Purification, crystallization, and preliminary X-ray analysis of a putative N-terminal acetyltransferase encoded by a *Salmonella*-infecting bacteriophage

Yongseong Hyun¹, Nayeon Ki¹, Nam-Chul Ha^{1*} and Hyun-Myung Oh^{2*}

¹Research Institute for Agriculture and Life Sciences, Center for Food and Bioconvergence, Department of Agricultural Biotechnology, CALS, Seoul National University, Seoul 08826, Republic of Korea

²Institute of Liberal Arts Education, Pukyong National University, Busan 48547, Republic of Korea

*Correspondence: marinebio@pknu.ac.kr, hanc210@snu.ac.kr

Bacteriophages infect host bacteria and control host metabolism through diverse mechanisms. We noted a gene harboring homology in the N-terminal acetyltransferase of a bacteriophage SPN3US, which infects *Salmonella*. The phage-encoded gene is phylogenetically remote from the typical N-terminal acetyltransferases whose structures are available. This study reports a preliminary analysis of the crystals of phage-encoded N-terminal acetyltransferase, including its purification and crystallization. The crystals diffracted X-rays to a 2.2 Å resolution, revealing that the crystals belong to *P*₆₂₂ or *P*₆₅₂₂ with unit cell parameters of *a* = 68.5 and *c* = 219.0 Å. We are now growing the selenomethionine-substituted crystals to obtain the phase information. The protein structure will provide molecular insights on how the phage hijacks the bacterial host metabolism.

INTRODUCTION

The emergence of antibiotic-resistant bacteria has driven bacteriophages (phages) as an alternative to conventional antibiotics (Burrowes et al., 2011). The narrow host range of the phages allows targeted bacterial killing without affecting the host beneficial microbiomes. The US FDA approved the use of phages as a feed additive and in the treatment of personalized bacteriophage therapeutics for serious MDR infections (Chan et al., 2013).

Phages infect bacteria via specific protein-protein interactions with host bacteria receptor molecules and then control host metabolism through diverse mechanisms using phage-encoded proteins (Weitz et al., 2013). For example, Srd protein of bacteriophage T4 stimulates activity of RNase E of *E. coli* to rapidly degrade host mRNAs (Qi et al., 2015). T4 Srh protein also resembles a segment of σ^{32} that interacts with host RNA polymerase (Mosig et al., 1998). *Salmonella enterica* is the most frequently identified human outbreak of food poisoning (Uzzau et al., 2000). Bacteriophage SPN3US was isolated from chicken feces using the host strain *Salmonella enterica* serovar Typhimurium LT2. The SPN3US phage belongs to the Myoviridae family, infecting host bacteria using flagella as receptors. Bacteriophages are of potential interest in preventing salmonellosis (Lee et al., 2011).

Many phage-encoded genes are involved in hijacking host metabolism to take advantage of phage proliferation in host bacteria. In the genome of SPN3US reported in 2011 (Lee et

al., 2011), we noted SPN3US_0088, annotated as a putative acetyltransferase (GenBank Accession AEP83921). Further analysis of the sequence showed that SPN3US_0088 (called SPN88 in this study) has sequence similarity to N-terminal acetyltransferases (NATs) (sequence identity: 21-23%). NATs belong to the superfamily of Gcn5-related N-acetyltransferases (GNAT), which contain a domain consisting of four conserved sequence motifs A-D. GNAT proteins catalyze acetyl transfer from the CoA donor to a primary amine of the acceptor substrates. However, the enzymatic activity of SPN88 remains unclear, with limited sequence similarity to typical GNAT proteins. The phage-encoded acetyltransferase is phylogenetically remote from the typical N-terminal acetyltransferases, whose crystal structures were determined.

Recent studies have revealed that bacterial protein acetylation plays a prominent role in central and secondary metabolism, virulence, transcription, and translation (VanDrissse and Escalante-Semerena, 2019). The role of SPN88 may be related to the modification of the host proteins to change the bacterial metabolism. As a first step to reveal the molecular function of the unique phage-encoded protein, this study reports the preliminary X-ray analysis of the crystals of SPN88, including the purification and crystallization steps.

