Evaluation of the carbohydrate recognition domain of the bacterial adhesin FimH: design, synthesis and binding properties of mannoside ligands

Oliver Sperling,† Andreas Fuchs† and Thisbe K. Lindhorst*†

Received 26th July 2006, Accepted 5th September 2006
First published as an Advance Article on the web 25th September 2006
DOI: 10.1039/b610745a

Fimbriae are proteinogeneous appendages on the surface of bacteria, which mediate bacterial adhesion to the host cell glycocalyx. The so-called type 1 fimbriae exhibit specificity for α-D-mannosides and, therefore, they are assumed to mediate bacterial adhesion *via* the interaction of a fimbrial lectin and α-D-mannosyl residues exposed on the host cell surface. This carbohydrate-specific adhesive protein subunit of type 1 fimbriae has been identified as a protein called FimH. The crystal structure of this lectin is known and, based on this information, the molecular details of the interaction of mannoside ligands and FimH are addressed in this paper. Computer-based docking methods were used to evaluate known ligands as well as to design new ones. Then, a series of new mannosides with extended aglycon was synthesized and tested as inhibitors of type 1 fimbriae-mediated bacterial adhesion in an ELISA. The results obtained were compared to the predictions and findings as delivered by molecular modeling. This study led to an improved understanding of the ligand–receptor interactions under investigation.

Introduction

Bacterial colonization of organs and cells respectively can cause severe problems for an organism. When *Escherichia coli* bacteria, which are commensal residents of the intestine, enter the urogenital tract and multiply in the bladder, for example, the body reacts with inflammation, such as in the case of cystitis.[1] Colonization of the stomach by *Helicobacter pylori* causes inflammation in the stomach (gastritis) as well as ulceration of the stomach or peptic ulcer disease.[2]

In order to understand the molecular circumstances of such processes, it is reasonable to begin with the investigation of the molecular interactions between bacteria and their host cells. These involve adhesion of bacteria to the glycocalyx, a complex nano-dimensioned layer of diverse glycoconjugates, which covers every eukaryotic cell. Bacteria enable adhesion to the glycocalyx by adhesive organells called fimbriae or pili.[3] Fimbriae are classified according to their carbohydrate specificity. The so-called type 1 fimbriae express a specificity for terminal α-mannosyl residues, which is mediated by a lectin called FimH forming a minor component of the fimbrial protein complex.[4] According to many studies, type 1 fimbriae are an important and critical factor for bacterial virulence, such as for uropathogenic *E. coli*.[5]

Since more than three decades, researchers have investigated various mannosides and oligosaccharides as inhibitors for type 1 fimbriae-mediated bacterial adhesion.[6] The purpose of this research has been twofold: (i) this work has been carried out to learn more about the molecular details of bacterial adhesion to the host cell glycocalyx and (ii) there has been a hope to develop carbohydrate-based antiadhesives, which might eventually be used as therapeutics against bacterial colonization and infection.[7]

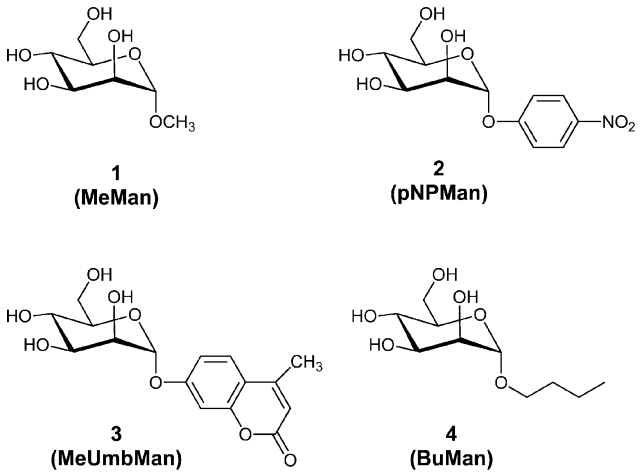
When type 1 fimbriae-mediated binding to various mannosides and mannose-containing ligands is tested in an ELISA (enzyme-linked immunosorbent assay), weak interactions are observed with IC_{50} values in the millimolar and low micromolar range. Attempts to improve mannose ligands as inhibitors of type 1 fimbriae-mediated bacterial adhesion followed two different approaches: (i) multivalent carbohydrate ligands have been designed (*cf.* preceding paper[8]) to utilize known multivalency effects,[9] and (ii) target design of ligands has been employed to optimize the interactions of a carbohydrate ligand to the carbohydrate recognition domain (CRD) of the bacterial adhesin.[10]

Rational design of carbohydrate ligands for FimH can been guided by information obtained from the crystal structure of the bacterial adhesin. The first X-ray structure of FimH with a ligand bound to its carbohydrate recognition domain (CRD), in complex with the chaperone FimC, was published in 1999,[11] and was followed by two other X-ray studies, which appeared in 2002[12] and 2005, respectively.[13] We have employed this structural information on FimH for our study on the design of carbohydrate ligands for FimH, which is reported here, together with their synthesis and biological testing.

Results and discussion

Well-known ligands for FimH are depicted in Scheme 1. The binding potency of methyl α-D-mannoside (**1**, MeMan) to FimH lies in the millimolar range as reflected by inhibition of hemagglutination or by ELISA, respectively.[14] On the other hand, α-D-mannosides carrying an aromatic aglycon such as *p*-nitrophenyl α-D-mannoside (**2**, *p*NPMan) and methylumbelliferyl α-D-mannoside (**3**, MeUmbMan), show a potency as inhibitors of type 1 fimbriae-mediated bacterial adhesion, which is increased by

Otto Diels Institute of Organic Chemistry, Christiana Albertina University of Kiel, Otto-Hahn-Platz 4, 24098, Kiel, Germany. E-mail: tklind@oc.uni-kiel.de; Fax: (+49)431 880 7410
† These authors have contributed equally.



Scheme 1 Standard mannoside ligands for the bacterial lectin FimH.

two orders of magnitude when compared to MeMan.[6] This can be understood on the basis of the crystal structure of the adhesive protein FimH. All three available X-ray analyses reveal a single carbohydrate recognition domain (CRD) on the tip of the FimH protein, with the size of a monosaccharide.

This CRD perfectly accommodates an α-D-mannosyl residue. It is mainly comprised of amino acids with hydrophilic side chains, including the *N*-terminal amino acid of the protein, PHE1 (Fig. 1a). These amino acid side chains and, in particular, the central aspartic acid ASP54 support complexation of the six-membered ring of α-D-mannosides within the CRD with the α-glycosidic linkage pointing outwards of the binding site.

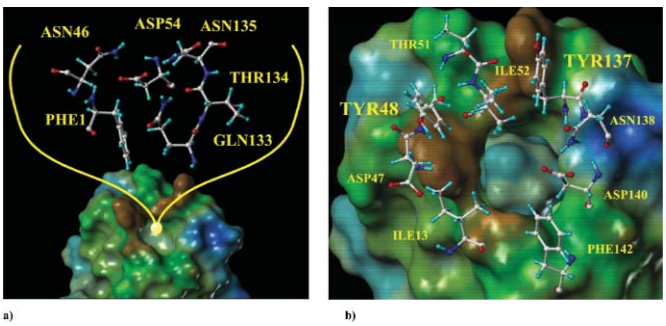


Fig. 1 Details of the carbohydrate recognition domain (CRD) of FimH as revealed by X-ray analysis.[12] Mannose is depicted in the binding pocket as surface representation. a) The interior of the carbohydrate binding pocket of FimH is mainly comprised of amino acids with hydrophilic side chains, including PHE1, the N-terminal amino acid of the protein. Especially, the central aspartic acid forms strong interactions with the complexed mannose or mannoside, respectively. b) The exterior of the FimH CRD is characterized by amino acids with hydrophobic side chains, defining a so-called "hydrophobic ridge". The amino acids TYR48 and TYR137 flank the entrance of the binding pocket forming a gate which has been called the "tyrosine gate".[11,12]

The exterior of the CRD on the other hand is characterized by amino acids with rather lipophilic side chains, defining a so-called 'hydrophobic ridge' at the entrance of the CRD[11,12] (Fig. 1b). Two aromatic amino acid side chains of this protein environment, the phenol moieties of TYR48 and TYR137, define a molecular entrance to the CRD, which has been named the 'tyrosine gate'.[11,12]

The aromatic aglycon of an appropriate mannoside ligand can establish favorable π–π-interactions with this tyrosine gate, thus leading to significantly improved affinities of mannosides such as *p*NPMan (**2**) and MeUmbMan (**3**) in comparison to MeMan (**1**).

