Biochemical and Biophysical Research Communications 394 (2010) 940–946

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Crystal structure and functional insight of HP0420-homolog from *Helicobacter felis*

Shunfu Piao [a], Xiao Ling Jin [a], Bo-Young Yun [a], Nahee Kim [b], Hyun-Soo Cho [b], Minoru Fukuda [c], Heeseob Lee [d], Nam-Chul Ha [a,*]

[a] College of Pharmacy and Research Institute for Drug Development, Pusan National University, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea
[b] Department of Biology, College of Science, Yonsei University, Seoul 120-749, Republic of Korea
[c] Tumor Microenvironment Program, Cancer Research Center, Burnham Institute for Medical Research, La Jolla, CA 92037, USA
[d] Department of Food Science and Nutrition, Pusan National University, Jangjeon-dong, Geumjeong-gu, Busan 609-735, Republic of Korea

ARTICLE INFO

Article history:
Received 8 March 2010
Available online 17 March 2010

Keywords:
Crystal structure
Helicobacter
Hot-dog fold
Structural genomics
HP0420

ABSTRACT

Helicobacter pylori infect more than half of the world's population and are considered a cause of peptic ulcer disease and gastric cancer. Recently, hypothetical gene HP0421 was identified in *H. pylori* as a cholesterol α-glucosyltransferase, which is required to synthesize cholesteryl glucosides, essential cell wall components of the bacteria. In the same gene-cluster, HP0420 was co-identified, whose function remains unknown. Here we report the crystal structure of HP0420-homolog of *H. felis* (HF0420) to gain insight into the function of HP0420. The crystal structure, combined with size-exclusion chromatography, reveals that HF0420 adopts a homodimeric hot-dog fold. The crystal structure suggests that HF0420 has enzymatic activity that involves a conserved histidine residue at the end of the central α-helix. Subsequent biochemical studies provide clues to the function of HP0420 and HF0420.

1. Introduction

Helicobacter pylori is a Gram-negative bacteria that lives in human stomach and may cause gastric chronic inflammation and even stomach cancer [1]. *Helicobacter* species have cholesteryl glucosides (CGs) as unique and essential cell membrane components [2,3]. CGs are synthesized from cholesterol, and thus cholesterol must be taken up from the host cell because the bacteria lack the genes responsible for its biosynthesis [4]. Recently, it was reported that the hypothetical gene HP0421 from *H. pylori* exhibits cholesterol α-glucosyltransferase activity that converts cholesterol into CGs and can be inhibited by the natural antibiotic mucin *O*-glycan in deeper portions of the gastric mucosa [5]. Bacterial growth is severely inhibited in the absence of the *HP0421* gene [6,7], suggesting a critical role for HP0421 in the survival of *H. pylori*.

During expression cloning of cholesterol α-glucosyltransferase from *H. pylori* and *H. felis*, Lee et al. identified HP0420 and its homolog in a single plasmid harboring genomic sequences for two open reading frames (HP0420 and HP0421) [6]. *H. felis*, which lives in cat stomach [8] and shows genomic similarity to *H. pylori*, has been used to study colonization, pathogenesis, and eradication of *H. pylori* [9,10]. The gene structures of HP0420 and HP0421 were conserved in both *Helicobacter* species. Unlike HP0421, HP0420 was not essential for the survival of the bacteria [6]. Although the gene structure suggested that the function of HP0420 might be associated with HP0421, such association remains to be elucidated [6]. The HP0420-homolog from *H. felis* (herein referred to as HF0420) shares high sequence identity (40%) with HP0420, indicating that HF0420 is a functional and structural homolog of HP0420 [6].

To gain insight into the function of the hypothetical protein HP0420 and its homologs from *Helicobacter* species, we determined the crystal structure of HF0420 from *H. felis* and performed subsequent biochemical studies.

2. Materials and methods

2.1. DNA construction and protein expression

DNA fragments encoding HF0420 (residues 1–141) and HP0420 (residues 1–142) were amplified from *H. felis* and *H. pylori* genomic DNA, respectively, using the polymerase chain reaction. The DNA fragments were inserted into the NcoI and XhoI sites of pProEX-HTa (Invitrogen, USA) containing the hexa-histidine tag and the TEV protease cleavage sites at the N-terminus. The recombinant HF0420 and HP0420 proteins were expressed in *Escherichia coli* strain BL21 (DE3) RIL in Luria–Bertani (LB) medium at 37 °C until the $OD_{600\ nm}$ reached 0.5. The protein was expressed by adding 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation at 5000 rpm for 15 min at 4 °C.

2.2. Protein purification of HF0420 and HP0420

To produce HF0420 and HP0420 proteins, harvested cells were suspended in lysis buffer containing 20 mM Tris (pH 8.0), 150 mM

* Corresponding author.
E-mail address: hnc@pusan.ac.kr (N.-C. Ha).

