Communication

Crystallisation and Preliminary Crystallographic Analysis of Helicobacter pylori Periplasmic Binding Protein YckK

Mohammad Mizanur Rahman 1,2,†, Daniel Phillip Germantsis 1,2,†, Mayra Alejandra Machuca 1,2, Abu Iftiaf Salah Ud-Din 1,2, and Anna Roujeinikova 1,2,3,*

1 Department of Microbiology, Monash University, Clayton, VIC 3800, Australia; mohammad.mizanur.rahman@monash.edu (M.M.R); dpger2@student.monash.edu (D.P.G.); mayra.machuca.perez@monash.edu (M.A.M.); abu.ud-din@monash.edu (A.I.S.U.-D.)
2 Infection and Immunity Program, Monash Biomedicine Discovery Institute, Clayton, VIC 3800, Australia
3 Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia
* Correspondence: anna.roujeinikova@monash.edu; Tel.: +61-39-902-9194
† These authors contributed equally to this work.

Academic Editor: Albert Guskov
Received: 7 September 2017; Accepted: 26 October 2017; Published: 29 October 2017

Abstract: Helicobacter pylori infection can lead to the development of gastric and duodenal ulcers and gastric cancer. In recent years, the efficacy of the standard therapy has been falling, necessitating ongoing efforts to identify new drug targets. Due to their important role in chemotaxis and nutrient uptake, periplasmic binding proteins (PBPs) represent potential targets for new antimicrobial agents that have not yet been fully explored and exploited. The H. pylori PBP YckK is homologous to polar amino acid-binding proteins from other bacteria. The yckK gene overlaps the gene tcyB—a gene annotated as a polar amino acid-transporting permease. Purified recombinant YckK behaved as a monomer in solution. Crystals of YckK were grown by the hanging drop vapour diffusion method using PEG 3350 as the precipitating agent. The crystals belong to the primitive triclinic space group P 1 with unit cell parameters a = 63.0, b = 63.5, c = 74.6 Å, α = 72.5, β = 68.3, γ = 69.4°. X-ray diffraction data were collected to 1.8 Å resolution using synchrotron radiation. Molecular replacement using this data revealed that the asymmetric unit contains three subunits: two in the open and one in the closed conformation.

Keywords: Helicobacter pylori; periplasmic binding protein; protein crystallisation

1. Introduction

Helicobacter pylori is a Gram-negative spiral-shaped bacterium that possesses a number of flagella located at the cell pole. The motility conferred by these flagella helps H. pylori to colonise the gastric mucosa and escape the highly acidic environment of the gastric lumen [1]. H. pylori infection is associated with chronic gastritis, and is an established risk factor for the development of gastric and duodenal ulcers. Infection is also associated with the development of more severe conditions such as MALT-lymphoma and gastric cancer [2]. In recent years, the efficacy of the standard triple therapy comprising one proton pump inhibitor, clarithromycin, and amoxicillin or metronidazole has been falling, largely due to increasing antibiotic resistance [3]. Although bismuth quadruple therapy can be used as an alternative in cases where standard treatment fails, this therapy still relies on the use of antibiotics to which resistance is growing, such as metronidazole [3]. Therefore, ongoing efforts to identify new drug targets for antibiotic development are needed.

Periplasmic binding proteins (PBPs), also known as solute-binding proteins, constitute a large family of ligand-binding proteins that serve as primary receptors in nutrient uptake and chemotaxis [4].
They consist of two globular domains separated by a flexible hinge region. Binding of a ligand in the cleft between the two domains results in its closure in a manner that was described as a “Venus flytrap” model [5]. PBPs bind a wide variety of ligands including sugars, amino acids, di- and oligopeptides, and tetrahedral oxyanions [4]. While many PBPs form components of ABC importers, some are associated with secondary transporters such as tripartite ATP-independent periplasmic (TRAP) transporters [6]. In addition to their role in active nutrient uptake, some PBPs (such as maltose-binding protein, for example [4]) function as primary receptors in chemotaxis, where upon binding to a ligand, they interact with methyl-accepting chemotaxis proteins (MCPs), generating signal. On account of their important role in chemotaxis and nutrient uptake, PBPs represent potential targets for new antimicrobial agents that have not yet been fully explored and exploited.

