Crystallization and preliminary X-ray analysis of the C-terminal fragment of Ski7 from *Saccharomyces cerevisiae*

Ji-Young Lee, Si Hoon Park,
Byung-Cheon Jeong[‡] and
Hyun Kyu Song*

Department of Life Sciences, Korea University,
145 Anam-ro, Seongbuk-gu, Seoul 136-701,
Republic of Korea

[‡] Present address: Department of Biophysics,
Department of Obstetrics and Gynecology,
UT Southwestern Medical Center, Dallas,
TX 75390, USA.

Correspondence e-mail: hksong@korea.ac.kr

Received 4 July 2014
Accepted 22 July 2014

Ski7 (superkiller protein 7) plays a critical role in the mRNA surveillance pathway. The C-terminal fragment of Ski7 (residues 520–747) from *Saccharomyces cerevisiae* was heterologously expressed in *Escherichia coli* and purified to homogeneity. It was successfully crystallized and preliminary X-ray data were collected to 2.0 Å resolution using synchrotron radiation. The crystal belonged to a trigonal space group, either $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 73.5$, $c = 83.6$ Å. The asymmetric unit contains one molecule of the C-terminal fragment of Ski7 with a corresponding crystal volume per protein mass (V_M) of $2.61 \text{ Å}^3 \text{ Da}^{-1}$ and a solvent content of 52.8% by volume. The merging R factor is 6.6%. Structure determination by MAD phasing is under way.

1. Introduction

In eukaryotic cells, RNA degradation is related to the processing, turnover and surveillance of all RNAs and thus plays a critical role in gene expression (Halbach *et al.*, 2013). Errors in gene transcription result in aberrant transcripts with premature stop codons or rare codons or in transcripts without stop codons (Klauer & van Hoof, 2012; Frischmeyer *et al.*, 2002). The process of mRNA surveillance ensures that aberrant mRNAs are degraded rapidly (Hilleren & Parker, 1999) and also prevents the synthesis of incomplete and potentially deleterious proteins (Jungwirth *et al.*, 2001). To minimize the formation of aberrant transcripts, different surveillance mechanisms are activated (Isken & Maquat, 2007). When a transcript lacks a stop codon, the ribosome stalls at the 3'-end of the mRNA (Maquat, 2002), triggering nonstop decay, which is a degradative process that proceeds in the 3'-to-5' direction and requires the cytoplasmic exosome, the associated Ski complex (Ski2/3/8) and superkiller protein 7 (Ski7) (Butler, 2002; Frischmeyer *et al.*, 2002).

The Ski complex consists of the RNA helicase Ski2 and the tetratricopeptide-repeat proteins Ski3 and Ski8 (Anderson & Parker, 1998). The genes encoding these superkiller proteins, the *ski* genes, were first identified by mutations causing overexpression of a lethal toxin that is encoded by double-stranded satellite RNA in virus-like particles (Toh-E *et al.*, 1978). Among the cofactors of the exosome, Ski7 is of particular interest because the two independent regions in the unique N-terminal domain of Ski7 have been shown to interact with both the exosome and the Ski complex (Araki *et al.*, 2001; Fig. 1a). Ski7 is a member of the family of eukaryotic elongation factor 1 α (eEF1A)-like GTP-binding proteins. The C-terminal domain of Ski7 (Fig. 1a) recognizes the completely or partially codon-free A site of the ribosome and is responsible for recruiting the exosome as well as the Ski complex (Ski2/3/8) to the 3'-end of nonstop mRNAs when degradation is initiated (Isken & Maquat, 2007; Atkinson *et al.*, 2008). The C-terminal domain of eEF1A family proteins is further divided into a GTPase domain, domain 2 and domain 3 (Fig. 1a). Although the C-terminal domain of Ski7 has characteristics similar to those of eRF3 and Hbs1, which form heterodimers with eRF1 and Dom34 (Paushkin *et al.*, 1997; Merkulova *et al.*, 1999; Ebihara & Nakamura, 1999; Taylor *et al.*, 2012; Chen *et al.*, 2010; Becker *et al.*, 2011; Lee *et al.*, 2007, 2010), no interactions with the C-terminal domain of Ski7 have been reported. Full-length

Ski7 has relatively low similarity to the related proteins eRF3 (from *Schizosaccharomyces pombe*; 15% identity) and Hbs1 (from *Saccharomyces cerevisiae*; 9% identity); in addition, there is no significant similarity between the C-terminal domains of these proteins (Fig. 1*b*). To date, biochemical and structural characterization of Ski7 is limited. Here, we report the expression, purification, crystallization and initial X-ray diffraction analysis of Ski7 (residues 520–747; Ski7C) from *S. cerevisiae*.

