

CRYSTALLIZATION P 67-71

# Purification, crystallization, and preliminary X-ray diffraction analysis of an S-formylglutathione hydrolase (VaSFGH) homolog from *Variovorax* sp. PAMC 28711

Jisub Hwang<sup>1,2</sup>, Min Ju Lee<sup>1</sup>, Sung Gu Lee<sup>1,2</sup>, Hackwon Do<sup>1</sup> and Jun Hyuck Lee<sup>1,2,\*</sup>

<sup>1</sup>Research Unit of Cryogenic Novel Material, Korea Polar Research Institute, Incheon 21990, Republic of Korea

<sup>2</sup>Department of Polar Sciences, University of Science and Technology, Incheon 21990, Republic of Korea

\*Correspondence: junhyucklee@kopri.re.kr

S-formylglutathione hydrolase (SFGH) is an esterase that hydrolyzes S-formylglutathione into formic acid and glutathione. As SFGHs are also able to hydrolyze thioesters as well as non-thioester substrates, they have attracted considerable attention as potential biocatalysts. Although the substrate specificity of various SFGHs has been determined, the detailed structural differences relating to substrate preference remain unclear. Here, we present overexpression, purification, and preliminary X-ray crystallographic data for an SFGH from *Variovorax* sp. PAMC 28711 (VaSFGH). The VaSFGH protein was over-expressed in *Escherichia coli* and successfully crystallized in 0.2 M sodium chloride, 0.1 M Bis-Tris:HCl (pH 6.5), and 20% (w/v) PEG 3350. A complete native X-ray diffraction dataset was collected up to 2.38 Å resolution and processed in the C2 space group with unit-cell parameters  $a = 53.2$  Å,  $b = 76.4$  Å,  $c = 199.9$  Å,  $\alpha = 90^\circ$ ,  $\beta = 90.2^\circ$ , and  $\gamma = 90^\circ$ . Moreover, VaSFGH exhibited higher esterase activity toward shorter-chain esters. Based on its structural determination, future studies will elucidate the substrate-binding mechanism and specificity of VaSFGH at the molecular level.

## INTRODUCTION

S-formylglutathione hydrolases (SFGHs) are widely found in various organisms including prokaryotes and eukaryotes. SFGHs catalyze the hydrolysis of S-formylglutathione to formic acid and glutathione and, therefore, it is thought that their biological functions may participate in formaldehyde detoxification (Harms et al., 1996). In addition, the previous studies have revealed that SFGHs can hydrolyze various non-thioester substrates, such as *p*-nitrophenyl esters (Gonzalez et al., 2006; van Straaten et al., 2009; Lee et al., 2019). SFGHs are serine hydrolases containing a conserved Ser-Asp-His catalytic triad with an  $\alpha/\beta$  hydrolase fold (Sun et al., 2014; Rauwerdink and Kazlauskas, 2015). Interestingly, a strictly conserved cysteine residue is located near the active site, and it is thought that the Cys residue may be involved in feedback inhibition by glutathione as a reaction product. This residue has also been proposed to serve a gatekeeping role in regulating the access of other substrates to the active site via disulfide formation with glutathione (Cummins et al., 2006; Legler et al., 2012; Chen et al., 2013).

In eukaryotes, several SFGHs have been characterized in the human liver, *Arabidopsis thaliana*, *Candida boidinii*, and *Saccharomyces cerevisiae* (Lee and Lee, 1986; Kordic et al., 2002; Yurimoto et al., 2003; Legler et al., 2008). Bacterial

SFGHs have also been identified and characterized in various species including *Paracoccus denitrificans*, *Escherichia coli*, *Agrobacterium tumefaciens*, and *Shewanella frigidimarina* (Harms et al., 1996; Gonzalez et al., 2006; van Straaten et al., 2009; Lee et al., 2019). Importantly, previous functional characterization of these enzymes shows that SFGHs have significantly different substrate specificities. For example, the SFGH from *A. tumefaciens* has greater affinity and catalytic efficiency toward C3 and C4 *p*-nitrophenyl substrates (van Straaten et al., 2009); SFGH from *E. coli* is only active toward  $\alpha$ -naphthyl acetate (Gonzalez et al., 2006); and the wild-type SFGH from *S. frigidimarina* (SfSFGH) prefers short-chain esters, such as *p*-nitrophenyl acetate. Mutation of Trp182 to Ala allowed SfSFGH to be active for substrates with a longer chain. It is thought that the W182A mutation increases the substrate-binding pocket and decreases the steric effect for larger substrates in SfSFGH (Lee et al., 2019). Collectively, these previous findings suggest that structural information on SFGHs is highly useful for protein engineering of bacterial SFGHs as industrial biocatalysts.