Based on this knowledge, we anticipated that mannosides with an extended aromatic aglycon could satisfy further interactions with the CRD exterior, thus improving binding to FimH. To allow extension of the aglycon moiety of an α-D-mannoside according to a general strategy, we have employed squaric acid diester (diethylsquarate, DES)[15] as a bifunctional linker,[16] and have tested it with 2-aminoethyl α-D-mannoside (**5**, Scheme 2) first. Reaction of the amine **5**[17] with DES in methanol gave literature-known **6**[18] in a smooth reaction, and eventually **9** could be obtained starting from *p*-aminophenyl α-D-mannoside (**7**)[19] in analogy.

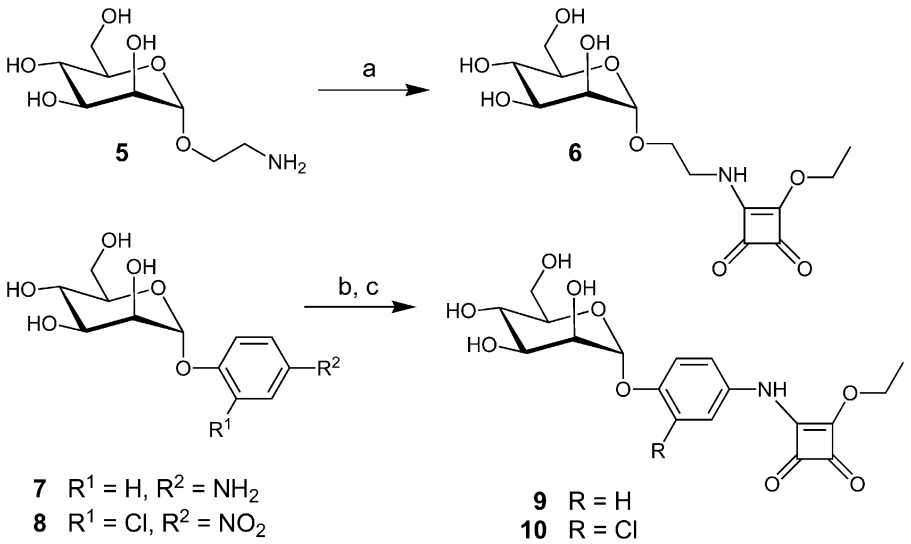
When tested in an ELISA, the potency of mannoside **6** as an inhibitor of type 1 fimbriae-mediated bacterial adhesion was poor as expected, whereas the inhibitory potency of the squaric acid monoester **9** was increased 1800-fold when compared to methyl α-D-mannoside (MeMan, **1**) and some 60 times higher than that of *p*-nitrophenyl α-D-mannoside (*p*NPMan, **2**) in the same assay (Table 1). Therefore, we concluded that the extended aglycon moiety of **9** can establish additional, favorable interactions at the entrance of the FimH CRD.

Consequently, we have employed computer-aided modeling to rationalize the measured inihibitory potencies of the various FimH ligands. We have used FlexX[20] flexible docking and consensus scoring[21] as implemented in Sybyl6.8[22] for docking of known

Table 1 FlexX scoring values as obtained by ligand docking are compared to the potencies of the different mannosides as inhibitors of the mannose-specific adhesion of *E. coli* to mannan as determined by ELISA. Lower scoring values correspond to higher affinities to the protein FimH. Relative inhibitory potencies (RIP) are relative to the IC_{50} value measured for methyl α-D-mannoside (**1**); thus the inhibitory potency of **1** has been defined as RIP ≡ 1. The listed RIP is an average value of the results of at least three independently performed ELISAs and therefore reported with its standard deviation

Mannosides investigated	Score[a]	RIP[b]	s.d.[c]
1 (MeMan)	−22.5	1	—
2 (*p*NPMan)	−24.9	31	13
3 (MeUmbMan)	−20.2	380	240
4 (BuMan)	−20.5	5.7	0.32
6	−26.5	2.7	0.43
8	−25.4	200	44
9	−29.2	1800	920
10	−26.9	6900	2300
11	−26.6	210	100
12	−20.7[d]	340	150
13	−22.0[e]	430	210
14	−24.3[f]	330	46

[a] If not otherwise indexed, values for top scorers are listed (if the respective conformation is reasonable, *i.e.* the mannosyl aglycon of the ligand is placed correctly inside the CRD). [b] RIP = relative inhibitory potency. [c] s.d. = standard deviation. [d] The highest seeded score with a reasonable conformation out of 90 docked solutions, which were considered for scoring, was hit number 12. [e] The highest seeded score with a reasonable conformation out of 90 docked solutions, which were considered for scoring, was hit number 60. [f] The highest seeded score with a reasonable conformation out of 90 docked solutions, which were considered for scoring, was hit number 67.



Scheme 2 Synthesis of the squaric acid ethylesters **6**, **9**, and **10**. Reagents and conditions: a) DES, MeOH, 10 h, 71%; b) from **7**: DES, DMF, 63%; c) from **8**: 1. Pd/C, H_2, DMF. 2. DES, DMF, 55% over two steps.

and potential carbohydrate ligands into the CRD of FimH and for scoring of the obtained ligand conformations. Docking was based on the published X-ray structure with D-mannose bound in the CRD.[12] This FimH CRD was held fixed during the minimization, whereas the sugar ligand was allowed to change its conformation freely under the influence of the force field. The ligand conformations delivered by FlexX as docking solutions are regarded as "unrelaxed". To release artificial strains, which might arise during the docking process, such conformations can be further proceeded in an energy minimization to deliver so-called "relaxed" conformations, which might differ from the unrelaxed solutions.[23]

Either unrelaxed or relaxed (or both) solutions of the docking have to be submitted to a scoring process to identify the most reasonable results. In this process, FlexX produces so-called scoring values for each docked ligand, which can be regarded as a rough estimate of its free binding energy. Low (more negative respectively) scoring values correlate with high affinities, higher scores reflect diminished binding potency. To validate FlexX scoring, in addition to the FlexX scoring function, a number of other scoring functions based on different algorithms were used for scoring of docking results and the scoring results obtained were evaluated according to the consensus scoring strategy.[24] This procedure allows a most reliable ranking of the docked solutions.

In the case of MeMan (**1**), docking with FlexX reproduces both conformation and orientation of the sugar ring in the CRD of FimH as observed by X-ray analysis of this lectin–sugar complex.[12] No difference between the unrelaxed and respective relaxed docking solutions was found in this case and all following ones. The mannosyl glycon is buried in the binding pocket and the α-configured methyl aglycon points outwards of the pocket. In the case of **2** (*p*NPMan), the mannosyl residue maintains its perfect fit within the CRD and an ideal geometry of the α-positioned aglycon allows optimal interactions of the aromatic *p*-nitrophenyl aglycon with the tyrosine gate of the CRD formed by TYR48 and TYR137 (Fig. 2). These lipophilic interactions of the *p*-nitrophenyl ring with the tyrosine gate lead to higher affinity to the receptor and this is

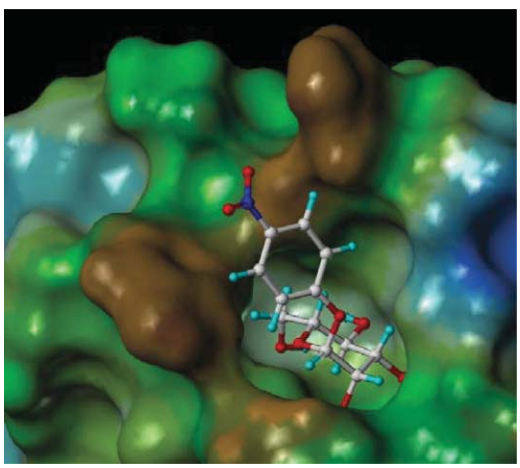


Fig. 2 Fit of top scoring conformation of *p*NPMan (**2**) depicted in the FimH CRD. The glycon moiety of the mannoside is buried in the rather deep binding pocket, whereas the phenyl aglycon sticks out of this pocket interacting with the hydrophobic tyrosine gate at the entrance of the CRD.

reflected in a scoring value of −24.9 which is significantly lower than that of MeMan (−22.5, Table 1). This finding corresponds to the measured inhibitiory potency of *p*NPMan, which is 31-times higher (RIP = 31) than that of MeMan (RIP = 1).