2. Materials and methods

2.1. Cloning

The construct of Ski7C from *S. cerevisiae* genomic DNA was PCR-amplified using the primers 5'-CGCGGATCCGAAACAACCTTGGAAGAGCCATTG-3' (*Bam*HI site in bold) and 5'-CTGCTGCTCGAGCTATTACTGGCATGCAATTCTGC-3' (*Xho*I site in bold). The PCR products were digested with the *Bam*HI and *Xho*I enzymes and were ligated into *Bam*HI- and *Xho*I-digested pRSF-GST vector. This produced a 460-amino-acid protein that contained the GST (glutathione *S*-transferase) tag and TEV (*Tobacco etch virus*) protease cleavage sequence (EDLYFQ). The molecular weight and calculated isoelectric point (pI) of Ski7C are 25 618 Da and 6.32, respectively.

2.2. Expression and purification of Ski7C

Ski7C was heterologously expressed in *Escherichia coli* BL21 (DE3) cells grown at 310 K to an OD_{600 nm} of 0.8 in Luria broth with kanamycin (50 µg ml⁻¹). Ski7C expression was induced by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside followed by overnight incubation at 291 K. After overnight growth, the cells were harvested by centrifugation at 7000 rev min⁻¹ for 20 min and the cell pellet was resuspended in PBS (phosphate-buffered saline) containing 1 mM 1,4-dithiothreitol (DTT). The cells were disrupted by ultrasonication

with 1 mM phenylmethylsulfonyl fluoride and one tablet of protease-inhibitor cocktail (Roche) and the lysate was centrifuged at 15 000 rev min⁻¹ for 2 h at 277 K. The supernatant was loaded onto GST resin (GE Healthcare) which had been pre-equilibrated with PBS. After washing the resin extensively with PBS, the protein was eluted from the resin with elution buffer (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM DTT, 10 mM reduced glutathione). The eluted protein was diluted with Q binding buffer (50 mM Tris-HCl pH 8.5, 0.5 mM TCEP) and applied onto an anion-exchange chromatography column (HiTrap Q, GE Healthcare). The target protein was eluted with a gradient of elution buffer (50 mM Tris-HCl pH 8.5, 0.5 mM TCEP, 1 M NaCl). After anion-exchange chromatography, TEV protease was added at a 1:25(w:w) ratio to the fusion protein and incubated overnight at 295 K in order to cleave the GST tag from the target protein. The protein digested with TEV protease was loaded onto Ni Sepharose HisTrap and GST beads to eliminate TEV protease containing the His tag and uncleaved GST-fused target protein, respectively. The target protein, as analyzed by 15%(w/v) SDS-PAGE, was collected and concentrated to about 2 ml using a 10.0 kDa cutoff Amicon Ultra centrifugal filter device (Millipore) at 2800 rev min⁻¹. The concentrated protein was further purified using a HiLoad 16/60 Superdex 75 pg column (GE Healthcare) pre-equilibrated with gel-filtration buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM TCEP). Fractions containing the target protein were confirmed by 15%(w/v) SDS-PAGE with Coomassie Blue staining.

2.3. Crystallization

The purified and concentrated protein (8 mg ml⁻¹) was used for initial crystallization screening. The hanging-drop vapour-diffusion method was performed using commercially available kits from Hampton Research, Emerald Bio and Molecular Dimensions (Boyle *et al.*, 1989). 1 µl reservoir solution and 1 µl protein solution were mixed for crystallization on 48-well VDX plates. Optimization of the

Figure 1
(a) Schematic representation of the domain organization of Ski7 from *S. cerevisiae*, eRF3 from *S. pombe* and Hbs1 from *S. cerevisiae*. The GTPase domain is shown in blue and domains 2 and 3 of the eEF1A-like GTP-binding domain are coloured green and orange, respectively. Unlike the eEF1A-like GTP-binding domains, the N-domains are species-specific. Two independent regions within the N-domain of Ski7, shown in yellow, are known to interact with the exosome and the Ski complex (Ski2/3/8). The N-domains of eRF3 and Hbs1 are coloured red and emerald. (b) Sequence comparison among the above proteins. All sequence identities are less than 20% among the proteins over the full length as well as in the C-terminal domains 2 and 3.

crystallization procedure was performed as follows: 1.5 µl protein solution was mixed with the same volume from 120 µl reservoir solution using the hanging-drop vapour-diffusion method on 24-well VDX plates. For improving the Ski7C crystals, an additive screen was performed with Hampton Research additive screen kits (Cudney & Patel, 1994). The crystallization plates were incubated at 292 K.