RESULTS AND DISCUSSION

We overexpressed the full-length SPN88 protein in the *E. coli* expression system. The resulting protein contained the

Table 1 | SPN88 production information

SPN88	
Source organism	<i>Salmonella</i> bacteriophage SPN3US
DNA source	Genome synthesis, including the SacI and XhoI sites GAGCTC /ATGCTGGTGATTAACTGGTGGAGATAAAATTCGGCGAACGTGTATCCGGAACCTGGTGAATGGGTGCG CGATCTGAACAGCATTGCGAAGAACCGATTAACTGACCATGTTTGTGAAGATGATATTGCGCGGCATTATGGCGT GGGAACCGGGCCATCTGGTGATATGGTGGTGCCGGAAGAAAGCCGCCGCGCGGCGTGGGCCGCTTTATGCTGAA ATATCTGCAGCAGAACAGCGATCGCAACATGTGAGCTGCCGCGTGCATCCGACCAACATTCCGGCGCTGGGCTTTT TCATCAGCAGGGCTTCAGATTGATCGCTGGTATATTGCCGCGGATGGCCAGCGCTATTTTCGCATGACCAACTATAACG TGATTAGCAGCCATAACCCGCCGGAAGAAAACTGCTGACCCATTATGCGGAAAGCGTGCCGATTTTTCTGAGCATGGC GGAAGGCCAGTTTGTG/ CTCGAG
Forward primer	NA
Reverse primer	NA
Expression vector	pET24a(+)
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino acid sequence of the construct produced	MASMTGGQQMGRGSEFELMLVINVEDKIPANVYPELVEWVRDLNSIREEPIKLTMFVEDDIVRGIMAWEPGHLVYMWVPEES RRGGVGRFMLKYLQQNSDRKHVSCRHVPTNIPALGFFHQQGFQIDRWYIADGQRYFRMTNYNVISSHTPPEEKLLTHYAES VPIFLSMAEGQFVLEHHHHHH

FIGURE 1 | The protein band corresponding to SPN88 on a 15% sodium dodecyl sulfate-polyacrylamide gel. The proteins in the left lane indicate size (kDa).

hexahistidine tag at the C-terminus in an uncleavable form (Table 1). The protein was purified from the cell lysate using Ni-NTA affinity chromatography. The pooled fractions were subjected to size exclusion chromatography for further purification. The final protein sample was concentrated to 6 mg/ml, which was featured in a purity of 95% as judged by SDS-PAGE (Figure 1).

We tested approximately 700 premade crystallization conditions in an automated protein crystallization screening device MOSQUITO. Initially, small protein crystals were obtained in a solution containing 0.4 M sodium malonate (pH 6.0), 0.1 M MES monohydrate (pH 6.0) and 0.5% (w/v) polyethylene glycol

FIGURE 2 | Crystals of the SPN88 protein. The approximate dimensions of the crystal are 150 × 100 × 100 μm.

10,000 in a sitting drop method in a 96-well plate. We manually optimized the crystallization condition using the hanging-drop diffusion method in 15-well plates using the hanging-drop vapor diffusion method. The final composition of the reservoir solution was 0.4 M sodium malonate (pH 6.0), 0.1 M MES monohydrate (pH 6.0), and 0.4% (w/v) polyethylene glycol 10,000, which yielded hexagonal prism-shaped crystals (Figure 2).

To collect the X-ray diffraction patterns using the crystals under cryogenic conditions, we transferred the crystals into a reservoir solution supplemented with 25% (v/v) glycerol and then incubated the crystals for 30 s. The crystals were rapidly frozen in liquid nitrogen by a loop made of polyimide (MiTeGen, USA). The datasets were collected at microfocusing beamline 11C of the Pohang Accelerator Laboratory (Pohang, Republic

of Korea) at a wavelength of 0.97941 Å. The resulting dataset featured 95.1% completeness at 2.2 Å resolution (Figure 3 and Table 2). The crystal belonged to the Bravais lattice of primitive hexagonal. Analyses of diffraction patterns along the *h*, *k*, and *l* axes revealed a *P*6₂22 or *P*6₅22 space group when systemic absence along the *c* axis was considered. The unit cell of the SPN88 crystal was *a* = 68.5 Å and *c* = 219.0 Å. Cell content analyses (Adams et al., 2010) suggested that the asymmetric unit contained one molecule of SPN88 with solvent contents of 65.3%. Based on phenix.xtriage analyses (Zwart et al., 2005), no twinning was suggested.

