The Three-Dimensional Structure of the Extracellular Adhesion Domain of the Sialic Acid-Binding Adhesin SabA from *Helicobacter pylori*

Siew Siew Pang¹, Stanley Thai Son Nguyen¹,², Andrew J. Perry, Christopher J. Day⁴, Santosh Panjikar⁵, Joe Tiralongo⁴, James C. Whisstock¹,²* and Terry Kwok¹,³*

¹Department of Biochemistry and Molecular Biology, ²ARC Centre of Excellence in Structural and Functional Microbial Genomics, ³Department of Microbiology, Monash University, Clayton Campus, Melbourne, VIC 3800 Australia

⁴Institute for Glycomics, Griffith University, Gold Coast Campus, QLD, 4222 Australia

⁵Australian Synchrotron, 800 Blackburn Rd, Clayton, Melbourne, VIC, 3168 Australia

Running title: *Crystal Structure of SabA Adhesin from Helicobacter pylori*

To whom correspondence should be addressed: James C. Whisstock, Department of Biochemistry and Molecular Biology, ²ARC Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton Campus, Melbourne, VIC, 3800, Australia. Tel: 61-4-18170585; Fax: 61-3-99029500; Email: james.whisstock@monash.edu or Terry Kwok, Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, VIC, 3800 Australia. Tel: 61-3-99029216; Email: terry.kwok@monash.edu.

Keywords: *Helicobacter pylori*; outer membrane protein; bacterial adhesion; X-ray crystallography; carbohydrate binding protein

**Background:** *Helicobacter pylori* SabA outer membrane protein is crucial for the bacteria to adhere to host stomach surface during chronic infection.

**Results:** The structure of the extracellular SabA adhesion region is determined by X-ray crystallography

**Conclusion:** SabA adhesion structure may provide insight into binding sites important for ligand receptor interactions

**Significance:** SabA is the first reported extracellular domain structure of the *H. pylori* OMP family

**ABSTRACT**

The gastric pathogen *Helicobacter pylori* is a major cause of acute chronic gastritis and the development of stomach and duodenal ulcers. Chronic infection furthermore pre-disposes to the development of gastric cancer. Crucial to *H. pylori* survival within the hostile environment of the digestive system are the adhesins SabA and BabA; these molecules belong to the same protein family and permit the bacteria to bind tightly to sugar moieties LewisB and sialyl-LewisX, respectively, on the surface of epithelial cells lining the stomach and duodenum. Currently, no representative SabA / BabA structure has been determined, however, hampering the development of strategies to eliminate persistent *H. pylori* infections that fail to respond to conventional therapy. Here, using X-ray crystallography, we show that the soluble extracellular adhesin domain of SabA shares distant similarity to the tetratricopeptide repeat (TPR) fold family. The molecule broadly resembles a golf putter in shape, with the head region featuring a large cavity surrounded by loops that vary in sequence between different *H. pylori* strains. The N-terminal and C-terminal helices protrude at right angles from the head domain and together form a shaft that connects to a predicted outer membrane protein (OMP)-like β-barrel trans-membrane domain. Using surface plasmon resonance, we were able to...
detect binding of the SabA adhesin domain to sialyl-LewisX and LewisX but not to LewisA, LewisB or LewisY. Substitution of the highly conserved glutamine residue 159 in the predicted ligand-binding pocket abrogates the binding of the SabA adhesin domain to sialyl-LewisX and LewisX. Taken together, these data suggest that the adhesin domain of SabA is sufficient in isolation for specific ligand binding.