In the case of the new squaric acid derivative **9**, docking revealed additional interactions of the extended aglycon moiety with the FimH CRD when compared to *p*NPMan (**2**). Orientation of the phenyl ring in **9** is as in **2** however, the squaric acid sub-structure with its planar geometry forms further interactions with the tyrosine gate (Fig. 3a). This moiety is also able to interact with the THR51 hydroxyl group on the distal end of the gate by hydrogen-bridging, and this is reflected by a diminished scoring value of −29.2 (Table 1). The affinity to FimH is reduced dramatically when the phenyl ring in **9** is replaced by an ethyl spacer such as in the case of **6** (score −26.5, RIP 2.7, Table 1). Visual inspection of the docking results for **6** reveals that the ethyl spacer is too short to

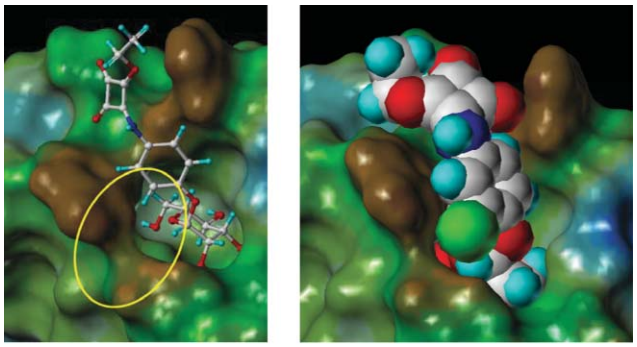


Fig. 3 Fit of top scoring conformations of designed mannoside ligands **9** and **10** in the CRD of FimH. a) Top scoring conformation of the squaric acid ethylester **9**. In addition to the phenyl ring, the squaric acid part also interacts with the tyrosine gate, thus increasing the contact area between protein and ligand. In addition, **9** interacts with the hydroxyl group of the distal THR51 by hydrogen-bridging. These interactions sum up, resulting in a higher affinity as reflected by lower IC_{50} values on the one hand and lower scoring values on the other. Circled in yellow, a depression in the hydrophobic ridge can be seen, which is not satisfied by ligand **9**. b) Top scoring conformation of the *o*-chloro-substituted squaric acid ethylester **10** depicted in the CRD of FimH as a CPK model. The chloro substituent in the *o*-position of the phenyl ring fits into a depression of the hydrophobic ridge, thus further increasing the contact area between protein and ligand, as the favorable interaction of the aglycon moiety of **10** remains the same as in the case of its analog **9**.

allow the squaric acid moiety to properly interact with the tyrosine gate and also H-bonding with the tyrosine OH-group is no more possible (Figure not shown).

Upon validation of the fit of mannoside **9** within the CRD of FimH, a lack of contact between protein and ligand in a depressed area of the hydrophobic ridge (encircled in yellow in Fig. 3a) became obvious. Thus, we have anticipated that the binding affinity of **9** might be further improved by introduction of a hydrophobic substituent in the *o*-position of the phenyl ring. Therefore, we started the synthesis of *o*-chlorophenyl mannoside **10** employing literature-known mannoside **8** (Scheme 2).[25] Reduction of the *p*-nitro group in **8** and subsequent coupling with DES could be carried out in a one-pot procedure in good yield without isolation of the intermediate amine. Testing results obtained with mannoside **10** supported our initial considerations. Its inhibitory potency was increased up to 6900-fold as compared to MeMan (**1**) and almost four-fold when compared to **9** (Table 1).

Interestingly, docking of **10** and scoring of the obtained ligand conformations delivered a relatively weak scoring value of −26.9, thus **10** scores worse than **9** (score −29.2) though it performs better in the ELISA. On the other hand, all top-ranking conformations of the *o*-chloro-substituted mannoside **10** reflect an increased lipophilic contact area between the ligand and the CRD (Fig. 3b). Thus, in the case of ligand **10** the "in silico" findings are only partly in agreement with the results obtained by biological testing. It should be kept in mind that the potent ligand **10** was discovered by rational concluding from the docking solutions for mannoside **9** rather than by unreflected ("blind") docking and scoring.

Then we set out to improve the binding affinity of ligands by further increasing the lipophilic contact surface between the ligand aglycon and the entrance of the binding pocket. We decided to attach various peptide side chains in the *p*-position of the aromatic phenyl ring of **9** by using the squaric acid linkage chemistry. A first ligation of DES with the amine **7** led to **9** and then a second ligation step with an amino acid or peptide derivative, respectively, was performed in basic media exploiting the capacity of squaric acid amide monoesters to react with a second amine only under basic conditions.[15] This procedure yielded the amino acid glycosides **11**, **12**, **13** and **14**, respectively (Scheme 3) in yields between 40 and 60%. Triglycine, tetraglycine, pentaglycine and tyrosine methylester were used in the last ligation step. The somewhat moderate yields are due to the reduced nucleophilicity of the N-terminal amino functions of the employed peptides.

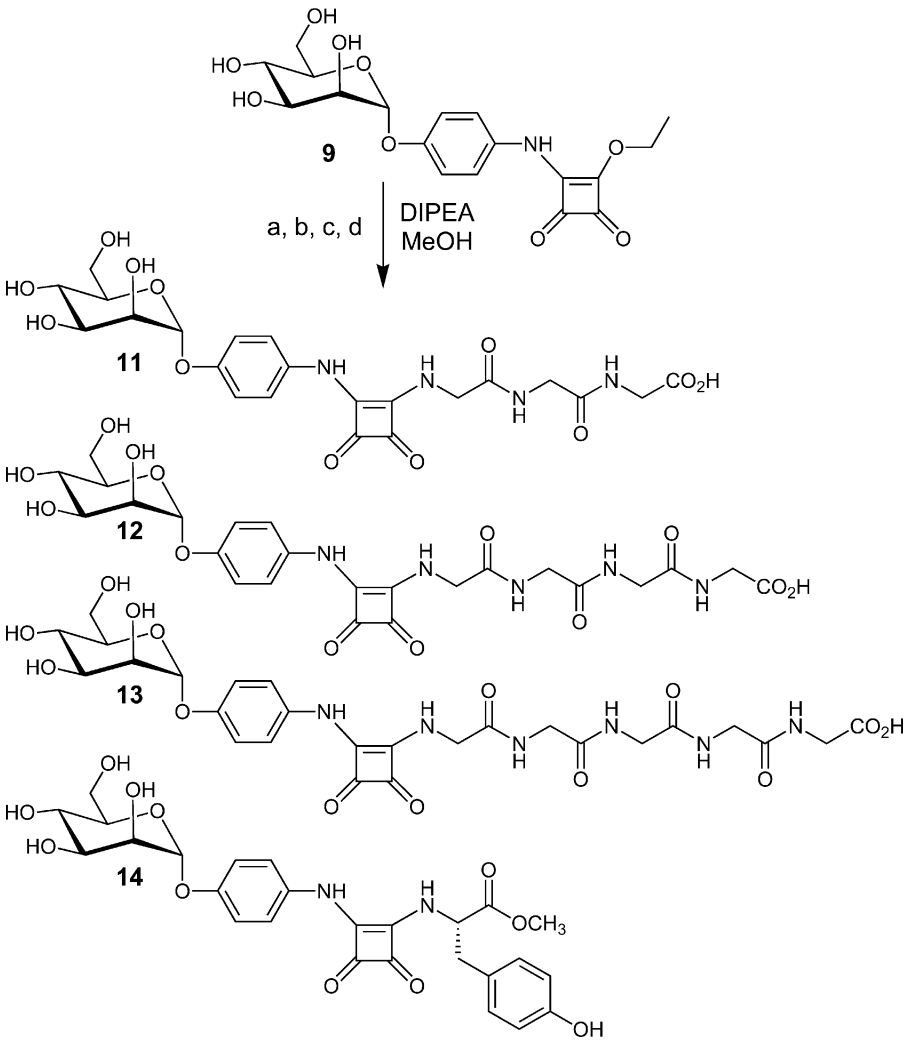
Docking of glycopeptides **11** to **13** reveals a complex situation. These glycopeptide derivatives with elongated peptide aglycon moieties carrying three, four and five glycin residues, respectively, result in top-ranking conformations, with excellent scoring values between −45 and −47, however the mannosyl glycon is not placed within the binding pocket anymore. In all three cases, top-ranking conformations refer to the molecule wound around the protein to maximize the ligand–receptor contact area. The increased lipophilic contact surface results in a very low scoring value however, the underlying docking solutions cannot be considered as reasonable, based on the initial assumption that ligand binding of FimH depends on complexation of an α-D-mannosyl residue within the CRD.

This finding is typical for docking of larger ligands, as rather large molecules tend to give higher scores. In addition, the number of rotable bonds reaches a value which is limiting for the FlexX algorithm, thus leading to unreliable results.[23] The situation in the case of the glycopeptide mimetic **14** is similar, except that, owing to the aromatic side chain in the aglycon moiety, none of the docked conformations can be oriented such that the molecule receives a convenient fit with regard to the protein surface of FimH.

