Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and X-ray data collection of HP0902 from *Helicobacter pylori* 26695

Dae-Won Sim,[a]‡ Jung Hyun Song,[b]‡ Woo Cheol Lee,[b]‡ Yoo-Sup Lee,[a] Hye-Yeon Kim[b]* and Hyung-Sik Won[a]*

[a]Department of Biotechnology, College of Biomedical and Health Science, Konkuk University, Chungju, Chungbuk 380-701, Republic of Korea, and [b]Division of Magnetic Resonance Research, Korea Basic Science Institute, Ochang, Chungbuk 363-883, Republic of Korea

‡ These authors contributed equally to this work.

Correspondence e-mail: hyeyeon@kbsi.re.kr, wonhs@kku.ac.kr

Received 14 June 2011
Accepted 25 September 2011

HP0902 from *Helicobacter pylori* 26695 belongs to the cupin superfamily of proteins, which encompasses proteins with a great diversity in function. In this work, two types of recombinant HP0902 protein were crystallized: one with an N-terminal His_6 tag (^{H6}HP0902) and the other with a C-terminal His_6 tag (HP0902^{H6}). The ^{H6}HP0902 crystal diffracted to 1.40 Å resolution and belonged to space group $P2_1$, with unit-cell parameters a = 33.5, b = 78.6, c = 41.4 Å. The HP0902^{H6} crystal belonged to space group $P4_32_12$ or $P4_12_12$ and diffracted to 2.5 Å resolution, with unit-cell parameters $a = b$ = 50.4, c = 142.0 Å.

1. Introduction

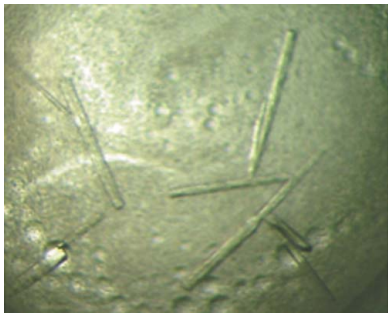
Infection by the gastric pathogen *Helicobacter pylori* induces severe gastric disorders, including chronic gastritis, peptic ulcers and stomach cancer. It is reasonable to consider the proteins secreted by *H. pylori* as potential virulence factors of the bacterium, since they can contribute to the gastric inflammation process. For example, the protein VacA has been identified as one of the critical determinants of virulence in *H. pylori*. VacA is secreted by the bacterium and mediates the pathogenesis of peptic ulceration and gastric cancer in human cells (Isomoto *et al.*, 2010). HP0902, another protein identified as being secreted by *H. pylori* (Kim *et al.*, 2002), has been hypothesized to interact with VacA (Sim, Ahn *et al.*, 2009). In addition, HP0902 is overexpressed in a mutant strain of *H. pylori* that lacks the *fdx*A gene, which regulates resistance to the antibiotic metronidazole (Mukhopadhyay *et al.*, 2003; Nam *et al.*, 2007).

Owing to its importance as a human pathogen, the genomes of several strains of *H. pylori* have been sequenced completely, the first of which was strain 26695 (Tomb *et al.*, 1997); moreover, structural genomics studies are currently being attempted by several groups around the world (Sim, Lee *et al.*, 2009). In this context, we initiated a structural study of HP0902, the function of which is unknown. HP0902 is a homodimeric 22 kDa protein. A previous NMR study (Sim, Ahn *et al.*, 2009; Sim, Lee *et al.*, 2009), which characterized its secondary structure, implied that this protein belongs to the cupin superfamily. Cupins are ubiquitous proteins that share a highly conserved β-barrel topology and many cupins have been identified as allergens (Mills *et al.*, 2002; Dunwell *et al.*, 2004). However, the cupin superfamily includes proteins with a wide variety of functions and sequences and is currently classified into 39 subfamilies in the Pfam database (Finn *et al.*, 2010). Thus, precise analysis of the three-dimensional structure of HP0902 is essential for its functional identification in terms of structural genomics. Since our previous NMR study was unsuccessful in determining its three-dimensional structure, we attempted X-ray crystallography of the protein. Here, we report the crystallization and preliminary X-ray analysis of the HP0902 protein from *H. pylori* 26695.