In this study, we cloned the SFGH gene from *Variovorax* sp. PAMC 28711 bacteria isolated from an Antarctic lichen. We then successfully overexpressed and purified VaSFGH. An enzyme activity assay using purified VaSFGH protein and *p*-nitrophenyl

esters shows that VaSFGH prefers a short acyl chain length between *p*-nitrophenyl ester substrates. To obtain structural information on VaSFGH, we also performed crystallization and initial X-ray crystallographic experiments. In the future, we expect that structural analysis of VaSFGH will provide useful information about its molecular mechanisms and substrate specificity as well as valuable directions for protein engineering.

## MATERIALS AND METHODS

### Expression and purification of recombinant VaSFGH protein

The gene encoding VaSFGH (NCBI Accession number: WP\_068632277.1) was synthesized with codon-optimized sequences (Bioneer, Korea). The synthesized gene fragment was amplified and ligated with the pET28a vector using *Nde*I and *Xho*I restriction enzymes. The recombinant plasmid encoding VaSFGH was transformed into *E. coli* BL21 (DE3) for protein expression (Table 1). Cell culture was conducted in 2 L of Luria Bertani medium with 50 µg/ml kanamycin at 37°C. When the optical density at 600 nm (OD<sub>600</sub>) reached 0.6, protein overexpression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were grown at 20°C for 24 h after induction. The cell culture was harvested by centrifugation at 6,000 rpm at 4°C for 20 min. The cell pellet was resuspended in cell lysis buffer containing 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 5 mM imidazole. Cells were disrupted by ultrasonication in ice, and the cell debris was removed by centrifugation at 16,000 rpm at 4°C for 50 min.

Ni-NTA agarose resin was packed in a column by gravity in advance, and then washed with distilled water and equilibrated using cell lysis buffer. The supernatant was loaded onto a column at a flow rate of 3 ml min<sup>-1</sup>. Proteins weakly bound to the Ni-NTA resin were subsequently washed with 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 30 mM imidazole. The poly-histidine-tagged protein was eluted with 20 mM Tris-HCl (pH 8.0), 200 mM NaCl and, 300 mM imidazole. The purified protein was then treated with 40 units of thrombin to remove the poly-histidine tag overnight at 4°C. Further size exclusion

chromatography was performed using HiLoad Superdex 200 pg column (Cytiva, USA) and new buffer condition with 20 mM Tris-HCl (pH 8.0) and 200 mM NaCl. The finally purified protein was concentrated up to 21 mg/ml at the volume of 5 ml using an Amicon-ultrafiltration filter unit with 10 kDa cutoff value. The purity of the protein was evaluated by 12% SDS-PAGE gel with a single band, and the concentrated protein was stored at -80°C.

### Crystallization and optimization

Crystallization was performed in a 96-well plate using the sitting-drop vapor diffusion method at 23°C. Six screening solution suites, MCSG I-IV (Anatrace, USA), JCSG, and Morpheus (Molecular Dimensions, UK), were used for primary screening. Next, 300 nl of the purified protein (21 mg/ml) was mixed with 300 nl of crystallization screening solution against 70 µl of reservoir solution using a ‘Mosquito’ crystallization robot (SPT Labtech, UK). The crystal was obtained in 0.2 M sodium chloride, 0.1 M Bis-Tris:HCl (pH 6.5), and 25% (w/v) PEG 3350 mixed solution (MCSG 1T #D2). To increase the size of the crystal, further optimization of the conditions was performed using the hanging-drop diffusion method. Subsequently, 1 µl of the purified protein was mixed with the same volume of the reservoir solution. The crystal for the X-ray diffraction experiment was obtained using 0.2 M sodium chloride, 0.1 M Bis-Tris:HCl (pH 6.5), and 20% (w/v) PEG 3350 mixed solution. The crystallization methods are summarized in Table 2.