To allow at least the comparison of the docking results obtained for **11**, **12**, **13** and **14** with those obtained with **1–4** and **6–10**, only those conformations out of 90 docking solutions were considered, which have the mannosyl aglycon positioned properly inside the CRD. From this collection of solutions, the highest ranking hit was included in Table 1. After all, the scores obtained for ligands **11–14** cannot be expected to correlate with the affinities deduced from ELISA.

Considering the results obtained for the ligands **1**, **2**, **3**, **4**, **6**, and **8–10**, it can be seen from Table 1 that evaluation of ligand quality by either docking and scoring of the docking results or by ELISA reflected corresponding trends in many, but not all cases. In the case of the methylumbelliferyl mannoside **3**, docking places this ligand at the last position of the ranking list of the above mentioned eight mannosides, whereas biological testing reflects the third-best inhibitory potency for **3** among the same eight ligands.

The rather weak scoring value of −20.2 obtained for **3** is the result of FlexX scoring based on the X-ray structure of FimH in complex with mannose.[12] When the most recent crystal structure for FimH, in which the lectin is complexed with butyl α-D-mannoside (**4**),[13] was used as the basis for docking, a different assessment of the binding energy of **3** resulted, delivering a score of −29.0, which correlates much better with its measured inhibitory potency. Searching for the critical difference between the two different X-ray structures of FimH, it can be seen that they differ only in the conformation of the tyrosine gate at the



Scheme 3 Synthesis of the squaric acid-linked glycopeptides. Reagents and conditions: for **11** a) triglycine, DIPEA, MeOH, 10 h, acidic ion exchange resin, 41%; for **12** b) tetraglycine, DIPEA, MeOH, 10 h, acidic ion exchange resin, 49%; for **13** c) pentaglycine, DIPEA, MeOH, 10 h, acidic ion exchange resin, 50%; for **14** d) L-tyrosine methylester, DIPEA, MeOH, 10 h, 62%.

entrance of the binding pocket (Fig. 4). In the case of the FimH crystal with α-mannose in the binding pocket, the conformation of this tyrosine gate can be considered as "open", whereas FimH in complex with butyl α-D-mannoside has crystallized with the tyrosine gate "closed". This is due to a change in the torsion angle in the TYR48 substructure CO–Cα–Cβ–C_{aryl} from 70.1° ("open") to 201.3° ("closed").

To estimate the impact of this difference on scoring, all docking studies were performed once again using the "closed" gate lectin structure as the receptor. In Table 2 the new ranking results are collected and compared to the older ones from Table 1, which were obtained with the "open" gate FimH structure. In Table 3 a ranking list for the investigated compounds is provided which was deduced from the results from three different approaches.

Table 2 Comparison of the scoring values for eight selected ligands as obtained from docking based on two different crystal structures

Ligand	Score mannose[a] "open" gate	Score BuMan (**4**)[b] "closed" gate	RIP based on MeMan
1 (MeMan)	−22.5	−23.3	1
2 (pNPMan)	−24.9	−27.4	31
3 (MeUmbMan)	−20.2	−29.0	380
4 (BuMan)	−20.5	−21.5	5.7
6	−26.5	−23.3	2.7
8	−25.4	−27.9	200
9	−29.2	−33.2	1800
10	−26.9	−32.8	6900

[a] Based on crystal structure of FimH complexed with α-mannose.[12] [b] Based on crystal structure of FimH complexed with *n*-butyl α-mannoside.[13]

Table 3 Ranking of compounds collected in Table 2 according to scoring and ELISA; ranking leads from weak to better inhibitors. Significant variations of the ranking resulting from different evaluation approaches concern especially compounds **3** and **6**

Ranking list (increasing potencies)	**According to scoring** with **"open" gate** structure[a]	**According to scoring** with **"closed" gate** structure[b]	**According to ELISA**
1.	**3**	**4**	**1**
2.	**4**	**1, 6**	**6**
3.	**1**		**4**
4.	**2**	**2**	**2**
5.	**8**	**8**	**8**
6.	**6**	**3**	**3**
7.	**10**	**10**	**9**
8.	**9**	**9**	**10**

[a] Based on crystal structure of FimH complexed with α-mannose.[12]

[b] Based on crystal structure of FimH complexed with *n*-butyl α-mannoside.[13]

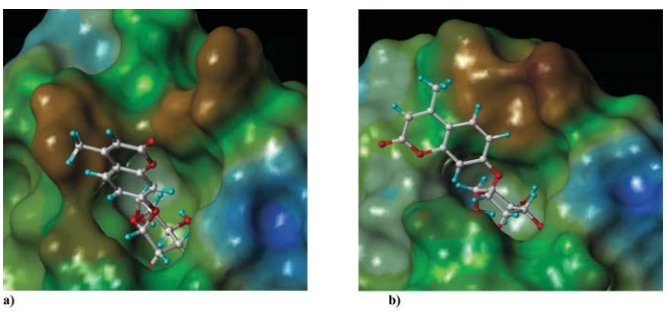


Fig. 4 Comparison of two different X-ray structures available for FimH. a) The crystal structure of FimH complexed with α-mannose is depicted (with the tyrosine gate "open").[12] This structure allows only a suboptimal interaction of the aromatic aglycon moiety of mannoside **3** (MeUmbMan) with the tyrosine residues at the entrance of the CRD. Consequently FlexX docking delivers a relatively high (less negative) scoring value (−20.2), and this weak score is in disagreement with the relatively high inhibitory potency of **3** as measured by ELISA. b) The crystal structure obtained of FimH complexed with butyl α-mannoside (**4**) is depicted (with the tyrosine gate "closed").[13] When this structure was used for ligand docking, the aglycon of **3** can be favorably aligned with the surface area of the protein CRD, leading to a more negative score of −29.0. This low score correlates with the experimental affinity of **3** as measured in the ELISA.

It can be seen that the scores obtained based on the *n*-butyl mannoside-complexed FimH structure are in fine accordance with the ranking as measured by ELISA, other than in case of our first study based on the X-ray analysis with FimH in complex with mannose. The discrepancy concerns especially methylumbelliferyl α-D-mannoside (**3**) and the squaric-acid modified mannoside **6**.

This can be reasoned by differences regarding the exposed lipophilic surface at the CRD entrance. In the case of the BuManFimH structure (Fig. 4b), the entire aromatic surface of the aglycon moiety of mannoside **3** can interact with the CRD entrance, while maintaining an optimal fit of the mannosyl glycon within the CRD. On the other hand, less favorable interactions are established between the same ligand and the "open-gate" structure (Fig. 4a). This increase in contact area changing from the "open" to the "closed" gate receptor conformation is even more pronounced in the case of the squaric acid derivatives **9** and **10**, resulting in drastically enhanced, that is lower, scoring values. Similar considerations may explain the differing values for **6**.

The docking and scoring results for the *n*-butyl mannoside-based crystal structure are in good agreement with the experimental ranking for the compounds with a more rigid aglycon. Inadequate ranking occurs again for compounds **11–14** having a very flexible aglycon. In the scoring process, the entropic penalty, which has to be paid upon binding of this type of highly flexible ligands outweighs the energetic gain, which is achieved through an increased contact surface between ligand and receptor.

It seems that most of the investigated ligands can interact more efficiently with the "closed" gate structure of FimH. However, the tyrosine gate is a very flexible region of the protein not enclosed or hindered by other protein parts. For the *n*-butyl mannoside **4** both protein conformations, "open" and "closed", can be found.[26] This might indicate a protein–ligand interaction of an "induced fit"-type, which depends on the type of bound ligand, forcing the protein to adopt a certain conformation.

Conclusions

We have combined our know-how in molecular modeling, carbohydrate synthesis and glycobiology to shed light on the molecular details of carbohydrate binding as mediated by bacterial type 1 fimbriae. We employed the structural information about the protein FimH, the adhesive sub-structure of type 1 fimbriae, which is available from, so far, three X-ray studies. These X-ray studies are consistent in that they reveal a single carbohydrate-binding site at the tip of FimH, which can accommodate an α-D-mannosyl residue, with the aglycon of an α-D-mannoside sticking out of the CRD. On the basis of this information about the adhesin FimH, multivalency effects as measured in various inhibition studies cannot be understood or rationalized (*cf.* preceding paper[8]). However, the differing potencies of α-D-mannosides with varying aglycon portions as inhibitors of type 1 fimbriae-mediated bacterial adhesion can be rationalized by comparison of the results obtained by computer-aided docking on one hand and ELISA on the other. In some of the investigated cases, especially in the case of **3**, the two sets of information (obtained by docking based on the ManFimH structure[12] and ELISA, respectively) were not congruent and this prompted us to compare the available crystal structrues of FimH. While the two X-ray studies published first[11,12] showed no significant difference with regard to the protein structure, the most recent X-ray analysis[13] revealed a different conformation of the two tyrosine residues at the periphery of the CRD, resembling a "closed" gate rather than an "open" one. Depending on which structure was employed as the basis for ligand docking, scoring of the docking results for a series of investigated ligands led to different ranking lists. This "in silico" result, which has been discussed in this paper, also points at the biological circumstances of ligand binding to FimH. It

has to be considered that the flexibility of the protein receptor in solution and its CRD, respectively, can significantly influence ligand binding, and this understanding may have important implications for the biology of carbohydrate-dependent bacterial adhesion, which awaits further investigation. Even the finding that fimbriae-mediated bacterial adhesion is often flow-regulated[27] might be reconsidered with regard to lectin flexibility.