The full genome sequences of different *H. pylori* strains revealed the presence of six to seven PBPs, one of which has been annotated YckK (the solute-binding protein of a polar amino acid ABC transporter, ~260 amino acids). An amino acid sequence similarity search in the RCSB Protein Data Bank with the sequence of YckK from *H. pylori* strain SS1 [7] revealed 54% overall sequence identity to a cysteine-binding protein from *Neisseria gonorrhoea* [8] (PDB ID 2YLN). In all sequenced *H. pylori* strains, the 3’ end of the yckK gene overlaps with the 5’ end of a gene on the same strand, annotated tcyB, a polar amino acid transport system permease, indicating that yckK and tcyB are co-transcribed. Furthermore, the YckK sequence shows 31% identity with that of a synthetic ancestral polar amino acid-binding protein (PDB ID 4ZV1) that binds L-arginine, L-histidine, L-lysine, and L-glutamine [9]. In addition, it shares 30% sequence identity with the arginine-binding protein from *Thermotoga maritima* [10] (PDB ID 4PRS) and 26% sequence identity with the basic amino acid binding protein from *Thermus thermophilus* [11] (PDB ID 3VV5). This suggests that *H. pylori* YckK may serve to deliver cystine and/or polar amino acids to the permease encoded by tcyB. This paper describes the cloning, purification, crystallisation, and preliminary X-ray diffraction analysis of recombinant YckK from *H. pylori* SS1, which is an important step towards elucidation of its ligand specificity and three-dimensional structure.

2. Materials and Methods

2.1. Gene Cloning and Overexpression

The gene encoding YckK from *H. pylori* strain SS1 (GenBank ID AQM65832.1), lacking the signal peptide (residues 1–19), was synthesised and cloned into the pet151/D-TOPO expression vector (Invitrogen, Waltham, MA, USA) by GenScript. This vector contains an N-terminal His₆ tag upstream of a tobacco etch virus (TEV) protease cleavage site. Recombinant protein used for crystallisation contained residues 20–256 of YckK preceded by six residues (GIDPFT) from the TEV protease cleavage site as a cloning artefact. The vector was transformed into *E. coli* strain BL21 (Novagen). Cells were grown in LB medium containing 50 µg/mL ampicillin at 310 K to an OD₆₀₀ of 0.6–1.0. Expression of YckK was then induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside, and growth was continued for a further 4 h. Cells were harvested by centrifugation at 6000 × g for 15 min at 277 K.

2.2. Purification

Protein was purified by following the procedure described in [12]. Briefly, *E. coli* cells were resuspended in buffer A (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM phenylmethanesulfonyl fluoride) and lysed using a C5 cell disruption system (Avestin, Ottawa, Canada). Cell lysate was then centrifuged at 10,000 × g for 15 min at 277 K to pellet debris. NaCl and imidazole were added to the supernatant to final concentrations of 500 mM and 20 mM, respectively. The sample was loaded onto a 5 mL Ni-NTA affinity column (GE Healthcare, Chicago, IL, USA) pre-equilibrated in buffer B (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole), washed with the same buffer and eluted with buffer C (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole). The hexahistidine tag was either left on or cleaved off using TEV protease (Invitrogen, Waltham, MA, USA) during overnight dialysis at
277 K against buffer D (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT, 1% (v/v) glycerol). NaCl and imidazole were added to the sample to final concentrations of 500 mM and 20 mM, respectively. Cleaved protein was separated from the hexahistidine tag, TEV, and any uncleaved protein by passing the sample through the Ni-NTA affinity column. The final step in YckK purification was gel filtration on a Superdex 200 HiLoad 26/60 size-exclusion column (GE Healthcare, Chicago, IL, USA) equilibrated with buffer E (20 mM Tris pH 8.0, 150 mM NaCl). The Bradford assay [13] was used to determine YckK concentration. Protein purity was assessed to be >95% using SDS-PAGE (Figure 1).

