Preliminary X-ray crystallographic studies on the *Helicobacter pylori* ABC transporter glutamine-binding protein GlnH

Mohammad M. Rahman[1], Mayra A. Machuca[1], Anna Roujeinikova[1,2,*]

[1] *Infection and Immunity Program, Monash Biomedicine Discovery Institute, Department of Microbiology, Monash University, Clayton, Victoria, Australia;*
[2] *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia.*

Summary	Periplasmic binding proteins (PBPs) of Gram-negative bacteria sense essential nutrients and mediate their uptake by ATP-binding cassette (ABC) transporters. The gene for a PBP of *H. pylori* SS1, annotated as GlnH, is located within the *glnPQH* operon encoding an ABC importer system. In this study, GlnH has been expressed in *E. coli* and purified to > 98% homogeneity. The recombinant protein was folded according to the circular dichroism (CD) analysis and behaved as a monomer in solution. Crystals of GlnH have been grown by the hanging-drop vapour-diffusion method using polyethylene glycol (PEG) 4000 as a precipitating agent. The crystals belonged to the primitive monoclinic space group $P2_1$ with unit cell parameters a = 38.67, b = 93.36, c = 64.13 Å, β = 93.72°. A complete X-ray diffraction data set was collected to 1.3 Å resolution from a single crystal using synchrotron radiation. Molecular replacement using this data revealed that the asymmetric unit contains a single molecule. This is a key step towards elucidation of the structural basis of the GlnH function.

Keywords: *Helicobacter pylori*, glutamine-binding protein, circular dichroism, protein crystallization, GlnH

1. Introduction

A significant portion of the world's population is infected with *Helicobacter pylori* (*1*), a motile, flagellated bacterium that colonizes the human gastric mucosa and is associated with chronic gastritis, gastric and duodenal ulcers (*2*). *H. pylori* infection has also been linked to an increased person's risk of developing mucosa-associated lymphoid tissue (MALT)-lymphoma (*3*), type 2 diabetes (*4*) and gastric adenocarcinoma (*5*). Currently, no vaccine is available against this pathogen; symptomatic patients are treated using a combination of at least two antibiotics plus a proton pump inhibitor/ bismuth citrate. In recent years, eradication rates have been falling due to the spread of resistance to some antibiotics (*6,7*). Therefore, there is an urgent need for identification of new protein targets for antibiotic development.

H. pylori takes up essential nutrients, such as amino acids, di- and oligopeptides and metal ions, from host cells (*6-8*). It scavenges the nutrients using various transport proteins including ATP-binding cassette (ABC) transporters (*9-11*). A canonical ABC transporter consists of two trans-membrane domains, two nucleotide-binding domains and a periplasmic substrate-binding protein (PBP), sometimes termed solute-binding protein (SBP) (*12*). PBPs bind their substrates at the periplasmic space with high specificity and then deliver them to the cognate trans-membrane domains located in the cytoplasmic membrane. Although PBPs are very diverse in terms of the their amino acid sequence, their structural fold is highly conserved (*13*). The structure of a typical PBP comprises two globular lobes joined by a flexible hinge region. Many PBPs operate *via* the so-called "Venus flytrap" mechanism, where ligand binding at the interface results in a large conformational change due to the closure of the two lobes around the ligand molecule (*14*). In addition to their role in nutrient sensing, some PBPs mediate chemotaxis towards the nutrient source (*15,16*). Furthermore, it has been shown that PBPs play

Address correspondence to:
Dr. Anna Roujeinikova, Infection and Immunity Program, Monash Biomedicine Discovery Institute, Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia; Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia.
E-mail: anna.roujeinikova@monash.edu

roles in bacterial virulence, and may, for some bacteria, be immunogenic, raising the notion that PBPs may be targeted for the development of antibacterial vaccines, therapies and new drugs (*14,17*).