*Helicobacter pylori* is a Gram-negative bacterium that chronically infects more than half of the world population and is a major aetiological factor of peptic ulcer, atrophic gastritis and gastric adenocarcinoma and mucosa-associated lymphoid tissue (MALT) lymphoma (1). The ability to attach to the gastric epithelium is crucial for *H. pylori* to colonize and establish lifelong chronic infection in the host. The sialic acid binding adhesin (SabA) is one of the best characterized *H. pylori* adhesins (2,3). It belongs to the Hop superfamily of outer membrane proteins (OMP), which also includes the LewisB antigen-binding adhesin, BabA, and the glycoprotein-binding adhesins AlpA and AlpB. The four proteins all comprise a similar domain structure – an N-terminal extracellular domain of unknown structure, and a C-terminal trans-membrane domain that is predicted to adopt an integral eight-stranded β-barrel similar to the OMP-like family of integral membrane proteins (Fig. 1A).

SabA interacts with sialyl-LewisX antigen, an important blood group antigen and glycosphingolipid that is rarely expressed in healthy gastric epithelium but is present in abundance in inflamed and cancerous gastric epithelium (3). The SabA – sialyl-LewisX interaction is therefore likely to play a pivotal role in *H. pylori* colonization during chronic infection, as chronic *H. pylori* infection is almost always associated with chronic active gastritis (1) Epidemiological data suggest that interaction between SabA and gastric sialyl-LewisX is also important for *H. pylori* colonization in patients with no or weak LewisB antigen expression (4).

SabA-mediated binding of *H. pylori* to sialyl glycoconjugates and sialyl gangliosides requires NeuAcα2-3Gal disaccharide as the minimal binding epitope and is favoured by extended and flexible glycan core chains (5). Furthermore, SabA variants in different *H. pylori* clinical isolates exhibit different affinities and specificities for the sialylated glycans, sialyl dimeric LewisX antigen, sialyl-LewisA and dialyllactosamine (5). Such variable SabA ligand-binding specificity might be a host adaptation mechanism by which *H. pylori* rapidly modulates its adherence properties to achieve optimal colonization and concomitantly evade host immune responses.

Apart from binding to gangliosides, SabA also mediates binding of *H. pylori* to laminin in a manner that is dependent on the terminal α2-3-sialic acid of the glycan moiety on laminin (6). The physiological significance of this interaction remains to be elucidated. SabA is also essential for binding of *H. pylori* to sialylated glycoconjugates on erythrocytes and neutrophils (5,7). The latter interaction leads to non-opsonic oxidative burst involving G-protein signalling and activation of phosphatidylinositol 3-kinase (8).

Our knowledge of the various important functions of SabA has so far been based on the results of competitive inhibition of *H. pylori* adherence using soluble sialyl-LewisX conjugates or the loss of functions observed with *H. pylori* sabA-deletion mutants. The molecular mechanisms underpinning the various functions of SabA protein therefore remain unknown. Currently, however, it is predicted that the extracellular adhesin domain of SabA contains key determinants for interaction with the host cell surface (9). Accordingly, in order to investigate the structural basis of SabA adhesin function and furthermore to provide structural insights across the SabA/BabA superfamily, we set out to determine the three-dimensional structure of the soluble SabA adhesin domain.

**EXPERIMENTAL PROCEDURES**

**Cloning, expression and purification**–The full-length SabA gene was amplified and isolated from the genome library of *H. pylori* strain 26695 (forward primer 5′-ACAACAAAAAACATTACTTTAAGG-3′, reverse primer 5′-CAAGCTCTTTTTTAAAAGGG-3′). The mature SabA protein contains an N-terminal extracellular adhesin region (residue 1-460), which shows no sequence homology to any known protein structures, and a predicted C-terminal β-barrel domain (residue 461-635) which anchors SabA to the outer membrane of *H. pylori*. For this
structural study, the N-terminal soluble domain of SabA was cloned into pET15 vector for overexpression. The corresponding mutants with Tyr148, Lys152, Gln159 and Gln162 substituted by alanine were generated by site-directed mutagenesis using primers shown in Supplementary Table 1. A hexa-histidine tag and a TEV protease site were introduced at the start of the constructs to facilitate purification and removal of the tag as required.