### Data collection and processing

Before crystal mounting, the optimized single crystal was briefly incubated in the 20% final glycerol concentration mixed with reservoir solution for cryoprotection against liquid nitrogen. Then, the crystal was mounted on a synchrotron facility at beamline 5C (BL-5C) of the Pohang Accelerator Laboratory (PAL, Korea). For complete coverage, 360 diffraction images were obtained with 1° oscillation per frame using an Eiger 9M detector (Dectris, Switzerland). After collecting the dataset,

**TABLE 1 |** Recombinant VaSFGH protein production information

VaSFGH	
Source organism	<i>Variovorax</i> sp. PAMC 28711
DNA source	Genomic DNA
Cloning vector	pET28a
Expression host	<i>Escherichia coli</i> BL21 (DE3)
Amino acid sequence	GSHMTDSIKTL SAHRSFGGVQHFHEHAS-REIGLPMRFAAYLPPQAEHGKVPALLYLA-GLTCNEETFVMVKAGAQRLLAAELGIALIAP-DTSPRGAHIDGESTSWDFGVGAGFYLDATA-APWAPNWRMESYLVDELLPLLAKTLPIDGD-RIGVFGHSMGGHGALTLALRHPGLFKLSAFA-PICAPTQCPWGHKAFTGYLGADTTRWIEHDAT-VLMQHQPVAPYPAGILIDQGLADKFLAEQLH-PHLLLEDACRAIGQPLTLRRHEGYDHYGFVQSF-MADHLAHHAQILNGAIGRPAARQP

**TABLE 2 |** Crystallization details

Method	Vapor diffusion
Plate type for screening	96-well sitting drop MRC plate (Molecular dimension, UK)
Plate type for optimization	24-well hanging drop plate (Molecular dimension, UK)
Temperature (°C)	23
Protein concentration (mg/ml)	21
Composition of protein solution	20 mM Tris-HCl (pH 8.0), 200 mM NaCl
Composition of reservoir solution	0.2 M sodium chloride 0.1 M Bis-Tris:HCl (pH 6.5) 20% (w/v) PEG 3350
Volume and ratio of drop	2 µl, 1:1
Volume of reservoir (µl)	500

HKL-2000 software (HKL Research Inc., USA) was used for data-processing, indexing, merging, and scaling. To generate an electron density map, the molecular replacement method was used with a template structure (PDB code: 3FCX).

### Substrate specificity assay

*p*-Nitrophenyl esters are frequently used as substrates for the carboxylesterase activity of SFGH family proteins. Different acyl-chain-length derivatives, such as *p*-nitrophenyl acetate (*p*-NA), *p*-nitrophenyl butyrate (*p*-NB), *p*-nitrophenyl valerate (*p*-NV), *p*-nitrophenyl octanoate (*p*-NO), *p*-nitrophenyl decanoate (*p*-ND), and *p*-nitrophenyl palmitate (*p*-NP), were prepared to determine the substrate specificity of VaSFGH. The standard assay solution contained 20 mM Tris-HCl (pH 8.0) and 200 mM NaCl consistent with the protein sample buffer. In the reaction mixture, 5  $\mu$ g (at a final concentration of  $\sim$ 1.56  $\mu$ M) of VaSFGH and three different concentrations of substrate (at a final concentration of 100, 200, or 500  $\mu$ M) were used in the microplates at 25°C for 10 min. The product of enzymatic hydrolysis, *p*-nitrophenol, showed a yellow color and was detected at 405 nm.

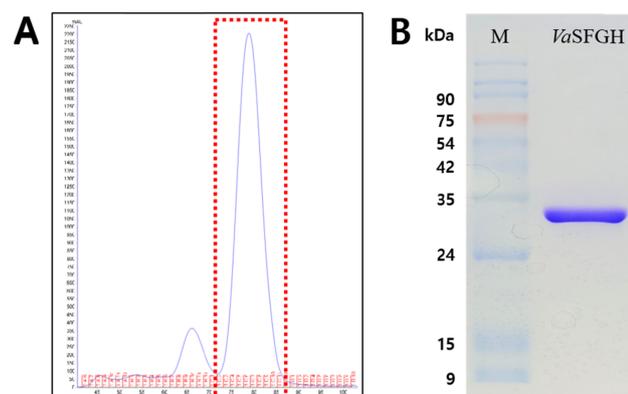
## RESULTS AND DISCUSSION

S-formylglutathione hydrolase protein is commonly found in prokaryotic and eukaryotic organisms and is thought to play an important role in the detoxification of formaldehyde. The VaSFGH gene was synthesized and cloned into the pET28a vector using *Nde*I and *Xho*I restriction enzymes. The overexpressed VaSFGH has an N-terminal poly-histidine tag and a thrombin protease recognition site to cleave the poly-histidine tag from the VaSFGH protein. Size-exclusion chromatography was performed to purify the VaSFGH and change the protein storage buffer after the Ni-NTA purification step and thrombin cleavage. Following several successful protein purification steps, a highly purified protein without a