Analysis of the whole genome sequences of different *H. pylori* strains suggested that this pathogen possesses seven to eight PBPs, of which one has been annotated as the ABC transporter glutamine-binding protein precursor, GlnH (*18-21*). In the genomic context, *glnH* is a part of the *glnPQH* operon, where *glnP* codes for an ABC transporter permease, and *glnQ* for an ATP binding subunit. Apart from this genome sequence-derived information, no studies have been published yet on *H. pylori* GlnH, and its natural ligand is not yet known. *H. pylori* GlnH shares only modest (24%) sequence identity with a well-characterized glutamine-binding protein from *E. coli* (GlnBP), which binds L-glutamine with a K_d of 0.5 μM (*22-25*). Interestingly, it shares 56% and 51% overall sequence identity with the cysteine-binding proteins Cj0982 from C. *jejuni* (*26*) and Ngo2014 from *N. gonorrhoea* (*27*), respectively. This suggests that GlnH may serve to deliver glutamine and/or cysteine to the permease encoded by *glnP*. In order to understand how GlnH interacts with its putative ligand and to elucidate its three-dimensional structure, we have initiated an X-ray crystallographic study of GlnH. This paper describes the cloning, purification, crystallization and the preliminary X-ray crystallographic analysis of recombinant GlnH from *H. pylori* strain SS1.

2. Materials and Methods

2.1. *Gene cloning and overexpression*

The signal sequence of *H. pylori* GlnH (NCBI ID WP_077232337.1) was predicted by using the online tools PrediSi (*28*) and Phobius (*29*). The coding sequence, lacking the signal peptide (residues 1-24), was codon-optimized, synthesized and ligated into the pet151/D-TOPO expression vector (Invitrogen, Waltham, MA, USA) by GenScript (Piscataway, USA). This vector harbors an N-terminal hexa-histidine tag followed by a V5 epitope and a tobacco etch virus (TEV) protease cleavage site (HHHHHHGKPIPNPL LGLDSTENLYFQGIDPFT). The recombinant protein used for thermal melts and crystallization comprised residues 25-277 of GlnH plus additional GIDPFT residues from the TEV cleavage site and the vector as a cloning artifact. The vector was transformed into *E. coli* BL21 DE3 (Novagen, Darmstadt, Germany) and cells were cultured at 37°C in Luria-Bertani (LB) medium supplemented with 50 μg/mL ampicillin (G-Bioscience, MO, USA). Overexpression of GlnH was induced with 1 mM of isopropyl β-*D*-1-thiogalactopyranoside (IPTG) (Astral Scientific, NSW, Australia) at an OD_{600} of 0.60, and growth was continued for a further 4 h. Cells were

then harvested by centrifugation at 4,800 × g for 15 min at 4°C.

2.2. *Purification*

Protein was purified by following the procedure described in (*30*). Briefly, harvested cells were resuspended in a buffer A (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1mM phenylmethylsulfonyl fluoride (Merck, Darmstadt, Germany)) and lysed using an EmulsiFlex-C5 high-pressure homogenizer (Avestin, Ottawa, Canada). Cell debris was then removed by centrifugation at 10,000× g for 15 min at 4°C. Imidazole (Sigma-Aldrich, MO, USA) and NaCl were added to the supernatant to final concentrations of 20 mM and 500 mM, respectively. The protein sample was loaded onto a 5 mL Ni-NTA affinity column (GE Healthcare, Chicago, IL, USA) pre-equilibrated with buffer B (20 mM Tris-HCl pH 8.0, 500 mM NaCl and 20 mM imidazole), washed with the same buffer and eluted with buffer C (20 mM Tris-HCl pH 8.0, 500 mM NaCl and 500 mM imidazole). The N-terminal His₆ tag was either left on, or cleaved off using TEV protease (Invitrogen, Waltham, MA, USA) during overnight dialysis against buffer D (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DL-Dithiothreitol (DTT) (Sigma-Aldrich, MO, USA) and 1% v/v glycerol (Astral Scientific, NSW, Australia)) at 4°C. Following the incubation with TEV, NaCl and imidazole were added to the sample to final concentrations of 500 mM and 20 mM, respectively, and the cleaved tag, TEV protease and any uncleaved protein were removed by passing the sample through the Ni-NTA affinity column. The final stage in the purification of both tagged and untagged versions of the protein was gel filtration using the Superdex 200 HiLoad 26/60 column (GE Healthcare, Chicago, IL, USA) equilibrated with buffer E (20 mM Tris-HCl pH 8.0, 150 mM NaCl), at a flow rate of 4 mL/min. The peak fractions of the eluate were pooled, protein concentration was determined using the Bradford assay (*31*) and purity was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.3. *Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis*