2. Materials and methods

2.1. Cloning, expression and purification

The ORF of HP0902 was amplified by PCR using the genomic DNA of *H. pylori* 26695 (ATCC 700392) as a template. The forward



and reverse oligonucleotide primers were 5′-G GAA TTC **CAT ATG** GAA GTG GTT CAT TTT TTA-3′ and 5′-CCG CCG **CTC GAG** TTA TTT TTT ACT TAA AGA TAG CCT-3′, respectively, where the *Nde*I and *Xho*I restriction-enzyme cleavage sites are shown in bold. The PCR products were digested with *Nde*I and *Xho*I and ligated into *Nde*I/*Xho*I-digested expression vector pET-15b (Novagen) in order to produce recombinant HP0902 with an N-terminal His_6 tag (^{H6}HP0902). The resulting construct contained 19 non-native residues at the N-terminus, including the His_6 tag and the thrombin cleavage site (bold): MGSSHHHHHHSSG**LVPRGS**. For cloning of recombinant HP0902 with a C-terminal His_6 tag (HP0902^{H6}), the reverse primer (5′-CCG CCG **CTC GAG** TTT TTT ACT TAA AGA TAG CCT-3′) did not contain a stop codon and the amplified DNA was inserted into the pET-21a vector (Novagen). This construct contained eight non-native residues at the C-terminus (LEHHHHHH), including a His_6 tag that facilitates protein purification. The constructed plasmids, which were verified *via* DNA sequencing, were transformed into *Escherichia coli* strain BL21 (DE3) pLysS. The transformed cells were grown in Luria–Bertani (LB) medium at 310 K. When the A_{600} reached about 0.6, protein expression was induced by adding IPTG to a final concentration of 1 m*M* and induction was continued for a further 4 h. Cells were harvested by centrifugation and resuspended in ice-cold buffer *A* (20 m*M* Tris–HCl, 10 m*M* NaCl, 5 m*M* imidazole pH 8.0). The cells were disrupted by sonication at 277 K and the supernatant was loaded onto a HisTrap FF column (GE Healthcare) pre-equilibrated with buffer *A*. After extensive washing with the same buffer, the bound protein was eluted using a gradient from 5 to 500 m*M* imidazole in buffer *A*. Fractions containing HP0902 were pooled and dialyzed in buffer *B* (20 m*M* Tris–HCl, 1 m*M* DTT pH 8.0) followed by loading onto a HiTrap Q FF column (GE Health-