The molecular replacement method was not successful, presumably because of limited sequence similarities to homologous proteins whose structures were deposited in the Protein Data Bank. We are currently growing selenomethionine-substituted crystals for phasing to resolve the crystal structure of SPN88.

MATERIALS AND METHODS

Construct design for protein purification

GNAT protein SPN3US_0088 (GenBank Accession AEP83921) was synthesized into the pTOPBlunt V2 vector by Macrogen, Inc. (Seoul, Republic of Korea). The synthesized nucleotides were subcloned in the pET24a(+) vector (Merck) using restriction enzymes *Sac*I and *Xho*I. The detailed sequence is listed in Table 2. The correctness of the cloning was checked by DNA sequencing (Macrogen, Inc. Daejeon, Republic of Korea).

Protein expression and purification

Escherichia coli BL21 (DE3) cells were used as the expression

host after transformation with the recombinant plasmid pET24a-SPN88. The cells were cultured in 1.5 L of Luria–Bertani (LB) medium supplemented with 50 µg/mL kanamycin at 37°C until an OD₆₀₀ of 0.6. The addition of 0.5 mM IPTG to the culture medium induced the expression of the SPN88 protein. The cells were further cultured at 30°C for 6 hours and harvested by centrifugation at 5,592 × *g* for 7 min at 4°C. The cell pellet was resuspended in 50 mL of lysis buffer containing 20 mM Tris (pH 8.0), 500 mM NaCl, and 2 mM β-mercaptoethanol. The cells were disrupted by sonication, and cell debris was removed by centrifugation at 18,800 × *g* for 30 min at 4°C.

Ni-NTA agarose resin (1 mL; Qiagen, Germany) was mixed with 50 mL of the cell lysate and gently rolled in a column for 50 min at 4°C. The resin was then washed with 250 mL of a buffer containing 20 mM Tris (pH 8.0), 500 mM NaCl, 20 mM imidazole (pH 8.0), and 2 mM β-mercaptoethanol. The protein was eluted with 20 mL of a buffer containing 20 mM Tris (pH 8.0), 500 mM NaCl, 250 mM imidazole (pH 8.0), and 2 mM β-mercaptoethanol. The pooled fractions were further purified by gel filtration chromatography (HiLoad™ 26/610 Superdex™ 200 pg; GE Healthcare, USA). Before chromatography, the column was pre-equilibrated with 20 mM Tris buffer (pH 8.0) containing 500 mM NaCl and 2 mM β-mercaptoethanol. The resulting protein solution was concentrated to 6 mg/mL using a Vivaspin centrifugal concentrator (30 kDa molecular-weight cutoff; Millipore, USA) and stored frozen at –80°C until use.

Crystallization

An automated crystal screening device (MOSQUITO Crystal; TTP Labtech, UK) was employed for the initial crystallization trials. The purified SPN88 protein was initially crystallized at 14°C using a

FIGURE 3 | A representative diffraction image. The resolution circles for 2.2 Å and 2.5 Å are shown as dotted lines.