SEC-MALS analysis involved an injection of 50 μL of the 100 μM GlnH-His₆ solution in buffer E onto a Superdex 200 5/150 HiLoad size-exclusion column (GE Healthcare, Chicago, IL, USA) pre-equilibrated with buffer E flowing at 0.15 mL/min. The eluate was passed through a Shimadzu HPLC 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). The MALS

detectors were normalized against 100 μM of bovine serum albumin (BSA). Data collection and analysis was performed with ASTRA6 (Wyatt Technology Corporation).

2.4. *Circular dichroism (CD) analysis*

The purified GlnH-His$_6$ was buffer-exchanged into 10 mM sodium phosphate pH 7.4. CD spectra were recorded using a J-815 spectropolarimeter (JASCO, Tokyo, Japan) at a protein concentration of 0.24 mg/mL in 2-mm path length cuvette (Starna Pty Ltd, NSW, Australia) at 25°C over a wavelength range of 200-250 nm with a scan rate of 20 nm/min. Spectra were recorded in triplicate and averaged. The secondary structure content was estimated by fitting the ellipticity data using the BeStSel server (*32*).

2.5. *Protein buffer optimization*

To assess the protein stability in different buffers, thermal melts were performed using a Rotor-Gene Q Real time PCR instrument (QIAGEN, Hilden, Germany). Purified GlnH in buffer E was concentrated to 12 mg/mL and then diluted 36-fold into a series of test buffers containing 10× SYPRO Orange reagent (Sigma-Aldrich, 5000× stock, catalogue number S5692). The samples were then thermally denatured by heating from 35°C to 90°C at a ramp rate of 0.5°C/min. Protein unfolding was monitored by following the SYPRO Orange fluorescence emission (λ_{ex} 530 nm/λ_{em} 555 nm). The unfolding data were fit to a derivation of the Boltzmann equation for the two-state unfolding model to obtain the midpoint of denaturation (the melting temperature T_m) (*33*) using GraphPad Prism. All experiments were performed in triplicate.

2.6. *Crystallization*

GlnH was concentrated to 20 mg/mL using an Amicon Ultracel 10 kDa cut-off concentrator (Merck, Darmstadt, Germany) and centrifuged at 4°C for 30 min at 13,200× g to clarify the solution. The crystallization screening was performed by the hanging-drop vapour-diffusion method using an automated Phoenix crystallization robot (Art Robbins Instruments, CA, USA) and commercial screens PEG/Ion HT and Crystal Screen HT (Hampton Research, CA, USA), JCSG+ Suite (QIAGEN, Hilden, Germany) and JBS HTS1 and HTS2 (Jena Bioscience, Jena, Germany). The initial crystallization droplets contained 100 nL of protein solution plus 100 nL of reservoir solution, and were equilibrated against 50 μL of reservoir solution in a 96-well Art-Robins CrystalMation Intelli-plate (Hampton Research, CA, USA). After one day, needle-like crystals appeared in condition B10 of the Crystal Screen HT containing 0.2 M sodium acetate, 0.1 M Tris-HCl pH

8.5 and 30% w/v polyethylene glycol (PEG) 4000. After two days, needle-like crystals also appeared in condition H3 of the JCSG+ Suite screen, containing 0.17 M ammonium sulfate, 25.5% w/v PEG 4,000 and 15% v/v glycerol. Manual optimization resulted in larger, rod-like crystals in a condition containing 0.2 M sodium acetate (Merck, Darmstadt, Germany), 0.05 M Tris-HCl pH 8.5 and 28% w/v PEG 4,000 (Sigma-Aldrich, MO, USA).