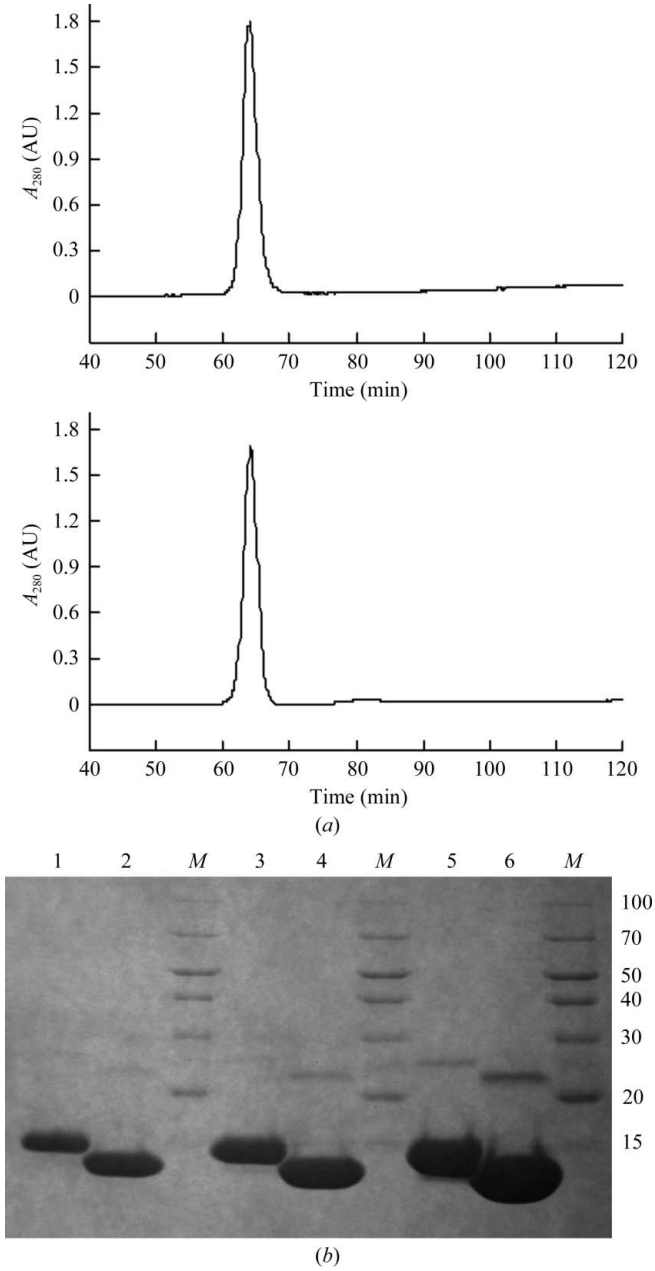


Figure 1
Purification of ^{H6}HP0902 and HP0902^{H6}. (*a*) Elution profiles of purified ^{H6}HP0902 (upper panel) and HP0902^{H6} (bottom panel) from a gel-filtration column (43 ml void volume and 120 ml total bed volume). The flow rate was 1 ml min^{-1} and a single peak was observed for the purified protein solution. (*b*) SDS–PAGE of purified ^{H6}HP0902 (lane 1, 12 µg; lane 3, 23 µg; lane 5, 58 µg) and HP0902^{H6} (lane 2, 17 µg; lane 4, 33 µg; lane 6, 83 µg). Small fractions of dimeric (between 20 and 30 kDa) proteins were detected even under denaturing conditions. Lane *M* contains molecular-weight markers (labelled in kDa).

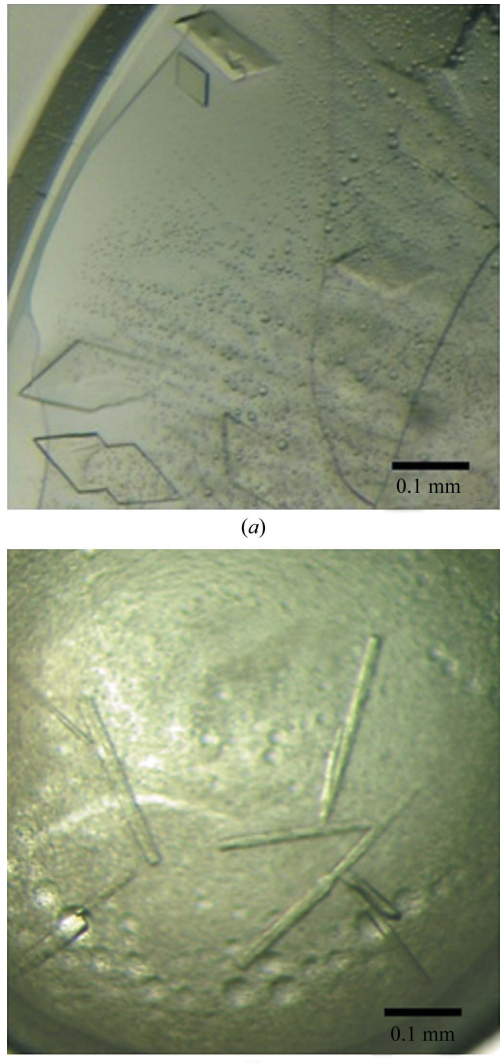


Figure 2
Native crystals of (*a*) ^{H6}HP0902 and (*b*) HP0902^{H6}. The average dimensions of the crystals were approximately 0.1 × 0.1 × 0.05 mm (*a*) and 0.2 × 0.05 × 0.05 mm (*b*).