![Coomassie Blue-stained 15% SDS-PAGE gel of recombinant YckK. Left to right: molecular weight markers, (1) Ni-NTA purified YckK<sub>tag</sub>, (2) Ni-NTA purified YckK, and (3) YckK after size-exclusion chromatography.](image)

**Figure 1.** Coomassie Blue-stained 15% SDS-PAGE gel of recombinant YckK. Left to right: molecular weight markers, (1) Ni-NTA purified YckK<sub>tag</sub>, (2) Ni-NTA purified YckK, and (3) YckK after size-exclusion chromatography.

### 2.3. Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS) Analysis

A 50 µL sample of the 3 mg/mL YckK<sub>tag</sub> solution in buffer E was loaded onto a Superdex 200 5/150 HiLoad size-exclusion column (GE Healthcare) pre-equilibrated with buffer E flowing at 0.15 mL/min. The eluate was passed through an in-line DAWN HELEOS light scattering detector, an Optilab T-rEX differential refractive index detector and a quasi-elastic light scattering detector (WyattQELS, Wyatt Technology Corporation, Santa Barbara, CA, USA). Normalisation was performed against bovine serum albumin (BSA), and data collection and analysis was carried out with ASTRA6 (Wyatt Technology Corporation). The results are presented in Table 1 and Figure 2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Polydispersity</th>
<th>Molecular Weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YckK&lt;sub&gt;tag&lt;/sub&gt;</td>
<td>1.0</td>
<td>29.4</td>
</tr>
<tr>
<td>BSA</td>
<td>1.0</td>
<td>63.2</td>
</tr>
</tbody>
</table>
Figure 2. Size-exclusion chromatography trace and molecular weight (MW) estimation of YckKtag. The peak shows the UV trace for fractions containing protein. The UV absorbance values, normalized to 0–100% scale, are shown along the left-hand y-axis. A line superimposed on the peak indicates the light-scattering-derived molecular weight (MW) of the eluted particles with the MW values shown on the right-hand y-axis.

2.4. Crystallisation, Preliminary X-ray Diffraction Data Collection, and Molecular Replacement

YckK was concentrated to 29 mg/mL using an Amicon Ultracel 10K cut-off concentrator (Merck Millipore, Darmstadt, Germany). Preliminary screening for crystallisation conditions was performed by the sitting drop vapour diffusion method using a Phoenix crystallisation robot (Art Robbins instruments, Sunnyvale, CA, USA) and commercial screens Crystal Screen HT and PEG/Ion HT (Hampton Research, Aliso Viejo, CA, USA). Crystallisation droplets contained 100 nL of protein solution mixed with 100 nL reservoir solution and equilibrated against 50 µL reservoir solution in a 96-well Art Robins CrystalMation Intelli-Plate (Hampton Research). Crystals formed after 1 day in the condition No 95 of the PEG/Ion HT screen, containing 1% (w/v) tryptone, 50 mM HEPES pH 7.0, and 12% (w/v) PEG 3350. Manual refinement resulted in an optimised crystallisation condition of 1% tryptone, 50 mM HEPES pH 6.0, 20% PEG 3350, 29 mg/mL YckK. For data collection at cryogenic temperatures, the crystal was briefly soaked in a cryoprotecting solution containing 1.2% tryptone, 60 mM HEPES pH 6.0, 24% PEG 3350, and 25% glycerol before being flash-cooled by plunging it in liquid nitrogen. X-ray diffraction data were collected at 100 K using an ADSC Quantum 210r detector (Area Detector Systems Corporation, Poway, CA, USA) at the MX1 beamline of the Australian Synchrotron; 360 images were taken using 0.5° oscillations. The data were integrated using iMosflm [14] and scaled using ALMLESS [15] from the CCP4 suite [16]. Statistics are presented in Table 2. Molecular replacement was performed using Phaser [17] with the structure of the cysteine-binding protein from N. gonorrhoea [8] as the search model. Model refinement was performed using Phenix [18].
Table 2. Data collection and processing statistics. Numbers in parentheses indicate values for the highest resolution shell.