2.7. *Data collection, processing and molecular replacement*

Prior to data collection, the crystals were briefly soaked in a cryoprotectant solution containing 0.1 M Tris-HCl pH 8.5, 0.2 M sodium acetate, 36% w/v PEG 4,000 and 10% v/v glycerol, and flash-frozen in liquid nitrogen. X-ray diffraction data were collected from a single crystal to 1.3 Å resolution using an ADSC Quantum 210r detector (Area Detector Systems Corporation, Poway, CA, USA) at the MX1 beamline of the Australian Synchrotron. A total of 360 images were taken using 0.5° oscillations. Data processing and scaling were performed using *iMosflm* (*34*) and *AIMLESS* (*35*), respectively from the CCP4 suite (*36*). The Matthews coefficient was calculated using *MATTHEWS_COEF* (*37*) from CCP4. Molecular replacement was performed using *Phaser* (*38*) from CCP4 with the structure of cysteine-binding protein from *C. jejuni* (PDB ID 1XT8) as a search model. Model refinement was carried out using *Phenix* (*39*).

3. Results and Discussion

The *H. pylori glnH* gene encodes a 277-a.a. long pre-protein containing a cleavable N-terminal signal peptide (a.a. 1-24). The recombinant GlnH that lacks the signal sequence was overexpressed in *E. coli* and purified to > 98% electrophoretic homogeneity as determined by Coomassie Blue-stained SDS-PAGE gel (Figure 1). The recombinant protein contained amino acid residues 25-277 of GlnH plus six N-terminal additional residues from the TEV cleavage site and the vector (GIDPFT). It migrated on SDS-PAGE with an apparent molecular weight (MW) of ~29 kDa, which is in agreement with the value calculated from the amino acid sequence (29.2 kDa).

Prior to the biophysical and crystallization experiments, we searched for an optimal buffer, in which GlnH is most stable, using a thermal shift assay. We monitored thermal unfolding of the protein samples prepared in different buffers as they were heated from 35°C to 90°C. The results of this assay (Figure 2) suggested that the protein is more stable at pH 6.5-8.0 than at pH ≤ 6.0, and that within the pH 6.5-8.0 range, the melting temperature of the protein in the gel filtration buffer E (20 mM Tris-HCl pH 8.0, 150 mM
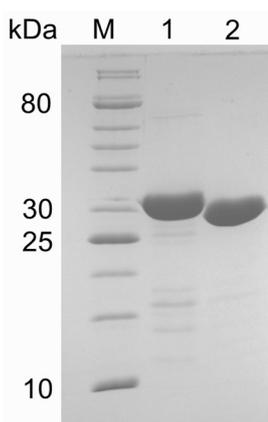
Figure 1. Coomassie blue-stained 15% SDS-PAGE gel of recombinant GlnH. M, molecular weight ladder; Lane 1, 15 μg of GlnH-His6 after the first Ni-NTA step; Lane 2, 15 μg of purified, untagged GlnH.
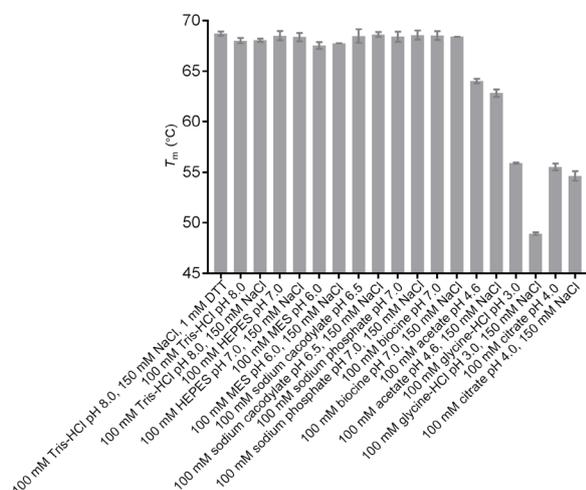
Figure 2. Effect of different buffers on the melting temperature T_m of GlnH. Results are means ± standard deviation (S.D.) for three independent replicates.