Table 2 | X-ray diffraction and cell content analysis

SPN88	
Data collection	
Beamline	PAL 11C
Wavelength (Å)	0.97941
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	0.1
Space group	<i>P</i> 6 ₂ 22 or <i>P</i> 6 ₅ 22
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	68.5, 68.5, 219.0
Resolution (Å)	50.0–2.20 (2.24–2.20)
Total No. reflections	15,921
<i>R</i> _{merge}	0.077 (0.324)
High resolution shell CC1/2	0.959
<i>I</i> /σ <i>I</i>	36.7 (5.5)
Completeness (%)	98.7 (95.1)
Redundancy	21.1 (11.3)

*The values in parentheses are for the highest resolution shell.

sitting-drop vapor-diffusion method. The protein solution (0.2 μ L, 6 mg/mL) was mixed with a reservoir solution (0.2 μ L) and then equilibrated against 60 μ L of the reservoir solution in a 96-well crystallization plate. We used seven matrix screening solutions (MCSG 1-4T, PEG RX, PEG Ion, and Hampton Research Crystal Screen) in the initial screening trials. Crystals grew in a reservoir solution containing 0.4 M sodium malonate (pH 6.0), 0.1 M MES monohydrate (pH 6.0), and 0.5% (w/v) polyethylene glycol 10,000. The crystallization conditions were optimized using the hanging-drop diffusion method at 14°C in 15-well crystallization plates. SPN88 was crystallized in solution containing 0.4 M sodium malonate (pH 6.0), 0.1 M MES monohydrate (pH 6.0), and 0.4% (w/v) polyethylene glycol 10,000.

Data collection and processing

For data collection under cryogenic conditions, SPN88 crystals were incubated in reservoir solution supplemented with 25% (v/v) glycerol for 30 s. Crystals were flash-frozen in liquid nitrogen at −196°C. The datasets were collected using the Pilatus3 6M detector (Dectris, Switzerland) on beamline 11C of the Pohang Accelerator Laboratory, Republic of Korea. The diffraction datasets were processed, merged, and scaled using the program HKL-2000 (Otwinowski and Minor, 1997). Table 2 shows the data collection statistics.

ACKNOWLEDGEMENTS

This research was supported by the Collaborative Genome Program for Fostering New Post-Genome Industry of the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (MOF) (No. 20180430).

Original Submission: Feb 7, 2021

Revised Version Received: Mar 12, 2021

Accepted: Mar 16, 2021

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczi, 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.
- Burrowes, B., Harper, D.R., Anderson, J., McConville, M., and Enright, M.C. (2011). Bacteriophage therapy: potential uses in the control of antibiotic-resistant pathogens. *Expert Rev Anti Infect Ther* **9**, 775–785.
- Chan, B.K., Abedon, S.T., and Loc-Carrillo, C. (2013). Phage cocktails and the future of phage therapy. *Future Microbiol* **8**, 769–783.
- Lee, J.H., Shin, H., Kim, H., and Ryu, S. (2011). Complete genome sequence of Salmonella bacteriophage SPN3US. *J Virol* **85**, 13470–13471.
- Mosig, G., Colowick, N.E., and Pietz, B.C. (1998). Several new bacteriophage T4 genes, mapped by sequencing deletion endpoints between genes 56 (dCTPase) and dda (a DNA-dependent ATPase-helicase) modulate transcription. *Gene* **223**, 143–155.
- Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
- Qi, D., Alawneh, A.M., Yonesaki, T., and Otsuka, Y. (2015). Rapid degradation of host mRNAs by stimulation of RNase E activity by Srd of bacteriophage T4. *Genetics* **201**, 977–987.
- Uzzau, S., Brown, D.J., Wallis, T., Rubino, S., Leori, G., Bernard, S., Casadesús, J., Platt, D.J., and Olsen, J.E. (2000). Host adapted serotypes of Salmonella enterica. *Epidemiol Infect* **125**, 229–255.
- VanDrise, C.M., and Escalante-Semerena, J.C. (2019). Protein acetylation in bacteria. *Annu Rev Microbiol* **73**, 111–132.
- Weitz, J.S., Poisot, T., Meyer, J.R., Flores, C.O., Valverde, S., Sullivan, M.B., and Hochberg, M.E. (2013). Phage-bacteria infection networks. *Trends Microbiol* **21**, 82–91.
- Zwart, P.H., Grosse-Kunstleve, R.W., and Adams, P.D. (2005). Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. *CCP4 Newsl* **43**, 27–35.