0006-291X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2010.03.087

NaCl and 2 mM β-mercaptoethanol, and were disrupted by sonication. The lysate was centrifuged at 13,000 rpm for 30 min at 4 °C. The resulting supernatant was loaded onto Ni–NTA agarose resin that was pre-equilibrated with lysis buffer. The resin was washed with lysis buffer supplemented with 20 mM imidazole and then eluted with lysis buffer supplemented with 200 mM imidazole. The fractions containing the HF0420 protein were pooled and treated with recombinant TEV protease overnight at room temperature to remove the hexa-histidine tag after addition of 10 mM β-mercaptoethanol. The reaction mixture was subsequently loaded onto a Q-anion exchange column (Hitrap-Q; GE Healthcare, USA) for further purification, and eluted with a gradient from 0 to 1 M NaCl. The collected fractions containing the target protein were pooled, concentrated, and separated on a HiLoad Superdex 200 gel-filtration column (GE Healthcare, USA) pre-equilibrated with lysis buffer. During the purification, the presence of the protein was confirmed by SDS–PAGE. The purified HF0420 and HP0420 proteins were concentrated to 60 mg/ml and 40 mg/ml, respectively, using Centriprep (Millipore, USA) and stored frozen at −80 °C until use.

2.3. Crystallization, data collection and structure determination

Crystals of the wild-type HF0420 protein were obtained by the vapor-diffusion technique at 14 °C. The initial crystallization screening was performed by the sitting-drop method with Crystal Screen HT, a high-throughput sparse-matrix screening kit (Hampton Research, USA). Crystals grown in the solution containing 0.1 M Tris–HCl (pH 8.5) and 2 M ammonium sulfate were directly chosen from the initial crystallization screening plate for data collection. For cryoprotection, the HF0420 crystals were soaked with the sticky oil Paratone-N. The data sets were collected on BL6C at Pohang Accelerator Laboratory with a CCD detector Quantum 210 (ADSC) at −173 °C. The diffraction data sets were processed and scaled with the HKL2000 package [11]. The crystal belongs to the space group $P2_12_12_1$ with cell dimensions of $a = 70.5$, $b = 70.3$, and $c = 58.8$ Å. Initial phases were determined by the molecular replacement package MOLREP [12] using the coordinates of Cj0977 (Protein Data Bank code 3BNV) as a search model. Model building was performed using the program COOT [13], and model refinement was conducted using the program CNS 1.2 [14]. PHENIX.REFINE [15] was applied at the final round of the model refinement.

Crystals of mutant HF0420 (C46A) were obtained by the same method as wild-type HF0420, and the crystallization conditions were optimized to produce high quality crystals in droplets containing 1 μl of protein solution and 1 μl of a precipitant solution consisting of 0.01 M cobalt (II) chloride hexahydrate, 0.1 M MES, pH 6.5, and 1.8 M ammonium sulfate. Data sets were collected as described above, and the crystal belonged to the space group $P2_12_12_1$ with cell dimensions of $a = 53.3$, $b = 67.1$, and $c = 70.2$ Å. Initial phases were determined by the molecular replacement package MOLREP [12] using coordinates of the wild-type HF0420 structure as a search model.

Crystallographic data statistics are summarized in Table 1. All figures were prepared with PYMOL [16].

2.4. Size-exclusion chromatography

Size-exclusion chromatography was performed at room temperature at a flow rate of 0.5 ml/min on Superdex S-200 HR 10/30 (GE Healthcare) equilibrated with 20 mM Tris buffer (pH 8.0) containing 150 mM NaCl and 2 mM β-mercaptoethanol. Five hundred microliters of each protein (100 μg) were injected onto the column.

2.5. Thermal stability

Thermal stability studies of HF0420 (wild-type and C46A) and HP0420 were performed by circular dichroism in a JASCO-J750 spectropolarimeter. Samples were prepared in 20 mM Tris, 0.15 M NaCl, pH 8.0, and thermal unfolding experiments were performed by monitoring the circular dichroism signal at 220 nm between 25 °C and 90 °C using a heating rate of 2 °C/min at a concentration of 0.25 mg/ml.

Table 1
X-ray data collection and refinement statistics.

Data set	Wild-type	C46A mutant
Source	Beamline 6C at PLS	Beamline 6C at PLS
Wavelength (Å)	1.0000	1.0000
Resolution limit (Å)	30–1.6 (1.66–1.60)[a]	50–1.8 (1.83–1.80)[a]
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell (Å)	$a = 70.6$, $b = 70.3$, $c = 53.9$ Å, $\alpha = \gamma = \beta = 90°$	$a = 53.3$, $b = 67.1$, $c = 70.2$ Å, $\alpha = \gamma = \beta = 90°$
Redundancy	8.5 (3.5)	9.5 (6.7)
R_{sym}[b] (%)	5.3 (40.0)[a]	8.2 (31.8)[a]
I/σ	49.4 (2.8)[a]	30.5 (3.5)[a]
Completeness (%)	98.1 (84.6)[a]	99.7 (99.4)[a]
Refinement		
Resolution range (Å)	30–1.6 (1.66–1.60)[a]	50–1.8 (1.83–1.80)[a]
R-factor (%)	20.8	18.0
R_{free}[c] (%)	23.3	21.6
Number of refined atoms		
Protein	1941	2012
Water	173	118
Average *B* value[b] (Å)	37.1	17.4
Rmsd for bonds (Å)	0.017	0.017
Rmsd for angles (°)	1.602	1.65
Ramachandran plot		
Most favored (%)	89.2	93.2
Additionally favored (%)	8.2	5.1

[a] The numbers in parentheses are statistics for the highest resolution shell.
[b] $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensity and the mean intensity of related reflections, respectively.
[c] R_{free} was calculated with 10% of the data set.