<table>
<thead>
<tr>
<th></th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mosaicity (°)</td>
<td>0.6</td>
</tr>
<tr>
<td>No. of crystals</td>
<td>1</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
</tr>
<tr>
<td>No. of images</td>
<td>360</td>
</tr>
<tr>
<td>Rotation per image (°)</td>
<td>0.5</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.95</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>49.3–1.80 (1.83–1.80)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95 (95)</td>
</tr>
<tr>
<td>Observed reflections</td>
<td>171,880 (8674)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>86,970 (4372)</td>
</tr>
<tr>
<td>(\text{Mean } I/\sigma(I))</td>
<td>8.9 (1.8)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>2.0 (2.0)</td>
</tr>
<tr>
<td>(R_\text{merge})</td>
<td>0.057 (0.327)</td>
</tr>
<tr>
<td>(CC_{1/2}) (%)</td>
<td>99 (73)</td>
</tr>
</tbody>
</table>

3. Results

3.1. Overexpression and Purification

Recombinant *H. pylori* SS1 YckK lacking the periplasm-targeting peptide (residues 1–19) was expressed in *E. coli* BL21 cells transformed with the pet151/D-TOPO expression vector. Electrophoretic homogeneity of the purified protein was more than 95% as determined by Coomassie Blue staining of an SDS-PAGE gel (Figure 1). The recombinant protein contained residues 20–256 plus six N-terminal amino acid residues from the TEV cleavage site (GIDPFT). To determine the oligomeric solution state and sample monodispersity, we carried out SEC-MALS analysis of YckK\textsubscript{tag}. The protein eluted as a single roughly symmetrical peak (Figure 2) with a polydispersity index value of 1.0, indicating that the sample was homogenous with respect to the molar mass. The molecular weight of YckK\textsubscript{tag} as determined by SEC-MALS analysis was 29.4 kDa. This value is close to the molecular weight calculated from the sequence of YckK\textsubscript{tag} (29.5 kDa), which suggests that YckK is monomeric in solution.

3.2. Crystallisation and Preliminary Crystallographic Analysis

Crystals of YckK (Figure 3) were obtained using a sparse matrix crystallisation approach. Data were collected from a single cryo-cooled crystal using beamline MX1 at the Australian Synchrotron (AS) to a resolution of 1.8 Å. Analysis of the data using the autoindexing routine in *iMosflm* indicated that the crystal belonged to the primitive triclinic space group *P*1 with unit cell parameters \(a = 63.0\), \(b = 63.5\), \(c = 74.6\) Å, \(\alpha = 72.5\), \(\beta = 68.3\), \(\gamma = 69.4°\). Calculation of the Matthews coefficient for three molecules in the asymmetric unit gave a value of 3.16 Å\(^3\) Da\(^{-1}\), corresponding to a solvent content of 61%. This value falls within the range observed for protein crystals [19]. A molecular replacement search was performed with *Phaser* using data at a resolution range of 20.0–2.5 Å. The monomers of the homologous protein from *N. gonorrhoeae* [8] in the closed (PDB ID: 2YLN) and open (PDB ID: 3ZSF) conformation were used as search models, and the protocol parameters were set to allow a search for a mix of the two. The molecular replacement (MR) solution contained three subunits: two in an open and one in the closed conformation. The log likelihood gain function of the solution was 649. After one round of XYZ, B refinement using the simulated annealing protocol in *Phenix*, the R factor value fell from 0.491 to 0.417 and Rfree fell from 0.494 to 0.447. Inspection of the electron density maps revealed good agreement between the MR solution and the density for the entire length of all three subunits. A representative part of the electron density map of the refined model is shown in Figure 4. Model building in *Coot* [20] and refinement in *Phenix* are in progress.
Experimental evidence for amino acid utilisation was provided by Stark et al. [27], who found that a mix of monomers, dimers, trimers, and tetramers in a concentration-dependent manner, with higher concentrations of protein favouring the formation of oligomers [25].