Table 1
X-ray data-collection statistics for ^{H6}HP0902 and HP0902^{H6}.

Values in parentheses are for the highest resolution bin.

Construct	^{H6}HP0902	HP0902^{H6}
Beamline	BL-17A, PF	BL-1A, PF
Wavelength (Å)	0.98	0.98
Space group	$P2_1$	$P4_12_12$ or $P4_32_12$
Unit-cell parameters (Å, °)	a = 33.5, b = 78.6, c = 41.4, β = 112.2	$a = b$ = 50.4, c = 142.0
Resolution range (Å)	71–1.40	71–2.50
No. of observations	250905	31274
No. of unique reflections	38998	6158
R_{merge}†	0.054 (0.362)	0.042 (0.310)
Multiplicity	6.4	5.1
Completeness (%)	99.8	98.2
Mean $I/\sigma(I)$	15.8 (3.6)	39.9 (5.09)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

care) pre-equilibrated with buffer *B*. The bound protein was eluted using a linear NaCl gradient (10–500 m*M*) in buffer *B*. Fractions containing HP0902, which eluted at around 150 m*M* NaCl, were concentrated to about 2 ml and applied onto a HiLoad 16/60 Superdex 75 column (Pharmacia) that had been equilibrated with the final buffer (10 m*M* Tris–HCl pH 8.0, 1 m*M* DTT). The purified solution of each construct was concentrated to around 35 mg ml^{-1}, as estimated *via* a typical BCA assay, for protein crystallization screening trials. The purified protein was judged to be >95% pure by SDS–PAGE and from the gel-filtration chromatogram (Fig. 1).

2.2. Crystallization

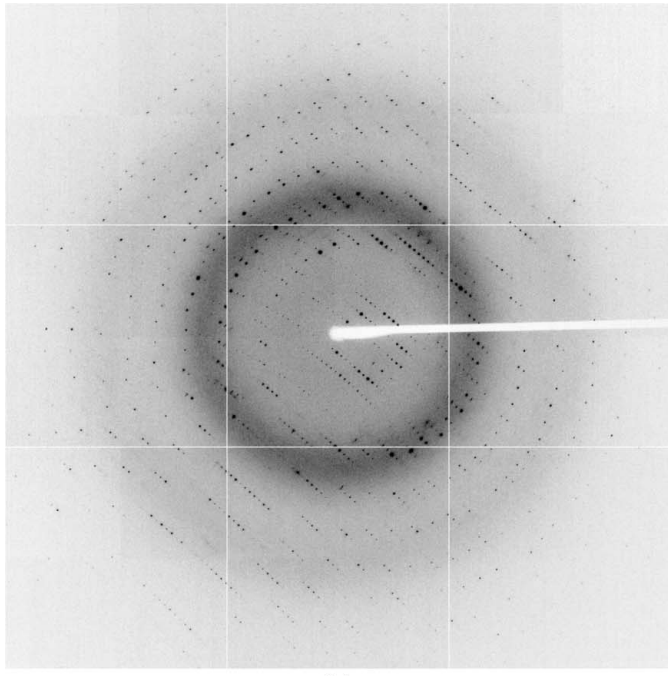
Crystallization of ^{H6}HP0902 and HP0902^{H6} was initially screened by the sitting-drop vapour-diffusion method at 293 K using a Mosquito crystallization robot (TTP LabTech). 0.2 µl protein solution (at 36 mg ml^{-1} for ^{H6}HP0902 and 34 mg ml^{-1} for HP0902^{H6}) was mixed with 0.2 µl reservoir solution and equilibrated against 70 µl of commercially available reservoir solutions from the Crystal Screen, Crystal Screen 2, Index (Hampton Research), Wizard I and Wizard II (Emerald BioSystems) screening kits. Initial crystals of ^{H6}HP0902 and HP0902^{H6} were formed using Index conditions G1 and G7, respectively. To obtain larger crystals, conditions were manually optimized by changing the PEG concentration and by using the hanging-drop vapour-diffusion method. Drops consisted of 1 µl protein solution mixed with 1 µl reservoir solution and were equilibrated against 1 ml reservoir solution. Well diffracting ^{H6}HP0902 crystals (Fig. 2*a*) were produced in 3 d at 293 K under the optimal conditions 25%(*w*/*v*) polyethylene glycol (PEG) 3350, 0.2 *M* sodium chloride, 0.1 *M* Tris–HCl pH 8.5. Diffraction-quality HP0902^{H6} crystals (Fig. 2*b*) were produced in 3 d under the condition 25%(*w*/*v*) PEG 3350, 0.2 *M* ammonium acetate, 0.1 *M* bis-tris pH 6.5.