The wild-type and mutant recombinant SabA proteins were expressed in E. coli Rosetta 2 (DE3) as insoluble inclusion bodies. The inclusion bodies were isolated, washed and solubilised in 50 mM Tris-HCl pH 8, 300 mM NaCl, 8 M guanidinium chloride. Protein purity and integrity were checked by SDS-PAGE. SabA was refolded by dialysis against 20 mM Tris-HCl pH 8, 50 mM NaCl. The refold solution was clarified by centrifugation and further purified by size-exclusion chromatography. The success of the refolding of wild-type SabA and SabA mutants was ascertained by circular dichroism (see Supplementary Fig. 2).

Selenomethionine-substituted SabA (SeMet-SabA) was expressed in E. coli BL21(DE3) as insoluble inclusion bodies. The inclusion bodies were isolated, washed and solubilised in 50 mM Tris-HCl pH 8, 300 mM NaCl, 8 M guanidinium chloride. Protein purity and integrity were checked by SDS-PAGE. SabA was refolded by dialysis against 20 mM Tris-HCl pH 8, 50 mM NaCl. The refold solution was clarified by centrifugation and further purified by size-exclusion chromatography. The success of the refolding of wild-type SabA and SabA mutants was ascertained by circular dichroism (see Supplementary Fig. 2).

Crystallization—For SabA crystallization, the N-terminal purification tag was first removed by digestion with the TEV protease. SabA protein was buffer-exchanged into 20 mM Tris-HCl pH 8, 50 mM NaCl and concentrated by ultra-filtration to 5 mg/mL. Initial SabA crystallization conditions were screened by the sparse-matrix approach using the hanging-drop vapour-diffusion technique. A single hit (100 mM sodium acetate, pH 4.6, 200 mM ammonium acetate, 30 % PEG 4000) was identified after one month. SabA crystallization was optimized by varying pH and PEG concentrations of hit condition, drop size and by microseeding. Typically, using the optimised conditions (200 mM sodium acetate, pH 4.6, 18-20 % PEG 4000), SabA crystals were visible within two days and grew to maximum size in two weeks.

For cryo-protection, SabA crystals were equilibrated in mother liquor supplemented with 30% glycerol or PEG 4000 for at least 30 minutes prior to flash-cooling in liquid nitrogen.

Data collection and processing—X-ray diffraction data for both the native SabA and SeMet-SabA were collected on beamline MX2 at the Australian Synchrotron (Table 1). Diffraction data were indexed, integrated and scaled using HKL2000 (10), MOSFLM (11) and SCALA (12) from the CCP4 package (13).

Structure determination and refinement—SeMet-SabA formed small rod crystals under similar precipitation conditions but using lower protein concentration (3 mg/mL). These cryo-protected crystals belong to the same space group and have similar unit cell parameters to the native SabA crystals. A fluorescence scan was carried out with a SeMet-SabA crystal around the selenium absorption edge to determine the wavelengths corresponding to the peak (0.9690 Å) and infection point (0.9795 Å). Anomalous diffraction data sets were collected with the SeMet-SabA crystals at the peak (2.90 Å) and infection point (3.3 Å) wavelengths (Table 1).