A second conclusion which can be drawn from this study with regard to ligand design is that a strategy counting on additional interactions mediated by the aglycon moiety of α-D-mannoside ligands to improve receptor binding has its limitations. Once the aglycon gets too large it may establish enough interactions in the periphery of the FimH CRD so that the mannosyl glycon gets disconnected from the binding pocket, resulting in an overall diminished affinity. In addition, in the case of mannosides with a highly flexible aglycon, the entropic penalty which has to be paid for the complexation with the receptor protein FimH compensates the enthalpic gain. Such an understanding of ligand binding to FimH will also be of importance once binding of multivalent carbohydrate ligands is under investigation.

It will be the aim of our future work to gain a better insight into the processes of bacterial adhesion, based on the results of this contribution and of the preceding paper[8]; and we will direct our attention to the mechanistic and biological implications of our research.

Experimental

General remarks

All chemicals and solvents were used without further purification, with the exception of methanol which was distilled prior to use. Optical rotations were measured with a Perkin Elmer 241 polarimeter (10 cm cells, sodium D line: 589 nm). NMR spectra were recorded at 300, 500 or 600 MHz on Bruker ARX 300 (300 MHz for ^{1}H, 75.47 for ^{13}C), Bruker DRX 500 (500 MHz for ^{1}H, 125.75 for ^{13}C) and Bruker ARX 600 (600 MHz for ^{1}H, 150.92 MHz for ^{13}C) instruments with Me_4Si ($\delta = 0$) as internal standard. Flash column chromatography was performed on silica gel 60 (230–400 mesh, Merck). Purifications on RP-Gel were carried out with a Büchi MPLC apparatus using Merck Licroprep RP18 coloums. MALDI-TOF MS spectra were measured on a Bruker Biflex apparatus at 19 kV Voltage with a sinapinic acid matrix (saturated sinapinic acid in 9 : 1 water–acetronitrile containing 0.1% trifluoroacetic acid). Optical densities (ODs) were measured on an Asys DigiScan 400 ELISA reader at 405 nm with the reference read to 492 nm. ELISA plates were incubated at 37 °C.

Reagents. Methyl α-D-mannoside (**1**) was purchased from Fluka, *p*-nitrophenyl α-D-mannoside (**2**) from Senn chemicals, methylumbelliferyl α-D-mannoside (**3**) and *n*-butyl α-D-mannoside (**4**) were prepared according to the literature.[13] F-shaped 96-well microtiter plates from Sarstedt. Mannan from *Saccharomyces cerevisiae* was purchased from Sigma and was used in 50 mM aq. Na_2CO_3 (1 mg ml^{-1}, pH 9.6). The peroxidase-conjugated goat anti-rabbit antibody (IgG, H + L) was purchased from Dianova. Skimmed milk was from Ulzena, Tween 20 from Roth, ABTS [2,2′-azidobis-(3-ethylbenzothiazoline-6-sulfonic acid)] from Fluka and thimerosal (2-(ethylmercuriothio) benzoic acid sodium salt) was from Merck. A recombinant type 1 fimbriated *E. coli* strain, *E. coli* HB101 (pPKl4),[28] was used and cultured as described earlier.[29]

Buffers. PBS (phosphate-buffered saline) was prepared by dissolving 8 g NaCl, 0.2 g KCl, 1.44 g $Na_2HPO_4{\cdot}2H_2O$ and 0.2 g KH_2PO_4 in 1000 ml of bidest. water (pH 7.2). PBSE was PBS buffer + 100 mg l^{-1} thimerosal, PBSET was PBSE buffer + 200 µl l^{-1} Tween 20. Substrate buffer was 0.1 M sodium citrate dihydrate, adjusted to pH 4.5 with citric acid. For preparation of the ABTS solution, ABTS (1 mg) was dissolved in substrate buffer (1 ml) and 0.1% H_2O_2 (25 µl per ml) was added.

Docking

For investigation of the binding mode of various FimH ligands, high resolution crystal structures of FimH as available in the PDB database (FimH–FimC complex with CHEGA: 1QUN,[30] FimH–FimC complex with mannose: 1KLF,[31] FimH with *n*-butyl α-D-mannoside: 1UWF[32]) were used. Docking was accomplished applying the automated flexible docking tool FlexX 1.10.0[20] along with Sybyl6.8[22] as interface. The receptor was defined as a sphere with a diameter of 10 Å around the mannose-binding pocket (specifying the carboxyl carbon atom of ASP54 as the centre), to assure that all possible interactions are considered. The CRD was described by the following amino acids: PHE1 ALA2 ILE13 GLY14 HIS45 ASN46 ASP47 TYR48 PRO49 GLU50 THR51 ILE52 ASP54 TYR55 ARG98 GLN133 THR134 ASN135 ASN136 TYR137 ASN138 SER139 ASP140 ASP141 PHE142 PHE144.

The crystal structures were used without further modification (no minimization and side chain relaxation). Docking was performed applying the FlexX standard parameters considering 30 docked solutions for scoring, unless otherwise stated. Formal charges were used for the docking routine. To produce a more robust evaluation of ligand–receptor interactions, a consensus of scoring functions was calculated.[21,33] The following scoring functions available to the CScore module in Sybyl6.8 were considered: FlexX original score, ChemScore,[34] DockScore,[35] GoldScore,[36] PMFScore.[37] The obtained solutions were relaxed and scored using the above mentioned scoring functions. The FlexX original scoring function was not available for re-scoring. Relaxation was performed using the Tripos FF and standard parameters available in Sybyl6.8. Rmsd values for the sugar moieties were calculated comparing the positions of the atoms of the mannose ring (except hydrogen) of a particular binding mode with the positions of the respective mannose ring taken from the experimental crystal structure using an spl-script.

Visualization of results. The protein is presented as Connolly surface representation[38] (solvent accessible surface) and colored according to its lipophilic potential.[39] Brown colors represent lipophilic areas, hydrophilic parts are shown in blue colors. Molecules are presented as ball and stick models (carbon = white, oxygen = red, nitrogen = blue, hydrogen = cyan, chlorine = green).

ELISA

To determine the potencies of the various mannoside derivatives tested as inhibitors of type 1 fimbriae-mediated adhesion of *E. coli*, an ELISA was used as published earlier.[29] Polystyrene microtiter plates were coated with mannan solution (100 µl per well) and

dried overnight at 37 °C. The plates were blocked once with 5% skimmed milk in PBSE for 30 min at 37 °C. The wells were washed with PBSE (150 µl) and then PBSE (50 µl) and inhibitor solutions (50 µl) were added. Inhibitor solutions were diluted serially two-fold in PBSE. Bacterial suspension (50 µl per well) was added and the plate was left at 37 °C for 1 h to allow sedimentation of the bacteria. Then each well was washed four times with PBSE (150 µl) and 50 µl of the first antibody (anti-fimA antibody, solution as optimized prior to the experiments) in 2% skimmed milk was added. The plates were incubated for 30 min and then washed twice with PBSET and the second antibody was added (50 µl). The plates were incubated for 30 min and then washed three times with PBSET and once with PBSE and substrate buffer. ABTS solution (50 µl) was added, incubated for 60 min at 37 °C. For ELISA controls, bacterial adhesion to blocked, uncoated microtiter plates was checked, and the reaction of the employed antibodies with yeast mannan was tested and found to be negligible. The low background was substracted when calculating the IC_{50} values. The percentage inhibition was calculated as OD(nI) − OD(I) × 100 × [OD(nI)]$^{-1}$ (nI: no inhibitor, I: with inhibitor).

IC_{50} values are average values from at least three independent assays and are listed together with their standard deviations. Relative inhibitory potencies (RIPs) are based on the IC_{50} value of methyl α-D-mannopyranoside (MeMan), with RIP (MeMan) ≡ 1.