2.3. Data collection

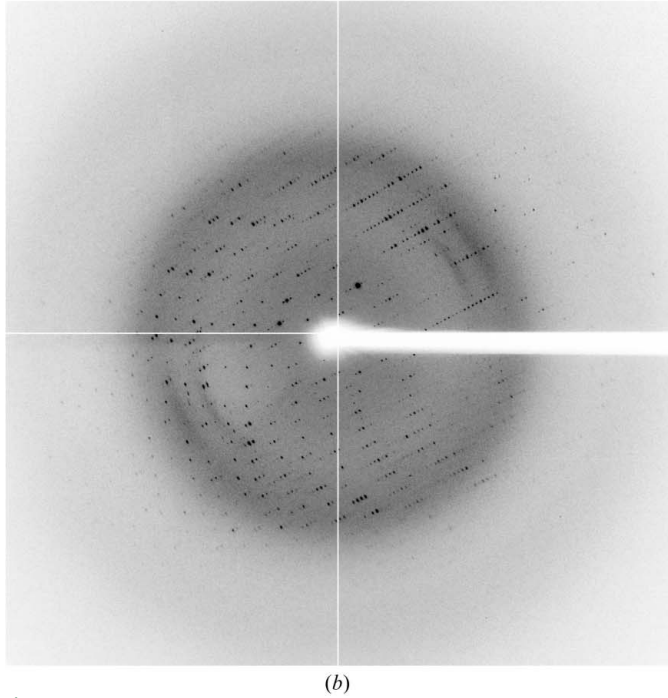
Prior to data collection, single crystals of ^{H6}HP0902 and HP0902^{H6} were picked up in nylon loops and flash-cooled in a nitrogen stream at 100 K. The native crystals were directly mounted under liquid nitrogen for diffraction studies under cryogenic conditions and no ice rings were observed. Additional cryoprotectant was not required as the high PEG content of the crystallization conditions prevented ice-ring formation. In order to find a suitable crystal for data collection, preliminary diffraction tests were performed using a MicroMax-007 HF X-ray generator and an R-AXIS IV^{++} imaging-plate area detector (Rigaku Americas). The final X-ray diffraction data for the native crystals (Fig. 3) were collected on the BL-17A and BL-1A beamlines of the Photon Factory (PF; Tsukuba, Japan). The collected diffraction data were integrated and scaled using the programs *MOSFLM* and *SCALA* from the *CCP*4 software package (Leslie, 1992; Evans, 2006; Winn *et al.*, 2011).

3. Results and discussion

As shown in Fig. 2, prism-shaped and rod-shaped crystals were finally obtained for ^{H6}HP0902 and HP0902^{H6}, respectively. The ^{H6}HP0902 crystal belonged to space group $P2_1$, with unit-cell parameters a = 33.5,



(*a*)



(*b*)

Figure 3
Diffraction images of (*a*) ^{H6}HP0902 and (*b*) HP0902^{H6} crystals.