Figure 3. Far-UV CD spectrum of *H. pylori* GlnH.
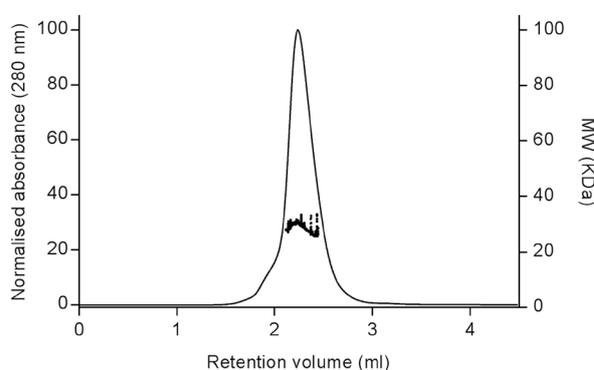
Figure 4. SEC-MALS analysis of GlnH-His₆. The peak (thin line) represents the UV trace for fractions containing protein. The UV absorbance values, normalized to 0-100% scale, are shown on the left-hand *y*-axis. The thick black line across the peak indicates the estimated molecular mass of the eluting protein particles (values are shown on the right-hand *y* axis).

Table 1. Molecular weights estimated by SEC-MALS analysis (BSA, bovine serum albumin)

Sample	Polydispersity	Molecular weight (kDa)
GlnH-His₆	1.0	34.4
BSA control	1.0	63.2

NaCl) does not differ much from that in other tested buffers. Therefore, to streamline the procedure, the protein for all subsequent experiments was prepared in the gel filtration buffer E.

Circular dichroism analysis (Figure 3) of GlnH-His₆ in the far-UV region of 200-250 nm involved fitting of the experimental spectrum for the purpose of secondary structure calculation using BeStSel server. The values derived from the spectrum (α 35%, β 21%) were very close to those predicted from the primary sequence analysis (α 34%, β 22%) using the YASPIN server (*40*), confirming that the recombinant protein is folded.

SEC-MALS analysis was performed to determine the oligomeric state of GlnH-His₆ in solution and the sample monodispersity. GlnH eluted as a single roughly symmetrical peak (Figure 4), the middle of which corresponds to a species with a polydispersity index value of 1.0. This indicates that the sample was largely homogenous with respect to the molecular mass.

The MW of Gln-His₆, as determined by SEC-MALS analysis (Table 1), was 34.4 kDa, which is close to the value calculated from the protein sequence (32.2 kDa), demonstrating that this protein is monomeric in solution. This result is in agreement with the findings that majority of PBPs behave as monomers in solution at physiologically relevant concentrations, including, for example, YckK (a different periplasmic binding protein from *H. pylori* (*41*)), ModA (a molybdate-binding protein from the archaeon *Methanosarcina acetivorans* (*42*)), SitA (a metal-binding PBP from the *Staphylococcus pseudintermedius* (*43*)) and OppA (a putative-oligopeptide binding PBP from *Thermotoga maritima* (*44*)).

Crystals of GlnH (Figure 5) were obtained using a sparse matrix crystallization approach. A complete
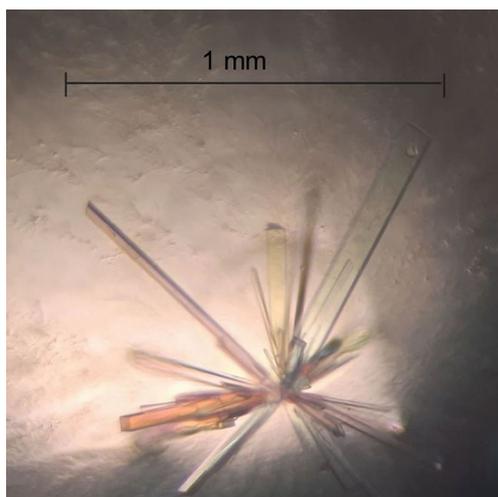
Figure 5. Crystals of GlnH from *H. pylori* SS1.

Figure 6. A representative diffraction image (0.5° oscillation) obtained for a GlnH crystal using an ADSC Quantum 210r detector on the MX1 station at the Australian Synchrotron, Victoria, Australia. The edge of the detector corresponds to the resolution of 1.36 Å.