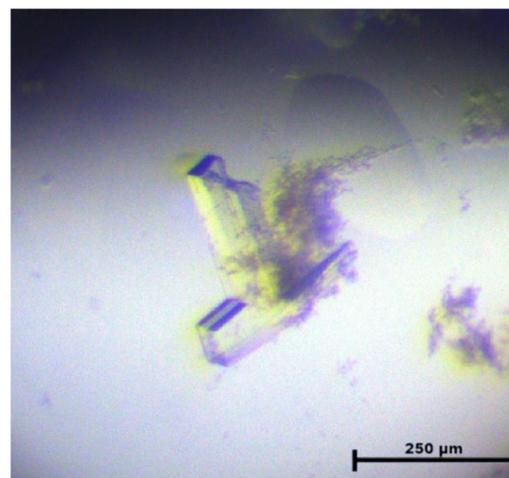
poly-histidine tag was obtained. The single protein band on SDS-PAGE gel indicated that the molecular weight of VaSFGH is approximately 32 kDa, corresponding to the molecular weight calculated based on a computed parameter using an amino acid sequence (Figure 1). The oligomeric state of VaSFGH in solution will be determined by analytical ultracentrifugation in our follow-up study. The purified protein was concentrated to 21 mg/ml using an ultrafiltration centrifugal filter tube. The protein sample for crystallization was successfully prepared by overexpression, purification, and concentration.

Initial crystal screening for VaSFGH was performed under 576 different conditions using commercial crystallization screening solution suites at 23°C. To optimize the crystal, the concentration of the precipitant in the reservoir solution was carefully controlled. The crystal for the X-ray diffraction experiment was obtained in 0.2 M sodium chloride, 0.1 M Bis-Tris:HCl (pH 6.5), and 20% (w/v) PEG 3350 mixed solution (Figure 2). A final concentration of 20% glycerol mixed with reservoir solution was used as a cryoprotectant to prevent freezing under liquid nitrogen. The full-coverage diffraction data were obtained at a resolution of 2.38 Å (Figure 3). The collected diffraction images were indexed, integrated, and scaled using the HKL2000 software package (Otwinowski and Miror, 1997). The crystal of VaSFGH belongs to the C2 space group with the following unit cell parameters:  $a = 53.2$  Å,  $b = 76.4$  Å,  $c = 199.9$  Å,  $\alpha = 90^\circ$ ,  $\beta = 90.2^\circ$ ,  $\gamma = 90^\circ$ . The two molecules were evaluated in an asymmetric unit as the volume of the unit cell is 812,625.94 Å<sup>3</sup> Da<sup>-1</sup>, and the Matthews coefficient value and solvent content were 3.17 and 61.21%, respectively (Matthews, 1968). The X-ray diffraction results are presented in Table 3.

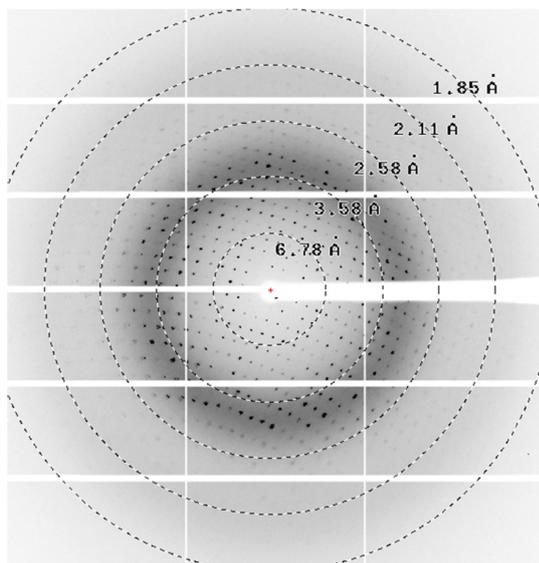
We determined the structure of the VaSFGH protein by molecular replacement using the CCP4i program suite (Vagin and Teplyakov, 1997). Esterase D from *Homo sapiens* (PDB



**FIGURE 1 | VaSFGH protein band on SDS-PAGE.** (A) The result of size-exclusion chromatography of VaSFGH. The red dotted box indicates the eluted peak of VaSFGH protein without poly-histidine tag. (B) The purified VaSFGH was diluted to 1 mg/ml and visualized on 12% SDS-PAGE.