Automated experimental phase calculations were carried out using Auto-Rickshaw (14) (15) by the 2-wavelength anomalous diffraction (2W-MAD) method. A solution was identified in the space group P3121 with one SabA molecule per asymmetric unit (solvent content of 80 %). The calculations identified 7 selenium sites using SHELXD (16). The sites were refined using MLPHARE/PHASER/BP3 (17) (18) (19). The resulting phase were subjected to phase extension and density modification using PIRATE (13) and the automated model building was performed using the program ARP/wARP (20) at 2.9 Å. The initial ARP/wARP model was fragmented and had wrong sequence assignment in the electron density map. Using the calculated positions of selenium and density-modified electron density map, the SabA structure model was built manually, and alternated with refinement cycles using Phenix (21) or REFMAC5 (22). Using the native dataset, the final SabA structure was refined to 2.2 Å with Rfactor and Rfree of 14.0 % and 16.0 %, respectively (Table 1). Final coordinates and structural factors have been deposited with the Protein Data Bank under accession code 4O5J.
Surface plasmon resonance analysis of interaction of SabA with Lewis antigens—Surface plasmon resonance (SPR) experiments were performed using a Biacore T100 biosensor system (GE-healthcare) at 25°C as previously described (23) with the following modifications. All analysis of purified His-tagged SabA N-terminal adhesin domain was performed in phosphate buffered saline (PBS) at pH 6.8 at a flow rate of 30 µl per min. SabA was diluted to 0.1 mg/mL and loaded on Flow Cell 2 of a Ni²⁺ NTA sensor chip with 10 min contact time. Flow Cell 1 was loaded with purified heat denatured (99 °C for 5 mins) SabA diluted to 0.1 mg/mL as reference. Serial dilutions of Lewis antigens (Elicityl SA, Crolles, France) were prepared to 0.00625, 0.0125, 0.025, 0.05, 0.1 mM and run using single-cycle kinetics as follows. After the injection of five glycan dilutions, the chip was regenerated with EDTA. Subsequently, the chip was re-loaded with Ni²⁺ and SabA before the injection of the next Lewis antigen to be tested. A 10 min dissociation time was allowed after the addition of each analyte. SPR signals were analyzed using the Biacore Evaluation software to determine KD.

RESULTS
Structure of SabA adhesin domain—The mature SabA adhesin protein (residues 1-460, lacking the trans-membrane domain) was expressed in E. coli, refolded from inclusion bodies and used in crystallization experiments. The crystal structure of the SabA adhesin was solved to 2.2 Å and reveals a predominantly α-helical molecule shaped rather like a golf putter (Fig. 1B) and comprising a “handle” and a “head” region. All residues are visible except for 26 residues from the N-terminus and 61 residues from the C-terminus. Residues 114 and 182-184, which are located at the tips of two highly mobile loops (as estimated through b-value analysis), are also not visible in electron density.

The handle domain (Fig. 1B, in blue) forms a rigid shaft that comprises the N- and C-terminal helices interacting to form an anti-parallel 2-helix coiled-coil bundle. The C-terminus helix and the disordered linker region (residue 400-460) connect to the predicted β-barrel trans-membrane domain (Fig. 1B).

The central head domain essentially comprises a bundle of α-helices, juxtaposed at approximately right angles to the handle domain. The core region of the head domain is a 4-helix antiparallel coiled-coil bundle (Helix-1-4; a TPR repeat; Fig. 1B in magenta). Except for Helix-1, which is interrupted in the middle by a short loop (Fig. 1B, marked with a *), the other three α-helices are continuous and are approximately 30 residues in length. The 4-helix bundle is a common structural scaffold motif found in a variety of proteins and, accordingly, DALI (24) and 3-D BLAST (25) searches reveal that the head domain is distantly homologous to a number of different proteins, e.g. hydrolases (1oR0; Z-score 8.3) (26), pectin methyl-esterase inhibitor (1x8z; Z-score 7.6) (27) and part of the BRO-1 domain (3um2; Z-score 5.7) (28). Figure 2 shows an example of the superposition between SabA and a pectin methyl-esterase inhibitor.

The connecting region between Helix-1 and -2 comprises 140 residues (Fig. 1B, in gold) and is made up of short α-helices, loops and a small β-sheet (residue 110-112 and 117-119). Predominant features within the Helix-1 / -2 connecting region include a pair of interacting ~20 residue α-helices (residue 144-163 [Helix-1a] and 191-209 [Helix-1b]) that sits on top of Helix-1 and -2 of the core 4-helix bundle. The loops joining Helix-1a/Helix-1b and Helix-1b / Helix-2 are both constrained by disulphide bonds, each of which is formed by a pair of closely spaced cysteines (4-5 amino acids apart; Cys135/Cys141 and Cys173/Cys178). Sequence alignments (Fig. 3) of SabA from different Helicobacter strains reveal both disulphide bonds are highly conserved in the SabA branch of the family. Interestingly, however, Cys141 is absent in the BabA branch of the family (Fig. 3). Instead, in BabA we note a conserved cysteine at position 108 (SabA numbering; Fig. 3, marked with a #); from a structural perspective we suggest this residue would be appropriately positioned to form a disulphide bond with the conserved Cys135.