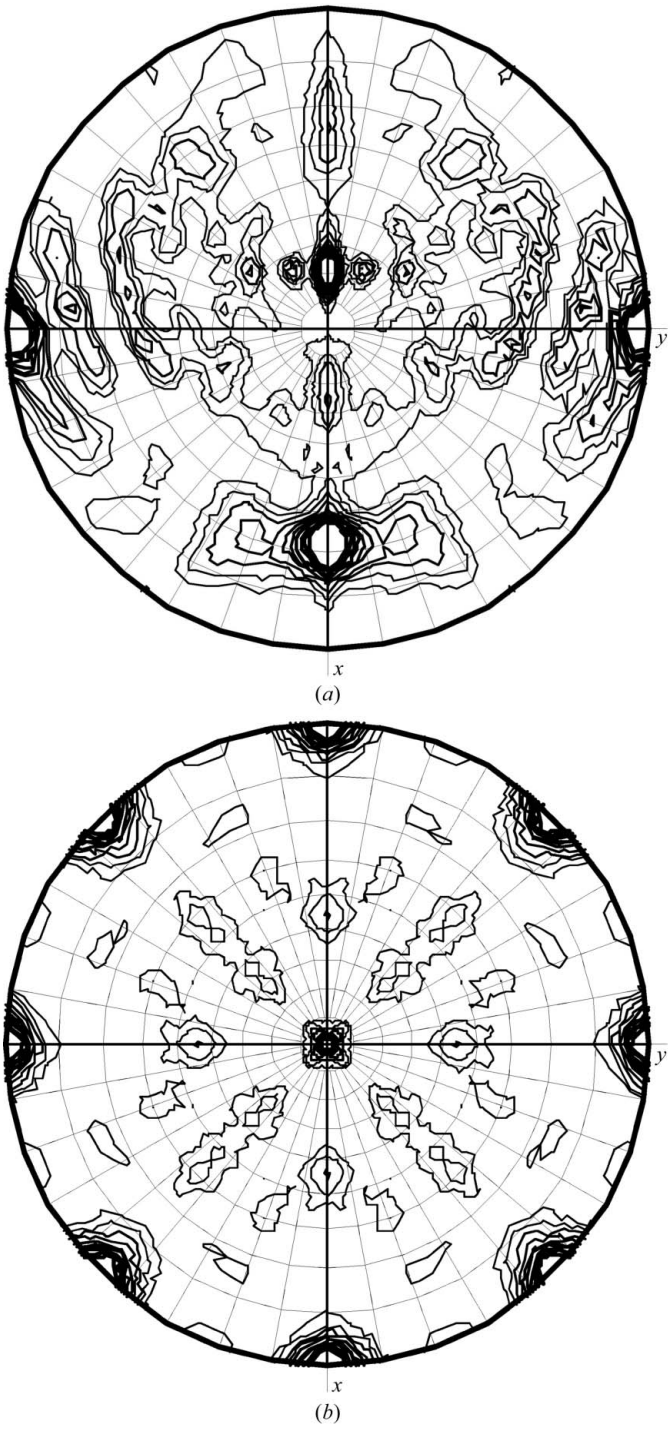


Figure 4
Self-rotation plots for the data sets of (*a*) ^{H6}HP0902 and (*b*) HP0902^{H6}. Self-rotation maps were generated using the *MOLREP* program (Vagin & Teplyakov, 2010) and contoured at the $\chi = 180^\circ$ section for the monoclinic (*a*) and tetragonal (*b*) crystal forms.