Table 2. Data-collection and processing statistics

Items	MX1 beamline, Australian Synchrotron
Detector	ADSC Quantum 210r Detector
Wavelength (Å)	0.95
Temperature (K)	100
Total rotation range (°)	180
Mosaicity (°)	0.33
Space group	$P2_1$
Unit cell parameters	
a, b, c (Å)	38.67 93.36 64.13
α, β, γ (°)	90.00 93.72 90.00
Resolution range (Å)	37.71-1.30 (1.32-1.30)
Observed reflections	397,406 (12,068)
Unique reflections	106,244 (3,795)
Mean $I/\sigma(I)$	15.8 (3.3)
Completeness (%)	95 (69)
Multiplicity	3.7 (3.2)
R_{merge}	0.046 (0.293)
$CC_{(1/2)}$ (%)	99 (85)
R_{meas}	0.062 (0.384)
R_{pim}	0.032 (0.210)

Values in parentheses are for the highest resolution shell.

the Matthews coefficient (V_M) using MW = 29.2 kDa gave values of 4.0 and 2.0 $Å^3 Da^{-1}$, assuming one or two subunits in the asymmetric unit, respectively (*45*). An automated molecular replacement (MR) was performed with *Phaser* using the coordinates of the cysteine-binding protein from *C. jejuni* (PDB ID 1XT8) as a search model. The MR solution contained 2 subunits in the asymmetric unit. After several rounds of XYZ and B refinement using the simulated annealing protocol in *Phenix*, the R and R_{free} values fell from 0.481 and 0.487 to 30.2 and 32.9, respectively, confirming that the MR solution is correct. Model building is currently under way.

Worku *et al*. (*46*) demonstrated that glutamine acts as a chemoattractant for the *H. pylori* isolates obtained from endoscopic biopsies of patients with non-ulcer dyspepsia or duodenal ulcer. Given the similarity of *H. pylori* GlnH to the glutamine-binding PBP of *E. coli*, it would be essential to establish in the future if glutamine is a natural ligand recognized by GlnH, and if not, what is, and what role GlnH plays in chemotaxis and in the nutrient uptake. Furthermore, determination of the crystal structure of GlnH in complex with its natural ligand will provide an insight into the structural basis of its ligand specificity that could be an important step towards the development of therapeutic small molecule inhibitors targeting chemotaxis or nutrient uptake.

data set was collected from a single cryo-cooled crystal using the Australian Synchrotron (AS) facility to a resolution of 1.3 Å (Figure 6, Table 2). Autoindexing of the diffraction data with *iMosflm* suggested that the crystal belonged to the primitive monoclinic symmetry, with unit cell parameters a = 38.67, b = 93.36, c = 64.13 Å, α = 90.00, β = 93.72, γ = 90.00°. The analysis of the scaled data using POINTLESS (*36*) showed systematic absences along the 0k0 axis, with only reflections with *k* = 2n present, which suggested that the crystals belong to space group $P2_1$. The average $I/\sigma(I)$ value was 15.8 for all reflections (resolution range 37.71-1.30 Å) and 3.3 in the highest resolution shell (1.32-1.30 Å). A total of 397,406 measurements were made of 106,244 independent reflections. Data processing gave an R_{merge} of 0.046 for intensities (0.293 in the highest resolution shell). The data was 95% complete (69% completeness in the outer shell). Calculation of

References

1. Mitchell H, Katelaris P. Epidemiology, clinical impacts and current clinical management of *Helicobacter pylori* infection. Med J Aust. 2016; 204:376-380.

2. Andersen LP. Colonization and infection by *Helicobacter pylori* in humans. Helicobacter. 2007; 12:12-15.

3. Pereira M-I, Medeiros JA. Role of *Helicobacter pylori* in gastric mucosa-associated lymphoid tissue lymphomas. World J Gastroenterol. 2014; 20:684-698.

4. Li JZ, Li JY, Wu TF, Xu JH, Huang CZ, Cheng D, Chen QK, Yu T. *Helicobacter pylori* infection is associated with type 2 diabetes, not type 1 diabetes: An updated meta-analysis. Gastroenterol Res Pract. 2017; 2017:5715403.