**FIGURE 2 | Crystal image of the VaSFGH protein.** Crystals for diffraction experiment were obtained in 0.2 M sodium chloride, 0.1 M Bis-Tris:HCl (pH 6.5), and 20% (w/v) PEG 3350 at 23°C.



**FIGURE 3 | X-ray diffraction pattern of VaSFGH protein crystal.** Diffraction image of VaSFGH crystal with resolution guide shown. The dotted black circles and bold text represent the resolution ranges.

**TABLE 3 | X-ray diffraction data collection statistics**

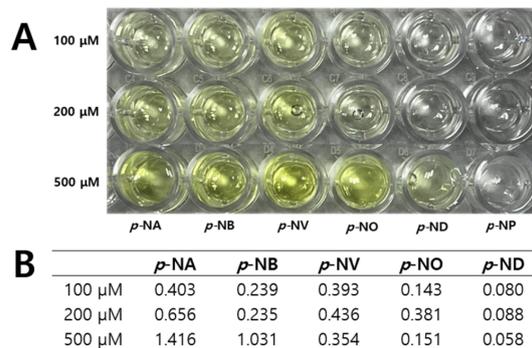
Data collection	
Wavelength (Å)	1
X-ray source	PAL 5C
Rotation range per image (°)	1
Exposure time (s)	0.1
Space group	C2
Unit-cell parameters (Å, °)	a = 53.2, b = 76.4, c = 199.9 α = 90, β = 90.2, γ = 90
Resolution range (Å) <sup>a</sup>	50–2.38 (2.42–2.38)
No. of observed reflections <sup>a</sup>	169,206 (167,843)
No. of unique reflections <sup>a</sup>	31,349 (1,671)
Completeness (%) <sup>a</sup>	95.7 (99.9)
Redundancy <sup>a</sup>	5.4 (6.2)
$R_{sym}^{a,b}$	0.141 (0.389)
$I/\sigma^a$	20.1 (4.2)

<sup>a</sup>Values in parentheses correspond to the highest resolution shell.  
<sup>b</sup> $R_{sym} = \sum_i \sum_j |I(h)_i - \langle I(h) \rangle| / \sum_i \sum_j I(h)_i$ , where  $I$  is the intensity of reflection  $h$ ;  $\sum_i$  is the sum over all reflections; and  $\sum_j$  is the sum over  $i$  measurements of reflection  $h$ .

code, 3FCX) was used as a template model, which shares a highly similar (52.48%) amino acid sequence identity (Wu et al, 2009). Model building is currently being performed using WinCoot (Emsley and Cowtan, 2004) along with structure refinement using Refmac5 (Murshudov et al., 2011) and Phenix.refinement (Adams et al., 2010).

**Substrate specificity of VaSFGH**

In general, SFGH-family enzymes are known to act as



**FIGURE 4 | Substrate specificity assay of VaSFGH.** Different lengths of *p*-nitrophenyl esters were used to assess carboxylesterase activity. (A) The *p*-nitrophenol shows a yellow color as a product. (B) The table shows the optical density (OD) values of *p*-nitrophenol detected at 405 nm wavelength. The OD values at three different concentration of each substrates are presented and the OD values were calculated with subtracting blank values (reaction buffer only).

carboxylesterases and thioesterases. *p*-Nitrophenyl esters are frequently used to measure carboxylesterase activity, and their substrate specificity is measured using different acyl-chain-length derivatives. Here, the substrate specificity of VaSFGH was measured using six different lengths of *p*-nitrophenyl esters, *p*-NA (C2), *p*-NB (C4), *p*-NV (C5), *p*-NO (C8), *p*-ND (C10), and *p*-NP (C16). Using colorimetric and spectrophotometric measurements, the amount of *p*-nitrophenol was measured with three different substrate concentrations (100, 200, and 500 μM) at 405 nm. The substrate preference of VaSFGH was shown for short acyl-chain-length substrates, and no activity was observed for long acyl-chain-length substrates, i.e., *p*-ND (C10) and *p*-NP (C16) (Figure 4).

**ACKNOWLEDGEMENTS**

We would like to thank the staff at the X-ray core facility of the Korea Basic Science Institute (KBSI; Ochang, Korea) and BL-5C of the Pohang Accelerator Laboratory (Pohang, Korea) for their kind help with data collection. This research was a part of the project titled “Development of potential antibiotic compounds using polar organism resources (15250103, KOPRI Grant PM21030)” funded by the Ministry of Oceans and Fisheries, Republic of Korea.