b = 78.6, c = 41.4 Å, while the HP0902^{H6} crystal belonged to space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 50.4$, c = 142.0 Å (Table 1). In both crystal forms the unit-cell parameters allowed two molecules of HP0902 in the asymmetric unit, which is consistent with our previous observation that the protein behaves as a homodimer in solution (Sim, Lee *et al.*, 2009). The Matthews coefficients calculated for the monoclinic ^{H6}HP0902 and tetragonal HP0902^{H6} crystals were 1.88 Å^3 Da^{-1} (35% solvent content) and 1.91 Å^3 Da^{-1} (36% solvent content), respectively (Matthews, 1968). Self-rotation function analysis (Fig. 4) of ^{H6}HP0902 suggested that the two molecules in the asymmetric unit were related by twofold noncrystallographic symmetry (NCS). NCS was not apparent in the tetragonal crystal form (HP0902^{H6}), which may result from overlap of the twofold NCS axis and one of the crystallographic twofold symmetries.

In our initial attempts to crystallize HP0902 without a His$_6$ tag, needle-shaped crystals were obtained for $^{\Delta H6}$HP0902, which was formed by cleavage of the N-terminal His$_6$ tag from ^{H6}HP0902, but attempts to improve the crystal quality failed (data not shown). However, despite being fused to a His$_6$ tag, both the ^{H6}HP0902 and HP0902^{H6} constructs readily crystallized; the position of the His$_6$ tag probably affected the crystal-packing process in favourable ways, thereby changing the shape of the crystals (which belong to different space groups and have different unit-cell parameters; Table 1). In particular, the crystal of the N-terminally His$_6$-tag-fused construct ^{H6}HP0902 diffracted to a high resolution of 1.40 Å. However, in many cupin-family proteins the N-terminus plays a critical role in intermolecular contacts for dimerization and these may have been affected by the N-terminal His$_6$ tag of the ^{H6}HP0902 construct. Thus, we additionally collected diffraction data from the HP0902^{H6} crystal, in which a His$_6$ tag was fused at the C-terminus. Since the HP0902^{H6} crystal diffracted to 2.5 Å resolution, which is sufficient for diffraction analysis, the two crystal structures should supplement and validate each other, thereby providing indubitable information on the native intact protein structure. Structure determination using molecular replacement is under way. The high-resolution crystal structure of HP0902 will be valuable in elucidating the function of this interesting protein secreted by the clinically important pathogen *H. pylori*.

This work was supported by the Korea Healthcare Technology R&D Project (A092006 to H-SW), Ministry for Health, Welfare and Family Affairs, Republic of Korea and in part by the Korean Membrane Protein Initiative program (to H-YK) of the Korean Ministry of Education, Science and Technology.

References

Dunwell, J. M., Purvis, A. & Khuri, S. (2004). *Phytochemistry*, **65**, 7–17.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L. L., Eddy, S. R. & Bateman, A. (2010). *Nucleic Acids Res.* **38**, D211–D222.
Isomoto, H., Moss, J. & Hirayama, T. (2010). *Tohoku J. Exp. Med.* **220**, 3–14.
Kim, N., Weeks, D. L., Shin, J. M., Scott, D. R., Young, M. K. & Sachs, G. (2002). *J. Bacteriol.* **184**, 6155–6162.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mills, E. N., Jenkins, J., Marigheto, N., Belton, P. S., Gunning, A. P. & Morris, V. J. (2002). *Biochem. Soc. Trans.* **30**, 925–929.
Mukhopadhyay, A. K., Jeong, J. Y., Dailidiene, D., Hoffman, P. S. & Berg, D. E. (2003). *J. Bacteriol.* **185**, 2927–2935.
Nam, W. H., Lee, S. M., Kim, E. S., Kim, J. H. & Jeong, J. Y. (2007). *Korean J. Life Sci.* **17**, 723–727.
Sim, D.-W., Ahn, H.-C. & Won, H.-S. (2009). *J. Korean Soc. Magn. Reson.* **13**, 108–116.
Sim, D.-W., Lee, Y.-S., Kim, J.-H., Seo, M.-D., Lee, B.-J. & Won, H.-S. (2009). *BMB Rep.* **42**, 387–392.
Tomb, J. F. *et al.* (1997). *Nature (London)*, **388**, 539–547.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.