Sequence / structure insights suggest functionally conserved portions of the SabA structure—We also analysed the structure of the SabA adhesin domain with respect to sequence conservation. The primary sequences of BabA (41 sequences) and SabA (33 sequences) from multiple H. pylori strains were aligned to identify identical residues that might be important for the function of these
adhesins (Fig. 3). We further used the web-based automatic ligand-binding-site prediction program POCASA (29) to predict pockets and cavities in the SabA head domain, with the rationale that these might represent potential ligand binding sites (Fig. 4A-B). Three adjacent pockets stretching from the non-structured tip of the head domain to the groove formed between Helix-1 and -4 were present. Interestingly, many of the identical SabA/BabA surface residues (Fig. 3, highlighted in red and Fig. 4C-D, cyan sticks) lined along the POCASA predicted pockets, suggesting that these cavities may play a role in the binding of host cell ligands. In particular, we noted a deep (~9 Å) positively-charged cavity (Fig. 4E), lined at the bottom by Trp97 and at one side Helix-1a with identical residues Lys152, Gln159 and Gln162 (Fig. 4C and D). In our structure, a glycerol molecule (cryo-protectant) is found in this cavity (Fig. 4E) forming hydrogen bond interactions with the side-chain of Lys152 and Gln159. We further note three other conserved residues – Ser80 and Pro81 map to the interruption in Helix-1 and Gly357 causes a slight kink in Helix-4 (Fig. 4C and D).

The interaction of SabA adhesin domain with binding partners—The SabA crystal structure was solved in the absence of any ligand molecule. Previous experimental studies suggested that the SabA ligands include sialic acid or sialyl-LewisX (5,6). Accordingly, we attempted to co-crystallize or soak-in sialic acid or sialyl-LewisX, however, these experiments were unsuccessful. We then performed surface plasmon resonance (SPR) experiments based on small molecule single cycle kinetics to examine the interaction of SabA with a variety of sugars. SabA bound sialyl-LewisX with a KD of 19.9 µM ±2.7 (Table 2 and Supplementary Fig. 3A). It also bound to LewisX by more than 50 and 20 folds, respectively (Table 2 and Supplementary Fig. 3). In contrast, the substitution of Tyr148, Lys152 or Gln162 by alanine (Y148A, K152A and Q162A, respectively) did not significantly alter the affinity of SabA for sialyl-LewisX although they reduced the binding of SabA to LewisX 3-20 folds (Table 2). Despite the obvious effects of the mutations on ligand binding to SabA, the mutations did not cause any global structural perturbation of SabA: gel filtration profiles (Supplementary Fig. 1) and circular dichroism spectra (Supplementary Fig. 2) of the SabA mutants are highly similar to that of their wild-type counterpart. Thus, we conclude that Gln159 is of critical importance for the binding of SabA to sialyl-LewisX and LewisX. Further, our data suggest that residues Tyr148 and Gln162 are essential only for the binding of SabA to LewisX but not to sialyl-LewisX. The charged side chain of Lys152, which forms a hydrogen bond with the bound glycerol in the crystal structure, surprisingly does not seem to be required for the binding of SabA to either sialyl-LewisX or LewisX.

