomol. Chem*. **2009**, *7*, 2013-2025; b) P. Bojarová, R. R. Rosencrantz, L. Elling and V. Kren, *Chem. Soc. Rev.* **2013,** *42,* 4774-4797; c) A. Bernardi, J. Jimenez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K.-E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penade ́s, F. Peri, R. J. Pieters, O. Renaudet, J.-L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schaffer, B. Turnbull, T. Velasco-Torrijos, S. Vidal,, S. Vincent, T. Wennekes, H. Zuilhof and A. Imberty, *Chem. Soc. Rev.* **2013,** *42,* 4709-4727; d) M. C. Galan, P. Dumy, O. Renaudet, *Chem. Soc. Rev.* **2013,** *42,* 4599-4612 e) V. Wittman, *Curr. Opin. Chem. Biol.* **2013,** *6, 982*-989; f) S. Behren, U. Westerlind, *Molecules* **2019,** *24, 1004-1035.*

[10]   a) E. M. Johansson, S. A. Crusz, E. Kolomiets, L. Buts, R. U. Kadam, M. Cacciarini, K.-M. Bartels, S. P. Diggle, M. Cámara, P. Williams, R. Loris, C. Nativi, F. Rosenau, K.-E. Jaeger, T. Darbre and J.-L. Reymond *Chem. Biol.* **2008,** *15*, 1249-1257; b) G. Michaud, R. Visini, M. Bergmann, G. Salerno, R. Bosco, E. Gillon, B. Richichi, C. Nativi, A. Imberty, A. Stocker, T. Darbre and J.-L. Reymond *Chem. Sci.* **2016,** *7*, 166-182; c) K. S. Bücher, N. Babic, T. Freichel, F. Kovacic and L. Hartmann, *Macromol Biosci* **2018**, *18*, 1800337; d) N. Berthet, B. Thomas, I. Bossu, E. Dufour, E. Gillon, J. Garcia, N. Spinelli, A. Imberty, P. Dumy, and O. Renaudet, *Bioconjugate Chem.* **2013,** *24*, 1598-1611.

[11]   a) R. Sommer, K. Rox, S. Wagner, D. Hauck, S. S. Henrikus, S. Newsad, T. Arnold, T. Ryckmans, M. Brönstrup, A. Imberty, A. Varrot, R. W. Hartmann and A. Titz, *J. Med. Chem.* **2019**, *62*, 9201-9216; b) 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 and A. Titz, *J. Am. Chem. Soc.* **2018**, *140,* 2537−2545.

[12]   C. M. Surawicz, and J. Alexander, *Nat. Rev. Gastroenterol. Hepatol.* **2011,** *8,* 330-339.

[13]   a) R. N. Pruitt and D. B. Lacy, *Frontiers in Cellular and Infection Microbiology* **2012,** *2*, 28; b) K. Aktories, *Nat. Rev. Micro.* **2011,** *9*, 487-498.

[14]   R. M. Cherian, C. Jin, J. Liu, N. G. Karlsson, and J. Holgersson, *Infect Immun* **2016**, *84,* 2842-2852.

[15]   a) C. Pett, M. Schorlemer and U. Westerlind, *Chemistry – A European Journal* **2013**, *19*, 17001-17010; b) C. Pett and U. Westerlind, *Chemistry – A European Journal* **2014**, *20*, 7287-7299.

[16]   a) P. Konradsson, U. E. Udodong and B. Fraser-Reid, *Tetrahedron Letters* **1990**, *31*, 4313-4316; b) P. Konradsson, D. R. Mootoo, R. E. McDevitt and B. Fraser-Reid, *Front. Cell. Infect. Microbiol.* **2017**, *7,* 387.

[17]   G. Just and K. Grozinger, *Synthesis* **1976**, *1976*, 457-458.

[18]   D. B. Bryan, R. F. Hall, K. G. Holden, W. F. Huffman and J. G. Gleason, *Journal of the American Chemical Society* **1977**, *99*, 2353-2355.

[19]   S. F. Martin, K. X. Chen and C. T. Eary, *Organic Letters* **1999**, *1*, 79-82.

[20]   a) C. Pett, H. Cai, J. Liu, B. Palitzsch, M. Schorlemer, S. Hartmann, N. Stergiou, M. Lu, H. Kunz, E. Schmitt, U. Westerlind, *Chem. Eur. J.* **2017**, 23, 3875; b) C. Pett, W. Nasir, C. Sihlbom, B. M. Olsson, V. Caixeta, M. Schorlemer, R. P. Zahedi, G. Larsson, J. Nilsson, and U. Westerlind, *Angew. Chem. Int. Ed.* **2018,** *57*, 9320-9324.

[21]   C. Rech, R.R. Rosencrantz, K. Krenek, H. Pelantová, P. Bojarová, C.E. Römer, F.-G. Hanisch, V. Kren, L. Elling, *Adv. Synth. Catal.* **2011,** *353,* 2492–2500

[22]   B. Sauerzapfe, D.J. Namdjou, T. Schumacher, N. Linden, K. Krenek, V. Kren, L. Elling, *J. Mol. Catal. B Enzym.* **2008,** *50,* 128–140.

[23]   D.A. Rasko, G. Wang, M. M. Palcic, and D. E. Taylor, *J. Biol. Chem.* **2000,** *275,* 4988–4994.

[24]   J. Ye, X. W. Liu, P. Peng, W. Yi and H. Cao, *ACS Catal.* **2016,** *6,* 8140–8144.

[25]   R. Sommer, S. Wagner, A. Varrot, C. M. Nycholat, A. Khaledi, S. Häussler, J. C. Paulson, A. Imberty and A. Titz, *Chemical science* **2016**, *7*, 4990-5001.