5. Kato M, Asaka M, Shimizu Y, Nobuta A, Takeda H, Sugiyama T. Relationship between *Helicobacter pylori* infection and the prevalence, site and histological type of gastric cancer. Aliment Pharmacol Ther. 2004; 20:85-89.

6. Nedenskov P. Nutritional requirements for growth of *Helicobacter pylori*. Appl Environ Microbiol. 1994; 60:3450-3453.

7. Weinberg MV, Maier RJ. Peptide transport in *Helicobacter pylori*: Roles of dpp and opp systems and evidence for additional peptide transporters. J Bacteriol. 2007; 189:3392-3402.

8. Haley KP, Gaddy JA. Metalloregulation of *Helicobacter pylori* physiology and pathogenesis. Front Microbiol. 2015; 6:911.

9. Fischer F, Robbe-Saule M, Turlin E, Mancuso F, Michel V, Richaud P, Veyrier FJ, De Reuse H, Vinella D. Characterization in *Helicobacter pylori* of a nickel transporter essential for colonization that was acquired during evolution by gastric *Helicobacter* species. PLoS Pathog. 2016; 12:e1006018.

10. Cao DM, Lu QF, Li SB, Wang JP, Chen YL, Huang YQ, Bi HK. Comparative genomics of *H. pylori* and non-pylori *Helicobacter* species to identify new regions associated with its pathogenicity and adaptability. BioMed Res Int. 2016; 2016:6106029.

11. Shaik MM, Cendron L, Salamina M, Ruzzene M, Zanotti G. *Helicobacter pylori* periplasmic receptor CeuE (HP1561) modulates its nickel affinity *via* organic metallophores. Mol Microbiol. 2014; 91:724-735.

12. Wilkens S. Structure and mechanism of ABC transporters. F1000Prime Rep. 2015; 7:14.

13. Tam R, Saier MH, Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol Rev. 1993; 57:320-346.

14. Felder CB, Graul RC, Lee AY, Merkle HP, Sadee W. The Venus flytrap of periplasmic binding proteins: An ancient protein module present in multiple drug receptors. AAPS Pharm Sci. 1999; 1:7-26.

15. Borrok MJ, Kiessling LL, Forest KT. Conformational changes of glucose/galactose-binding protein illuminated by open, unliganded, and ultra-high-resolution ligand-bound structures. Protein Sci. 2007; 16:1032-1041.

16. Dunten P, Mowbray SL. Crystal structure of the dipeptide binding protein from *Escherichia coli* involved in active transport and chemotaxis. Protein Sci. 1995; 4:2327-2334.

17. Garmory HS, Titball RW. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. Infect Immun. 2004; 72:6757-6763.

18. Tomb JF, White O, Kerlavage AR, *et al*. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature. 1997; 388:539-547.

19. McClain MS, Shaffer CL, Israel DA, Peek RM, Cover TL. Genome sequence analysis of *Helicobacter pylori* strains associated with gastric ulceration and gastric cancer. BMC genomics. 2009; 10:3-3.

20. Draper JL, Hansen LM, Bernick DL, Abedrabbo S, Underwood JG, Kong N, Huang BC, Weis AM, Weimer BC, van Vliet AH, Pourmand N, Solnick JV, Karplus K, Ottemann KM. Fallacy of the unique genome: Sequence diversity within single *Helicobacter pylori* strains. mBio. 2017; 8:e02321-16.

21. Baltrus DA, Amieva MR, Covacci A, Lowe TM, Merrell DS, Ottemann KM, Stein M, Salama NR, Guillemin K. The complete genome sequence of *Helicobacter pylori* strain G27. J Bacteriol. 2009; 191:447-448.

22. Masters PS, Hong JS. Genetics of the glutamine transport system in *Escherichia coli*. J Bacteriol. 1981; 147:805-819.

23. Willis RC, Iwata KK, Furlong CE. Regulation of glutamine transport in *Escherichia coli*. J Bacteriol. 1975; 122:1032-1037.