Streptococcus pneumoniae in the uptake of carbohydrates in a dimer [10]. The monomeric state of YckK also differs from the oligomeric state of the PBP involved in the uptake of carbohydrates in Thermotoga maritima [23], and Tp32, the methionine binding PBP from the archaeon Methanosarcina acetivorans [22], OppA, the oligopeptide binding PBP from Thermotoga maritima [23], and Tp32, the methionine binding PBP from Treponema pallidum [24]. In contrast, the T. maritima arginine binding protein, with which YckK shares 30% sequence identity, has been found to behave as a dimer [10]. The monomeric state of YckK also differs from the oligomeric state of the PBP involved in the uptake of carbohydrates in Streptococcus pneumoniae (SP0092), which has been shown to exist as a mix of monomers, dimers, trimers, and tetramers in a concentration-dependent manner, with higher concentrations of protein favouring the formation of oligomers [25].

Homology searches suggest that YckK is likely a polar amino acid binding protein that interacts with its transcriptionally linked importer, TcyB. Amino acid utilisation in H. pylori has been investigated.

**Figure 3.** Crystals of *H. pylori* SS1 YckK. Scale bar corresponds to 1 mm.

**Figure 4.** A representative part of the electron density map of the molecular replacement solution after one round of XYZ, B refinement using the simulated annealing protocol in Phenix. The contour level is 1 $\sigma$.

4. Discussion and Conclusions

SEC-MALS experiments revealed that the isolated recombinant *H. pylori* YckK behaves in solution as a monomer. This is in agreement with the observation that many other PBPs are monomeric in solution at physiologically relevant concentrations, including SitA, a metal binding PBP from *Staphylococcus pseudintermedius* [21], MaModA, the molybadate binding PBP of the archaeon *Methanosarcina acetivorans* [22], OppA, the oligopeptide binding PBP from *Thermotoga maritima* [23], and Tp32, the methionine binding PBP from *Treponema pallidum* [24]. In contrast, the *T. maritima* arginine binding protein, with which YckK shares 30% sequence identity, has been found to behave as a dimer [10]. The monomeric state of YckK also differs from the oligomeric state of the PBP involved in the uptake of carbohydrates in *Streptococcus pneumoniae* (SP0092), which has been shown to exist as a mix of monomers, dimers, trimers, and tetramers in a concentration-dependent manner, with higher concentrations of protein favouring the formation of oligomers [25].

Homology searches suggest that YckK is likely a polar amino acid binding protein that interacts with its transcriptionally linked importer, TcyB. Amino acid utilisation in *H. pylori* has been investigated...
by Schilling et al. [26], who used in silico modelling to predict the minimal media requirements of *H. pylori* strain 26695. The amino acids that were suggested to be required for growth were alanine, arginine, histidine, isoleucine, leucine, methionine, phenylalanine, valine, and cysteine. Experimental evidence for amino acid utilisation was provided by Stark et al. [27], who found that a continuous culture of *H. pylori* strain NCTC 11637 utilised all amino acids in a nutritionally defined medium with the exception of lysine, threonine and glycine. Reynolds and Penn [28] have found that *H. pylori* strain NCTC 11637 and an additional nine clinical isolates required arginine, histidine, isoleucine, leucine, methionine, phenylalanine and valine for growth in a defined medium, while alanine and serine were also required by most strains. These studies suggested that while amino acid requirement for *H. pylori* growth is strain-specific, all strains require at least five amino acids with a non-polar side chain (isoleucine, leucine, methionine, phenylalanine, and valine) and at least two amino acids with a polar side chain (arginine and histidine). It is possible that the YckK/TcyB system serves to import polar amino acids from the external environment.