***p*-[*N*-(2-Ethoxy-3,4-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 9.** Mannoside **7**[19] (800 mg, 2.95 mmol) was dissolved in MeOH (10 ml), diethyl squarate (500 µl, 3.38 mmol) was added and the reaction mixture was stirred at rt for 1 d. Then, the solvent was removed *in vacuo* and the residue was purified by two subsequent chromatographic steps, first by flash chromatography on silica gel (MeOH–ethyl acetate, 1 : 1) followed by chromatography on RP-18 (MeOH–H_2O, 1 : 1) to furnish the title mannoside as white lyophilisate. Yield: 1.12 g (2.84 mmol, 63%); $[\alpha]_D^{20}$ +82.3 (*c* 1.0, DMSO). δ_H (500 MHz, D_6-DMSO, D_2O-exchange) 7.27 (2H, bs, aryl-H), 7.07 (2H, d, ^3J = 9.0, aryl-H), 5.31 (1H, d, $^3J_{1,2}$ = 1.9, H-1), 4.73 (2H, q, ^3J = 7.1, OCH_2CH$_3$), 3.80 (1H, dd, $^3J_{2,3}$ = 3.4, H-2), 3.65 (1H, dd, $^3J_{3,4}$ = 9.0, H-3), 3.58 (1H, dd, ^3J = 1.7, $^2J_{6a,6b}$ = 11.7, H-6a), 3.57–3.41 (3H, m, H-5, H-6b, H-4), 1.39 (3H, t, OCH$_2$CH_3) ppm; δ_C (50.32 MHz, 313 K, D_6-DMSO) 187.96, 183.26 (C=O squaric acid), 177.73, 169.32 (C=C squaric acid), 153.22 (man-*O*-C_{aryl}), 132.28 (*p*-C_{aryl}), 121.05 (2 × *m*-C_{aryl}), 117.39 (2 × *o*-C_{aryl}), 99.31 (C-1), 74.82 (C-5), 70.67 (C-3), 70.01 (C-2), 69.30 (CH_2), 66.83 (C-4), 61.10 (C-6), 15.51 (CH_3) ppm; MS (MALDI-TOF): *m*/*z*: 418.5 [M + Na]$^+$ (calcd. 418.1) for $C_{18}H_{21}NO_9$ (MW 395.12), HRMS (ESI), found: 418.1109 [M + Na]$^+$, $C_{18}H_{21}NO_9$ requires 418.1107 [M + Na]$^+$.

***o*-Chloro-*p*-[*N*-(2-ethoxy-3,4-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 10.** Mannoside **8** (290 mg, 0.86 mmol) was dissolved in MeOH (10 ml) and stirred for 1 h with Pd/C-catalyst (100 mg) under a hydrogen atmosphere (1 bar) at rt. The suspension was filtered and diethyl squarate (300 µl, 2.03 mmol) was added. The reaction mixture was stirred overnight at rt, then the solvent was removed *in vacuo* and the residue was purified by MPLC on RP-18 (MeOH–H_2O, 1 : 3) to obtain the title compound as a white lyophilisate. Yield: 200 mg (0.47 mmol, 55%); $[\alpha]_D^{20}$ +50.0 (*c* 0.15 in DMSO); δ_H (500 MHz, D_6-DMSO) 7.49 (1H, bs, aryl-H), 7.33 (1H, d, $^3J_{ar.}$ = 9.0, aryl-H), 7.24 (1H, d, aryl-H), 5.39 (1H, d, $^3J_{1,2}$ = 1.7, H-1), 4.76 (2H, q, ^3J = 7.0, OCH_2CH$_3$) 3.88 (1H, dd, $^3J_{2,3}$ = 3.2, H-2), 3.71 (1H, dd, $^3J_{3,4}$ = 9.1, H-3), 3.40–3.63 (4H, m, H-5, H-6a, H-6b, H-4), 1.41 (3H, t, OCH$_2$CH_3) ppm; δ_C (50.32 MHz, 313 K, D_6-DMSO) 186.60, 183.41 (C=O squaric acid, very small peaks), 178.20, 169.31 (C=C squaric acid), 148.13 (man-*O*-C_{aryl}), 133.53 (*p*-C_{aryl}), 122.93 (*m*-C_{aryl}), 121.16 (*m*-C_{aryl}), 119.45 (*o*-C_{aryl}-Cl), 118.27 (*o*-C_{aryl}), 99.85 (C-1), 75.21 (C-5), 70.58 (C-3), 69.90 (C-2), 69.39 (CH_2), 66.63 (C-4), 61.01 (C-6), 15.47 (CH_3) ppm; MALDI-ToF MS: *m*/*z* = 452.0 [M + Na]$^+$ (calcd. 452.1) for $C_{18}H_{20}ClNO_9$ (M = 429.08), HRMS (ESI), found: 452.0799 [M + Na]$^+$.

***p*-[*N*-(4-Triglycyl-2,3-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 11.** The squaric acid monoester **9** (100 mg, 0.25 mmol) and triglycine (50 mg, 0.26 mmol) were dissolved in MeOH (10 ml) and DIPEA (100 µl) was added. The reaction mixture was stirred overnight at rt and subsequently acidified with Amberlite IR-120 ion exchange resin. The solvent was removed *in vacuo* and the residue purified by MPLC on RP-18 (MeOH–H_2O, 1 : 3) to furnish the title compound as a white lyophilisate. Yield: 55 mg (0.10 mmol, 41%); $[\alpha]_D^{20}$ +66.7 (*c* 0.15 in DMSO); δ_H (500 MHz, D_6-DMSO, D_2O-exchange) 7.36 (2H, d, ^3J = 9.0, aryl-H), 7.08 (2H, d, aryl-H), 5.29 (1H, d, $^3J_{1,2}$ = 1.8, H-1), 4.37 (2H, d, ^3J = 4.9, NH-CH_2COOH), 3.73–3.83 (5H, m, H-2, 2 × NH-CH_2-CONH), 3.67 (1H, dd, $^3J_{2,3}$ = 3.3, $^3J_{3,4}$ = 8.8, H-3), 3.62 (1H, dd, $^3J_{5,6a}$ = 1.7, $^2J_{6a,6b}$ = 11.3 Hz, H-6a), 3.41–3.50 (3H, m, H-4, H-5, H-6b) ppm; δ_C (125.76 MHz, D_6-DMSO) 183.56, 180.67 (C=O squaric acid), 171.07, 171.02, 168.95 (C=O), 163.84 (2 × C=C squaric acid), 152.28 (i-C_{aryl}), 133.54 (*p*-C_{aryl}), 119.35 (2 × *o*-C_{aryl}), 117.93 (2 × *m*-C_{aryl}), 99.45 (C-1), 74.92, 70.68, 70.12, 66.76 (C-2, C-3, C-4, C-5), 61.09 (C-6), 45.93, 41.84, 40.56, 3 × Gly-CH_2) ppm; MALDI-ToF MS: *m*/*z* = 561.4 [M + Na]$^+$ (calcd. 561.1) for $C_{22}H_{26}N_4O_{12}$ (M = 538.15), HRMS (ESI), found: 561.1424 [M + Na]$^+$, $C_{22}H_{26}N_4O_{12}$ requires 561.1439 [M + Na]$^+$.

***p*-[*N*-(4-Tetraglycyl-2,3-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 12.** The squaric acid monoester **9** (150 mg, 0.38 mmol) and tetraglycine (50 mg, 0.36 mmol) were dissolved in a 1 : 1 mixture of MeOH and H_2O (10 ml) and DIPEA (100 µl) was added. The reaction mixture was stirred overnight at rt and subsequently acidified with Amberlite IR-120 ion exchange resin. The solvent was removed *in vacuo* and the residue purified by MPLC on RP-18 (MeOH–H_2O, 1 : 3) to furnish the title compound as a white lyophilisate. Yield: 104 mg (0.23 mmol, 49%); $[\alpha]_D^{20}$ +71.6 (*c* 0.25 in DMSO); δ_H (600 MHz, D_6-DMSO, D_2O-exchange) 7.35 (2H, d, ^3J = 8.9, aryl-H), 7.07 (2H, d, aryl-H), 5.29 (1H, m, $^3J_{1,2}$ = 1.5, H-1), 4.37 (2H, s NCH_2COOH), 3.72–3.83 (7H, m, H-2, 3 × NH–CH_2-CONH), 3.66 (1H, dd, $^3J_{2,3}$ = 3.4, $^3J_{3,4}$ = 9.1, H-3), 3.59 (1H, dd, $^3J_{5,6a}$ = 1.5, $^2J_{6a,6b}$ = 11.5, H-6a), 3.39–3.50 (3H, m, H-4, H-5, H-6b) ppm; δ_C (150.92 MHz, D_6-DMSO) 183.59, 180.77 (C=O squaric acid), 171.22, 169.19, 168.98, 168.61 (4 × C=O), 163.86 (2 × C=C squaric acid), 152.33 (i-C_{aryl}), 133.59 (*p*-C_{aryl}), 119.41 (2 × *o*-C_{aryl}), 117.98 (2 × *m*-C_{aryl}), 99.46 (C-1), 74.99, 70.71, 70.17, 66.76 (C-2, C-3, C-4, C-5), 61.10 (C-6), 46.00, 42.11, 41.75, 40.61, 4 × Gly-CH_2) ppm; MALDI-ToF MS: *m*/*z* = 617.8 [M + Na]$^+$ (calcd. 618.2) for $C_{24}H_{29}N_5O_{13}$ (M = 595.2), HRMS (ESI), found: 618.1684 [M + Na]$^+$, $C_{24}H_{29}N_5O_{13}$ requires 618.1654 [M + Na]$^+$.