24. Shen QC, Simplaceanu V, Cottam PF, Ho C. Proton nuclear magnetic resonance studies on glutamine-binding protein from *Escherichia coli*. Formation of intermolecular and intramolecular hydrogen bonds upon ligand binding. J Mol Biol. 1989; 210:849-857.

25. Hsiao CD, Sun YJ, Rose J, Wang BC. The crystal structure of glutamine-binding protein from *Escherichia coli*. J Mol Biol. 1996; 262:225-242.

26. Muller A, Thomas GH, Horler R, Brannigan JA, Blagova E, Levdikov VM, Fogg MJ, Wilson KS, Wilkinson AJ. An ATP-binding cassette-type cysteine transporter in *Campylobacter jejuni* inferred from the structure of an extracytoplasmic solute receptor protein. Mol Microbiol. 2005; 57:143-155.

27. Bulut H, Moniot S, Licht A, Scheffel F, Gathmann S, Saenger W, Schneider E. Crystal structures of two solute receptors for L-cystine and L-cysteine, respectively, of the human pathogen *Neisseria gonorrhoeae*. J Mol Biol. 2012; 415:560-572.

28. Hiller K, Grote A, Scheer M, Munch R, Jahn D. PrediSi: Prediction of signal peptides and their cleavage positions. Nucleic Acids Res. 2004; 32:W375-379.

29. Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol. 2004; 338:1027-1036.

30. Woon AP, Tohidpour A, Alonso H, Saijo-Hamano Y, Kwok T, Roujeinikova A. Conformational analysis of isolated domains of *Helicobacter pylori* CagA. PloS One. 2013; 8:e79367.

31. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976; 72:248-254.

32. Micsonai A, Wien F, Bulyaki E, Kun J, Moussong E, Lee YH, Goto Y, Refregiers M, Kardos J. BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. Nucleic Acids Res. 2018; 46:W315-w322.

33. Orwig SD, Lieberman RL. Biophysical characterization of the olfactomedin domain of myocilin, an extracellular matrix protein implicated in inherited forms of glaucoma. PloS One. 2011; 6:e16347.

34. Battye TG, Kontogiannis L, Johnson O, Powell HR,

Leslie AG. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011; 67:271-281.

35. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013; 69:1204-1214.

36. Winn MD, Ballard CC, Cowtan KD, *et al*. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011; 67:235-242.

37. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci. 2003; 12:1865-1871.

38. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr. 2005; 61:458-464.

39. Adams PD, Afonine PV, Bunkoczi G, *et al*. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213-221.

40. Lin K, Simossis VA, Taylor WR, Heringa J. A simple and fast secondary structure prediction method using hidden neural networks. Bioinformatics. 2005; 21:152-159.

41. Rahman M, Germantsis D, Machuca M, Ud-Din A, Roujeinikova A. Crystallisation and preliminary crystallographic analysis of *Helicobacter pylori* periplasmic binding protein YckK. Crystals. 2017; 7:330.

42. Chan S, Giuroiu I, Chernishof I, Sawaya MR, Chiang J, Gunsalus RP, Arbing MA, Perry LJ. Apo and ligand-bound structures of ModA from the archaeon *Methanosarcina acetivorans*. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010; 66:242-250.

43. Abate F, Malito E, Cozzi R, Lo Surdo P, Maione D, Bottomley Matthew J. Apo, Zn^{2+}-bound and Mn(2+)-bound structures reveal ligand-binding properties of SitA from the pathogen *Staphylococcus pseudintermedius*. Biosci Rep. 2014; 34:e00154.

44. Yoon HJ, Kim HJ, Mikami B, Yu YG, Lee HH. Crystal structure of a putative oligopeptide-binding periplasmic protein from a hyperthermophile. Extremophiles. 2016; 20:723-731.

45. Matthews BW. Solvent content of protein crystals. J Mol Biol. 1968; 33:491-497.

46. Worku ML, Karim QN, Spencer J, Sidebotham RL. Chemotactic response of *Helicobacter pylori* to human plasma and bile. J Med Microbiol. 2004; 53:807-811.