DISCUSSION

In this study we present the structure of the SabA adhesin domain. These data reveal a “club shaped” molecule that, on first inspection, presents a number of features consistent with a proposed function in interaction with host-cell surface glycoproteins. For example, we note the presence of a highly conserved cavity in the SabA head domain that would provide an attractive binding site for a potential ligand. Although, to
date, and frustratingly, we have been unable to crystallize the SabA adhesin domain in the presence of a bound carbohydrate ligand, we were able to detect the binding of SabA protein to sialyl-Lewis$^X$ ($K_D$ of 19.9 ± 2.7 µM) by small molecule single cycle kinetics SPR, which is a highly sensitive method for analysing molecular interactions. Surprisingly, we also detected interaction between SabA and Lewis$^X$ which occurred with weak affinity ($K_D$ of 50.4 ± 8.2 µM). In contrast, binding to SabA was not detected using Lewis$^A$, Lewis$^B$ and Lewis$^Y$ as the analyte. These results are in agreement with previous reports that the sialic acid moiety is required for optimal binding of SabA to sialyl-Lewis$^X$ and that SabA binding is highly specific to Lewis$^X$ glycans (3,5,6). Additionally, this study shows for the first time that the adhesion domain of SabA in isolation is fully functional and sufficient for selective binding to sialyl-Lewis$^X$ or Lewis$^X$.

In order to understand the molecular basis of SabA ligand binding, we mutated four highly conserved residues selected based on the analysis of the primary and tertiary structures of SabA, and the structure of the predicted ligand-binding cavities. Of the four mutations tested, Q159A has the most significant effect, abolishing SabA binding to both sialyl-Lewis$^X$ and Lewis$^X$. These data suggest that Gln159 plays a critical role in ligand binding. In contrast, while Lys152 is conserved, the K152A mutation does not cause significant changes in the binding affinity to sialyl-Lewis$^X$ or Lewis$^X$.

Substituting Tyr148 and Gln162 with alanine significantly reduced Lewis$^X$-binding (3 to >20-fold), but had little or no effect on sialyl-Lewis$^X$-binding. Tyr148 and Gln162 are found on either end of the putative ligand-binding pocket and form part of the surface of the cavity. The side chains of these residues also form multiple Van der Waals interaction with other parts of the protein (Tyr148 with the loop region; Gln162 with residues in Helix-1 and Helix-4). It is plausible that changing these residues to the smaller alanine may alter the shape of the binding pocket in such a way that binding to the smaller Lewis$^X$ trisaccharide becomes less favourable compared to the binding of the larger sialyl-Lewis$^X$ tetrasaccharide. This remains to be ascertained by further mutagenesis of SabA and structural studies on SabA-Lewis$^X$ and SabA-sialyl Lewis$^X$ complexes.

In summary, our study shows that the N-terminal domain of SabA functions as a sugar-binding adhesion domain, in which a cavity lined by conserved amino acids likely serves as a highly selective ligand-binding site. Notably, this region consists of amino acid residues that are not only conserved among SabA orthologs but also between SabA and BabA. Thus, the findings in this study have significant implications for understanding the structure-function relationship of not only SabA but also that of BabA. It would be of interest to test whether mutating the corresponding residues in BabA will influence the specificity or affinity of Lewis$^B$-binding to BabA.

Previous studies suggested that SabA-mediated binding of *H. pylori* to sialyl glycoconjugates and sialyl gangliosides requires NeuAcα 2-3Gal disaccharide as the minimal binding epitope (5), whereas the results in this study indicate that the adhesion domain of SabA binds to not only sialyl-Lewis$^X$ but also Lewis$^X$. This difference in glycan-binding specificity between the purified SabA adhesion domain and SabA expressed on intact bacteria could be due to differences in steric hindrance, hydrophobicity and/or charges of the local environments of SabA. Alternatively, other domains/regions of SabA may play a role in fine-tuning the glycan-binding specificity. Future investigations are required to further understand the molecular basis of the ligand-binding specificity of SabA and hence its role for pathogenesis in the complex gastric environment during chronic *H. pylori* infection. Such knowledge will be instrumental for the development of new drugs against *H. pylori*-induced chronic gastritis and gastric cancer.
REFERENCES


Crystal structure of SabA adhesin from Helicobacter pylori

for the validation of an X-ray diffraction experiment. Acta crystallographica. Section D, Biological crystallography 61, 449-457


Acknowledgements – JCW is an Australian Research Council Federation Fellow and Honorary National Health and Medical Research Council of Australia Principal Research Fellow. We thank the Australian synchrotron (MX-2 Beamline) and staff for beamline support.