***p*-[*N*-(4-Pentaglycyl-2,3-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 13.** The squaric acid monoester **9** (100 mg, 0.25 mmol) and pentaglycine (74 mg, 0.24 mmol) were dissolved in a 1 : 2 mixture of MeOH and H_2O (10 ml). DIPEA (100 µl) was added. The reaction mixture was stirred overnight at rt and subsequently acidified with Amberlite IR-120 ion exchange resin. The solvent was removed *in vacuo* and the residue purified by MPLC on RP-18 (MeOH–H_2O, 1 : 3) to furnish the title compound as a white lyophilisate. Yield: 78 mg (0.12 mmol, 50%); $[\alpha]_D^{20}$ +73.3 (*c* 0.25 in DMSO); δ_H (600 MHz, D_6-DMSO, D_2O-exchange) 7.35 (2H, d, $^3J_{Har.}$ = 8.9, aryl-H), 7.07 (2H, d, aryl-H), 5.28 (1H, m, $^3J_{1,2}$ = 1.6, H-1), 4.36 (2H, s NCH_2COOH), 3.72–3.82 (9H, m, H-2, 4 × NH–CH_2-CONH), 3.65 (1H, dd, $^3J_{2,3}$ = 3.4, $^3J_{3,4}$ = 9.1, H-3), 3.59 (1H, dd, $^3J_{5,6a}$ = 1.6, $^2J_{6a,6b}$ = 11.5, H-6a), 3.39–3.52 (3H, m, H-4, H-5, H-6b) ppm; δ_C (150.92 MHz, D_6-DMSO) 183.60, 180.77 (C=O squaric acid), 171.21, 169.19, 169.11, 169.08, 168.63 (5 × C=O), 163.86 (2 × C=C squaric acid), 152.33 (i-C_{aryl}), 133.58 (*p*-C_{aryl}), 119.39 (2 × *o*-C_{aryl}), 117.97 (2 × *m*-C_{aryl}), 99.45 (C-1), 74.98, 70.69, 70.16, 66.75 (C-2, C-3, C-4, C-5), 61.10 (C-6), 45.99, 42.33, 42.03, 41.75, 40.61, 5 × Gly-CH_2) ppm; MALDI-ToF MS: m/z = 674.6 $[M + Na]^+$ (calcd. 675.2) for $C_{26}H_{32}N_6O_{14}$ (M = 652.2), HRMS (ESI), found: 675.1860 $[M + Na]^+$, $C_{26}H_{32}N_6O_{14}$ requires 675.1869 $[M + Na]^+$.

***p*-[*N*-(4-L-Methyltyrosyl-2,3-dioxocyclobut-1-enyl)amino]phenyl α-D-mannopyranoside 14.** The squaric acid monoester **9** (20 mg, 50 µmol) and L-tyrosine methylester (10 mg, 51 µmol) were dissolved in a 4 : 1 mixture of MeOH and H_2O (1 ml). DIPEA (10 µl) was added and the reaction mixture was stirred at rt overnight. Then, the solvent was removed *in vacuo* and the residue purifed by MPLC on RP-18 (MeOH–H_2O, 1 : 4) to give the title compound as a white lyophilisate. Yield: 17 mg (31 µmol, 62%); $[\alpha]_D^{20}$ +41.5 (*c* 0.20 in DMSO); δ_H (500 MHz, D_6-DMSO) 7.32 (2H, d, ^3J = 8.7, aryl-H), 7.06 (2H, d, aryl-H), 6.96 (2H, d, ^3J = 8.5, aryl-H, Tyr), 6.67 (2H, d, aryl-H, Tyr), 5.29 (1H, d, $^3J_{1,2}$ = 1.7, H-1), 5.02 (1H, m, CHCH$_2$), 3.82 (1H, dd, $^3J_{2,3}$ = 3.3, H-2), 3.71 (3H, s, COOCH_3), 3.67 (1H, dd, $^3J_{3,4}$ = 9.0, H-3), 3.62 (1H, dd, $^3J_{5,6a}$ = 1.6, $^2J_{6a,6b}$ = 11.4, H-6a), 3.39–3.51 (3H, m, H-4, H-5, H-6b), 3.08 (1H, m, CHCH_2), 2.99 (1H, m, CHCH_2) ppm; δ_C (125.76 MHz, D_6-DMSO) 183.08, 181.23 (C=O squaric acid), 171.29 (C=O), 168.04, 164.09, (C=C squaric acid), 156.34 (C_{Ar}-OH), 152.36 (Man-*O*-C_{aryl}), 133.40 (*p*-$C_{arylTyr}$), 130.39 (*p*-C_{ArpAP}, 125.76 (*m*-$C_{arylTyr}$), 119.51 (*m*-C_{ArpAP}), 117.91 (*o*-$C_{arylTyr}$), 115.26 (*o*-C_{ArpAP}), 99.44 (C-1), 74.97 (C-5), 70.68 (C-3), 70.15 (C-2), 66.73 (C-4), 61.08 (C-6), 57.39 (CH), 52.46 (CH_3), 38.49 (CH_2) ppm; MALDI-ToF MS: m/z = 567.4 $[M + Na]^+$ (calcd. 567.2) for $C_{26}H_{28}N_2O_{11}$ (M = 544.2), HRMS (ESI), found: 567.1597 $[M + Na]^+$, $C_{26}H_{28}N_2O_{11}$ requires 567.1585 $[M + Na]^+$.

Acknowledgements

This work was supported by the DFG (Deutsche Forschungsgemeinschaft) in the frame of SFB 470.