**CONFLICT OF INTEREST**

The authors declare no conflicts of interest.

Original Submission: Oct 21, 2021  
 Revised Version Received: Nov 13, 2021  
 Accepted: Nov 15, 2021

**REFERENCES**

Adams, P.D., Afonine, P.V., Bunkóczy, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve,

- R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.
- Chen, N.H., Couñago, R.M., Djoko, K.Y., Jennings, M.P., Apicella, M.A., Kobe, B., and McEwan, A.G. (2013). A glutathione-dependent detoxification system is required for formaldehyde resistance and optimal survival of *Neisseria meningitidis* in biofilms. *Antioxid Redox Signal* **18**, 743-755.
- Cummins, I., McAuley, K., Fordham-Skelton, A., Schwoerer, R., Steel, P.G., Davis, B.G., and Edwards, R. (2006). Unique regulation of the active site of the serine esterase S-formylglutathione hydrolase. *J Mol Biol* **359**, 422-432.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.
- Gonzalez, C.F., Proudfoot, M., Brown, G., Korniyenko, Y., Mori, H., Savchenko, A.V., and Yakunin, A.F. (2006). Molecular basis of formaldehyde detoxification. Characterization of two S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J Biol Chem* **281**, 14514-14522.
- Harms, N., Ras, J., Reijnders, W.N., van Spanning, R.J., and Stouthamer, A.H. (1996). S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *J Bacteriol* **178**, 6296-6299.
- Kordic, S., Cummins, I., and Edwards, R. (2002). Cloning and characterization of an S-formylglutathione hydrolase from *Arabidopsis thaliana*. *Arch Biochem Biophys* **399**, 232-238.
- Legler, P.M., Kumaran, D., Swaminathan, S., Studier, F.W., and Millard, C.B. (2008). Structural characterization and reversal of the natural organophosphate resistance of a D-type esterase, *Saccharomyces cerevisiae* S-formylglutathione hydrolase. *Biochemistry* **47**, 9592-9601.
- Legler, P.M., Leary, D.H., Hervey, W.J.T., and Millard, C.B. (2012). A role for His-160 in peroxide inhibition of *S. cerevisiae* S-formylglutathione hydrolase: evidence for an oxidation sensitive motif. *Arch Biochem Biophys* **528**, 7-20.
- Lee, C.W., Yoo, W., Park, S.H., Le, L., Jeong, C.S., Ryu, B.H., Shin, S.C., Kim, H.W., Park, H., Kim, K.K., Kim, T.D., and Lee, J.H. (2019). Structural and functional characterization of a novel cold-active S-formylglutathione hydrolase (SfSFGH) homolog from *Shewanella frigidimarina*, a psychrophilic bacterium. *Microb Cell Fact* **18**, 140.
- Lee, E.Y., and Lee, W.H. (1986). Molecular cloning of the human esterase D gene, a genetic marker of retinoblastoma. *Proc Natl Acad Sci U S A* **83**, 6337-6341.
- Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-497.
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-367.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326.
- Rauwerdink, A., and Kazlauskas, R.J. (2015). How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of  $\alpha/\beta$ -hydrolase fold enzymes. *ACS Catal* **5**, 6153-6176.
- Sun, Y., Yin, S., Feng, Y., Li, J., Zhou, J., Liu, C., Zhu, G., and Guo, Z. (2014). Molecular basis of the general base catalysis of an  $\alpha/\beta$ -hydrolase catalytic triad. *J Biol Chem* **289**, 15867-15879.
- Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J Appl Cryst* **30**, 1022-1025.
- van Straaten, K.E., Gonzalez, C.F., Valladares, R.B., Xu, X., Savchenko, A.V., and Sanders, D.A. (2009). The structure of a putative S-formylglutathione hydrolase from *Agrobacterium tumefaciens*. *Protein Sci* **18**, 2196-2202.
- Wu, D., Li, Y., Song, G., Zhang, D., Shaw, N., and Liu, Z.J. (2009). Crystal structure of human esterase D: a potential genetic marker of retinoblastoma. *FASEB J* **23**, 1441-1446.
- Yurimoto, H., Lee, B., Yano, T., Sakai, Y., and Kato, N. (2003). Physiological role of S-formylglutathione hydrolase in C(1) metabolism of the methylotrophic yeast *Candida boidinii*. *Microbiology (Reading)* **149**, 1971-1979.