FOOTNOTES

#Joint first authors; these authors contributed equally to the work.
*Whisstock and Kwok are joint senior and corresponding authors – correspondence to James.Whisstock@monash.edu or Terry.Kwok@monash.edu
1Department of Biochemistry and Molecular Biology, 2ARC Centre of Excellence in Structural and Functional Microbial Genomics, 3Department of Microbiology, Monash University, Clayton Campus, Melbourne, VIC, 3800 Australia
4Australian Synchrotron, 800 Blackburn Rd, Clayton, Melbourne, VIC, 3168 Australia
5Institute for Glycomics, Griffith University, Gold Coast Campus, QLD, 4222 Australia
6The abbreviations used are: OMP, outer membrane protein; TPR, tetratricopeptide repeat; SabA, sialic acid binding adhesin; BabA, Lewis\textsuperscript{B} antigen-binding adhesin; SetMet-SabA, Selenomethionine-substituted SabA; 2W-MAD, 2-wavelength anomalous diffraction; SPR, surface plasmon resonance; $K_D$, dissociation constant
FIGURE LEGENDS

Figure 1. A. Schematic illustrating the predicted domain structure of SabA. B. Crystal structure of SabA N-terminal adhesion region. The handle is in blue, the TPR region in magenta and the remainder of the molecule in orange. Broken lines indicate disordered regions. The two conserved disulphide bonds are represented as cyan stick. The short loop interruption in Helix-1 is marked with a *. The C-terminal helix is connected to a trans-membrane - beta-barrel domain, which anchors the SabA adhesion region to the outer membrane of H. pylori.

Figure 2. Structural superposition of SabA (magenta and grey) and pectin methylesterase inhibitor (green) around the TPR portion as calculated by DALI.

Figure 3. Primary sequence alignment of SabA and BabA N-terminal adhesion regions. The alignment was generated with Clustal Omega using 33 SabA and 41 BabA sequences from various H. pylori strains (KEGG database). The figure shows a representative of the complete alignment (5 sequence from each protein). Residues in yellow are identical in all SabA and BabA, red are identical surface residues in the predicted binding pocket, cyan are cysteine residues involved in disulphide bonds in SabA. The position of the alternative conserved cysteine residue in BabA is marked with a #. S80 and P81 (@) map to the interruption in Helix-1 and G357 (§) causes a slight kink in Helix-4. Y148, K152, Q159 and Q162 were mutated to alanine in this study are marked with *.

Figure 4. Ligand binding pockets calculated by an automatic ligand-binding-site prediction program (POCASA). A-B. The binding pockets (dark grey surface) are predicted mostly towards to the tip of the SabA head domain. C-D. Viewing from the top towards the predicted ligand binding pockets (removed in D for clarity). A number of highly conserved residues (cyan sticks), in particular S80, P81, W97, Y148, K152, Q159 and Q162, line the surface of these pockets. Y148, K152, Q159 and Q162 were mutated to alanine in this study are marked with *. A glycerol molecule (pink sticks) is observed in the crystal structure and forms hydrogen bond interactions with K152 and Q159. E. Electrostatic potential surface representation of the SabA head domain showing the predicted pockets (grey dots) generated using CCP4mg (30). A glycerol molecule (pink stick) binds in a positively-charged cavity.
Table 1. SabA X-ray data collection and structure refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>SetMet derivative</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (K)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SetMet derivative</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Space group</td>
<td>P3;21</td>
<td>P3;21</td>
<td>P3;21</td>
</tr>
<tr>
<td>Unit cell parameters (Å)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a=b</td>
<td>179.08</td>
<td>178.90</td>
<td>180.04</td>
</tr>
<tr>
<td>c</td>
<td>64.53</td>
<td>64.42</td>
<td>64.47</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9690</td>
<td>0.9795</td>
<td>0.9537</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>89.5-2.9(3.06-2.9)</td>
<td>89.5-3.3(3.4-3.3)</td>
<td>90.0-2.2(2.3-2.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100(100)</td>
<td>100(100)</td>
<td>100(100)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>239968(35408)</td>
<td>164853(24138)</td>
<td>556070(81900)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>26604(3843)</td>
<td>18899(2707)</td>
<td>60966(8825)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>9.0(9.2)</td>
<td>8.7(8.9)</td>
<td>9.1(9.3)</td>
</tr>
<tr>
<td>R_pim (%)</td>
<td>6.7(33.7)</td>
<td>9.7(42.0)</td>
<td>4.4(35.7)</td>
</tr>
<tr>
<td>R_meas (%)</td>
<td>20.2(103.4)</td>
<td>29.1(128.0)</td>
<td>13.4(109.1)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>10.0(2.7)</td>
<td>8.7(2.5)</td>
<td>14.2(2.4)</td>
</tr>
<tr>
<td>Mosiacity (°)</td>
<td>0.54</td>
<td>0.60</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Refinement