References

1 (*a*) A. Andreau, M. Xercavins and F. Fernandez, *Med. Clin. (Barcelona)*, 1989, **92**, 409–12; (*b*) J. A. Snyder, A. L. Lloyd, C. V. Lockatell, D. E. Johnson and H. L. T. Mobley, *Infect. Immun.*, 2006, **74**, 1387–1393.
2 (*a*) B. J. Marshall and R. M. Warren, *Lancet*, 1984, **16**, 1311–1315; (*b*) M. J. Blaser, *Sci. Am.*, 2005, **292**, 38–45.
3 (*a*) E. H. Beachey, *J. Infect. Dis.*, 1981, **143**, 325–345; (*b*) C. Bavington and C. Page, *Respiration*, 2005, **72**, 335–344; (*c*) J. Berglund and S. D. Knight, *Adv. Exp. Med. Biol.*, 2003, **535**, 33–52.
4 (*a*) F. G. Sauer, M. Barnhart, D. Choudhury, S. D. Knight, G. Waksman and S. J. Hultgren, *Curr. Opin. Struct. Biol.*, 2000, **10**, 548–556; (*b*) M. Vetsch, C. Puorger, T. Spirig, U. Grauschopf, E.-U. Weber-Ban and R. Glockshuber, *Nature*, 2004, **431**, 330–332.
5 (*a*) A. B. Jonson, S. Normark and M. Rhen, *Contrib. Microbiol.*, 2005, **12**, 67–89; (*b*) F. K. Bahrani-Mougeot, E. L. Buckles, C. V. Lockatell, J. R. Hebel, D. E. Johnson, C. M. Tang and M. S. Donnenberg, *Mol. Microbiol.*, 2002, **45**, 1079–1093.
6 (*a*) N. Sharon, *FEBS Lett.*, 1987, **217**, 145–157; (*b*) N. Firon, I. Ofek and N. Sharon, *Carbohydr. Res.*, 1983, **120**, 235–249; (*c*) N. Nagahori, R. T. Lee, S.-L. Nishimura, S. Pagé, R. Roy and Y. C. Lee, *ChemBioChem*, 2002, **3**, 836–844; (*d*) A. Patel and T. K. Lindhorst, *Carbohydr. Res.*, 2006, **341**, 1657–1668.
7 (*a*) T. K. Lindhorst, *Spek. Wiss.*, **2000**, 16–20; (*b*) I. Ofek, D. L. Hasty and N. Sharon, *FEMS Immunol. Med. Microbiol.*, 2003, **38**, 181–191; (*c*) E. Fan and E. A. Merritt, *Curr. Drug Targets: Infect. Disord.*, 2002, **2**, 161–167.
8 M. Dubber, O. Sperling and T. K. Lindhorst, *Org. Biomol. Chem.*, 2006, DOI: 10.1039/b610741a.
9 (*a*) C. R. Bertozzi and L. L. Kiessling, *Science*, 2001, **291**, 2357–2364; (*b*) M. Mammen, S.-K. Choi and G. M. Whitesides, *Angew. Chem., Int. Ed.*, 1998, **37**, 2755–2794; (*c*) R. T. Lee and Y. C. Lee, *Glycoconjugate J.*, 2000, **17**, 543–551; (*d*) T. K. Lindhorst, *Top. Curr. Chem.*, 2002, **218**, 201–235; (*e*) L. L. Kiessling, J. E. Gestwicki and L. E. Strong, *Angew. Chem., Int. Ed.*, 2006, **45**, 2348–2368; (*f*) see also ref. 8.
10 (*a*) P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Amstrong, H. Ling, N. S. Pannu, R. J. Read and D. R. Bundle, *Nature*, 2000, **403**, 669–672; (*b*) T. K. Mandal and C. Mukhopadhyay, *Protein Eng.*, 2002, **15**, 979–986; (*c*) A. Larsson, S. M. C. Johansson, J. S. Pinkner, S. J. Hultgren, F. Almqvist, J. Kihlberg and A. Linusson, *J. Med. Chem.*, 2005, **48**, 935–945.
11 D. Choudhury, A. Thompson, V. Stojanoff, S. Langerman, J. Pinkner, S. J. Hultgren and S. Knight, *Science*, 1999, **285**, 1061–1066.
12 C. S. Hung, J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. Defusco, C. G. Auguste, R. Strouse, S. Langermann, G. Waksman and S. J. Hultgren, *Mol. Microbiol.*, 2002, **44**, 903–918.
13 J. Bouckaert, J. Berglund, M. Schembri, E. D. Genst, L. Cools, M. Wuhrer, C.-S. Hung, J. Pinkner, R. Slättegård, A. Zavialov, D. Choudhury, S. Langermann, S. J. Hultgren, L. Wyns, P. Klemm, S. Oscarson, S. D. Knight and H. D. Greve, *Mol. Microbiol.*, 2005, **55**, 441–455.
14 (*a*) N. Firon, S. Ashkenazi, D. Mirelman, I. Ofek and N. Sharon, *Infect. Immun.*, 1987, **55**, 472–476; (*b*) T. K. Lindhorst, C. Kieburg and U. Krallmann-Wenzel, *Glycoconjugate J.*, 1998, **15**, 605–613; (*c*) N. Röckendorf, O. Sperling and T. K. Lindhorst, *Aust. J. Chem.*, 2002, **55**, 87–93.
15 L. F. Tietze, M. Arit, M. Beller, K.-H. Glüsenkamp, E. Jähde and M. F. Rajewsky, *Chem. Ber.*, 1991, **124**, 1215–1221.
16 (*a*) E. Fan, Z. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde and W. G. J. Hol, *J. Am. Chem. Soc.*, 2000, **122**, 2663–2664; (*b*) B. Benaissa-Trouw, D. J. Lefeber, J. P. Kamerling, J. F. G. Vliegenthart, K. Kraaijeveld and H. Snippe, *Infect. Immun.*, 2001, **69**, 4698–4701; (*c*) N. Nagahori, R. T. Lee, S.-L. Nishimura, D. Pagé, R. Roy and Y. C. Lee, *ChemBioChem*, 2002, **3**, 836–844.
17 A. Y. Chernyak, G. V. M. Sharma, L. O. Kononov, P. R. Krishna, A. B. Levinskii, N. K. Kochetkov and A. V. R. Rao, *Carbohydr. Res.*, 1992, **223**, 303–309.
18 M. Izumi, S. Okumura, H. Yuasa and H. Hashimoto, *J. Carbohydr. Chem.*, 2003, **5**, 317–329.
19 O. Westphal and H. Feier, *Chem. Ber.*, 1956, **89**, 582–588.
20 (*a*) M. Rarey, B. Kramer, T. Lengauer and G. Klebe, *J. Mol. Biol.*, 1996, **261**, 470–489; (*b*) M. Rarey, B. Kramer and T. Lengauer, *J. Comput. Aided Mol. Des.*, 1997, **11**, 369–384; (*c*) B. Kramer, M. Rarey and T. Lengauer, *Proteins: Struct., Funct., Genet.*, 1999, **37**, 228–241; (*d*) M. Rarey, S. Wefing and T. Lengauer, *J. Comput. Aided Mol. Des.*, 1996, **10**, 41–54.
21 R. D. Clark, A. Strizhev, J. M. Leonard, J. F. Blake and J. B. Matthew, *J. Mol. Graphics Modell.*, 2002, **20**, 281–295.

22 Tripos Associates, 1699 South Hanley Rd., Suite 303, St. Louis, MO 63144. SYBYL, version 6.8, Tripos Inc. (1699 South Hanley Rd, St. Louis, Missouri, 63144).
23 D. Hoffmann, B. Kramer, T. Washio, T. Steinmetzer, M. Rarey and T. Lengauer, *J. Med. Chem.*, 1999, **42**, 4422–33.
24 P. S. Charifson, J. J. Corkery, M. A. Murcko and W. P. Walters, *J. Med. Chem.*, 1999, **42**, 5100–5109.
25 (*a*) A. Vervoort and C. K. De Bruyne, *Carbohydr. Res.*, 1970, **12**, 277–280; (*b*) J. De Prijcker, A. De Bock and C. K. De Bruyne, *Carbohydr. Res.*, 1978, **60**, 141–53.
26 A. Bergh, B.-G. Magnusson, J. Ohlsson, U. Wellmar and U. J. Nilsson, *Glycoconjugate J.*, 2001, **18**, 615–621.
27 (*a*) R. R. Isenberg and P. Barnes, *Cell*, 2002, **110**, 1–4; (*b*) L. M. Nilsson, W. E. Thomas, E. Trintchina, V. Vogel and E. V. Sokurenko, *J. Biol. Chem.*, 2006, **281**, 16653–16663.
28 P. Klemm, B. J. Jørgensen, I. van Die, H. de Ree and H. Bergmans, *Mol. Gen. Genet.*, 1985, **199**, 410–414.
29 T. K. Lindhorst, S. Kötter, U. Krallmann-Wenzel and S. Ehlers, *J. Chem. Soc., Perkin Trans. 1*, 2001, 823–831.
30 PDB ID: 1QUN; D. Choudhury, A. Thompson, V. Stojanoff, S. Langerman, J. Pinkner, S. J. Hultgren and S. Knight, *Science*, 1999, **285**, 1061–1066.
31 PDB ID: 1KLF; S. Hung, J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. Defusco, C. G. Auguste, R. Strouse, S. Langermann, G. Waksman and S. J. Hultgren, *Mol. Microbiol.*, 2002, **44**, 903–915.
32 PDB ID: 1UWF; J. Bouckaert, J. Berglund, M. Schembri, E. D. Genst, L. Cools, M. Wuhrer, C.-S. Hung, J. Pinkner, R. Slättegård, A. Zavialov, D. Choudhury, S. Langermann, S. J. Hultgren, L. Wyns, P. Klemm, S. Oscarson, S. D. Knight and H. D. Greve, *Mol. Microbiol.*, 2002, **44**, 903–915.
33 P. S. Charifson, J. J. Corkery, M. A. Murcko and W. P. Walters, *J. Med. Chem.*, 1999, **42**, 5100–5109.
34 M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini and R. P. Mee, *J. Comput. Aided Mol. Des.*, 1997, **11**, 425–445.
35 I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge and T. E. Ferrin, *J. Mol. Biol.*, 1982, **161**, 269–288.
36 (*a*) G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–748; (*b*) G. Jones, P. Willett and R. C. Glen, *J. Mol. Biol.*, 1995, **245**, 43–53.
37 I. Muegge and Y. C. Martin, *J. Med. Chem.*, 1999, **42**, 791–804.
38 (*a*) M. L. Connolly, *Science*, 1983, **221**, 709–713; (*b*) M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548.
39 (*a*) E. Audry, J. P. Dubost, J. C. Colleter and P. Dallet, *Eur. J. Med. Chem.*, 1986, **21**, 71–72; (*b*) P. Furet, A. Sele and N. C. Cohen, *J. Mol. Graphics*, 1988, **6**, 182–189; (*c*) W. Heiden, G. Moeckel and J. Brickmann, *J. Comput. Aided Mol. Des.*, 1993, **7**, 503.