A number of amino acids have been found to act as chemoattractants of *H. pylori*. Worku et al. [29] reported that alanine, arginine, asparagine, glutamine, glycine, histidine, leucine, proline, tyrosine, and valine acted as chemoattractants of *H. pylori* isolates obtained from duodenal ulcer or dyspepsia patients. Serine has also been shown to behave as a chemoattractant for strains 26695 and 43504, and there is evidence that strain 26695 recognises arginine as a chemoattractant [30]. In contrast to the clinical isolates studied by Worku et al. [29], Cerda et al. [30] have found no chemotactic response to alanine, histidine, or leucine in strains 26695 and 43504. Cysteine, glutamate, methionine and tryptophan also failed to elicit a chemotactic response. Furthermore, while Worku et al. [29] found glutamic and aspartic acids to behave as weak chemorepellents, Cerda et al. [30] reported that aspartic acid behaves as a chemoattractant of both *H. pylori* 26695 and 43504. Given the role of polar amino acids as chemotactic signals, it would be important to establish in the future if YckK plays a role in chemotaxis, and if so, which methyl-accepting chemotaxis protein it signals through.

Although the natural ligand of YckK is not yet known, the procedure for the production of the recombinant protein detailed here will allow the use of high-throughput screening methods to determine its ligand specificity. One such approach is a fluorescence-based thermal shift assay described by McKellar et al. [31]. This method involves detecting an increase in the heat-induced unfolding of a protein upon ligand binding. The technique uses a fluorescent dye that is quenched by water in a solution of folded protein. As the temperature of the solution increases, heat-induced unfolding of the protein results in non-specific binding of the dye to exposed hydrophobic regions of the protein. The dye is subsequently unquenched, and an increase in fluorescence is observed. The melting temperatures of the free protein and the protein in the presence of a potential ligand can then be compared to determine whether ligand binding has occurred. McKellar et al. [31] have demonstrated the ability of this simple and efficient technique to rapidly identify protein ligands using small amounts of a purified recombinant chemoreceptor sensing domain, a fluorescent dye, a commercially available compound screen, and an RT-PCR machine. Following ligand identification, the ability to crystallise YckK will allow determination of the structural basis of ligand recognition through the analysis of the ligand-bound structure. The fact that YckK in both the open and closed conformations was found within the same asymmetric unit is consistent with the Venus flytrap model of ligand binding, in which the two lobes of the protein close in around the ligand. Structure determination using the available X-ray diffraction data will allow a detailed analysis of the conformational transition of YckK. Furthermore, insight into the ligand binding mechanism of recombinant YckK provided by the determination of its crystal structure in complex with the ligand can be used in the development of therapeutic small molecule inhibitors that blocks nutrient uptake.

**Acknowledgments:** We thank Danuta Maksel and Robyn Gray at the Monash Macromolecular Crystallisation Facility for their assistance in setting up automated crystallisation screens. Part of this research was undertaken on the MX1 beamline of the AS, Victoria, Australia. We thank the AS staff for their assistance with data collection. Mayra A. Machuca is indebted to the Departamento Administrativo de Ciencia, Tecnología e Innovación COLCIENCIAS for a doctoral scholarship.
Author Contributions: All authors conceived and designed the experiments. Mohammad Mizanur Rahman, Daniel Phillip Germantsis, Mayra Alejandra Machuca and Anna Roujeinikova performed the experiments. All authors analysed the data. Daniel Phillip Germantsis and Mayra Alejandra Machuca drafted the manuscript. Anna Roujeinikova edited the manuscript. All authors reviewed and approved the final version.

Conflicts of Interest: The authors declare no conflict of interest.

References