Non-hydrogen atoms          3011
Protein                     2838
Water                       159
Resolution, Å               90.0-2.2
R_factor, %                 14.0
R_free, %                   16.0

Rms deviations from ideality

Bond lengths, Å             0.02
Bond angles, Å              1.53
Chirality, °                0.09
Planarity, °                0.007
Dihedrals, °                14.4

Ramachandran plot

Preferred regions, %        96.4
Allowed regions, %          2.8
Disallowed regions, %       0.8

B factors, Å²

Average main chain          55.037
Average side chain and water 59.137
Average all atoms           57.122
Table 2. Affinity of SabA for Lewis antigens determined by surface plasmon resonance

<table>
<thead>
<tr>
<th></th>
<th>Sialyl LewisX</th>
<th>LewisX</th>
<th>(K_D) determined (^{b,c})</th>
</tr>
</thead>
<tbody>
<tr>
<td>SabA wild-type</td>
<td>19.9 (\mu)M ± 2.7</td>
<td>50.4 (\mu)M ± 8.2</td>
<td></td>
</tr>
<tr>
<td>SabA Y148A</td>
<td>21.1 (\mu)M ± 5.6</td>
<td>151.7 (\mu)M ± 30.5</td>
<td></td>
</tr>
<tr>
<td>SabA K152A</td>
<td>24.5 (\mu)M ± 2.1</td>
<td>62.8 (\mu)M ± 13.6</td>
<td></td>
</tr>
<tr>
<td>SabA Q159A</td>
<td>NCDI(^a)</td>
<td>NCDI(^a)</td>
<td></td>
</tr>
<tr>
<td>SabA Q162A</td>
<td>31.7 (\mu)M ± 7.1</td>
<td>NCDI(^a)</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\): NCDI: No concentration-dependent interaction was detected (\(K_D > 1\)mM).

\(^{b}\): No interaction was observed between SabA and SabA mutants with Lewis\(^{A}\), Lewis\(^{B}\) and Lewis\(^{Y}\).

\(^{c}\): Data shown are mean \(K_D\) values determined from four independent experiments.
Figure 1

A

Signal peptide

Extracellular SabA adhesion region

β-barrel transmembrane domain

B

Head domain

Handle domain

Extracellular

H. pylori outer membrane

β-barrel domain

Periplasmic space
Crystal structure of SabA adhesin from Helicobacter pylori

Figure 2
Crystal structure of SabA adhesin from Helicobacter pylori

Figure 4

A

B

180°

Top view

C D

E
The three-dimensional structure of the extracellular adhesion domain of the sialic acid-binding adhesin SabA from Helicobacter pylori
Siew Siew Pang, Stanley Thai Son Nguyen, Andrew J. Perry, Christopher J. Day, Santosh Panjikar, Joe Tiralongo, James C. Whisstock and Terry Kwok

J. Biol. Chem. published online December 27, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.513135

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/12/27/M113.513135.DC1