2.4. X-ray data collection

The optimized crystals were transferred from the mother liquor to a cryoprotectant solution consisting of the reservoir solution with 26–30% (v/v) ethylene glycol. A CryoLoop (Hampton Research; 20 µm; approximately 0.15 mm in size) was used for crystal mounting. All X-ray experiments were performed under a liquid-nitrogen stream generated by an Oxford Cryojet at 100 K. An ADSC Q315r charge-coupled device detector on beamline 5C at the Pohang Accelerator Laboratory (PAL, Republic of Korea) was used to collect diffraction data from the Ski7C crystals. The collected data were processed and scaled using the *HKL-2000* software package (Otwinowski & Minor, 1997). The crystals diffracted to 2.0 Å resolution (Table 1).

Figure 2 SEC-MALS results for Ski7C. The solid and dotted lines represent the gel-filtration elution profile and the measured molecular weight, respectively. The experimentally determined molecular weight of Ski7C is 32.6 kDa (calculated as 25.6 kDa), which shows the presence of a monomer in solution. The peak represents pure Ski7C as shown by the SDS-PAGE results. The fractions indicated by numbers (1, 2, 3) were collected for crystallization. Lane M contains molecular-weight markers (labelled in kDa): BSA (66 kDa), MBP-K8 (45 kDa), His-Gal8 (35.3 kDa), His-TEV (29 kDa), GST (24 kDa), His-Gal8N (21 kDa), Gal8N (17.5 kDa), lysozyme (14 kDa) and ubiquitin (8.5 kDa).

Figure 3 Crystals of Ski7C. The scale bar represents 100 µm.

Table 1

Data-collection statistics.

Values in parentheses are for the highest resolution bin.

X-ray source	5C, PAL
Wavelength (Å)	0.97933
Space group	<i>P</i> ₃ ₁ 21 or <i>P</i> ₃ ₂ 21
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 73.5, <i>c</i> = 83.6
Resolution (Å)	50–2.00 (2.06–2.00)
Mosaicity (°)	0.419
Unique reflections	18060
Total reflections	133828
Completeness (%)	99.3 (100)
Multiplicity	7.4 (7.7)
Average <i>I</i> /σ(<i>I</i>)	50.42 (4.02)
<i>R</i> _{merge} † (%)	6.6 (65.7)
<i>R</i> _{p.i.m.} ‡ (%)	2.7 (25.4)
CC _{1/2}	0.97 (0.82)

† *R*_{merge} = $\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where *I_i(hkl)* is the intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl) \rangle$ is the corresponding average value for all *i* measurements. ‡ *R*_{p.i.m.} (the precision-indicating merging *R* factor) = $\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

3. Results and discussion

Ski7C was heterologously expressed in *E. coli* and purified. The yield of purified Ski7C was about 1.3 mg per litre of bacterial culture. Domains 2 and 3 of eEF1A family proteins are involved in protein–protein interactions to form heterodimers (Pittman *et al.*, 2009; Cho *et al.*, 2012). In addition, it should be noted that domain 2 of human elongation factor 1Bγ forms a homodimer in solution (Vanwetswinkel *et al.*, 2003). To determine whether Ski7C forms a dimer, a size-exclusion chromatography with multi-angle light scattering (SEC-MALS) experiment was performed, which showed that Ski7C behaves as a monomer in solution (Fig. 2).

Purified Ski7C was successfully crystallized within a day (Fig. 3) in a buffer consisting of 0.1 M HEPES pH 7.5, 10% (w/v) polyethylene glycol 8000, 8% (w/v) ethylene glycol. Preliminary X-ray data at 2.0 Å resolution were collected. The crystals belonged to a trigonal space group, either *P*₃₁21 or *P*₃₂21, with unit-cell parameters *a* = *b* = 73.5, *c* = 83.6 and a merging *R* factor of 6.6%. Table 1 shows the statistics of the data collected using synchrotron radiation. The Matthews coefficient (*V_M*) of 2.61 Å³ Da^{−1} suggests the presence of one molecule per asymmetric unit, with a solvent content of 52.8% (Matthews, 1968). We are currently trying to solve the crystal structure using molecular replacement (MR; Bunkóczi *et al.*, 2013). Although the sequence identity is not high (Fig. 1*b*), domains 2 and 3 of Hbs1 and eRF3 are predicted to share structural homology with those of Ski7C; therefore, the MR technique could be applied to obtain initial phases using the structures of these proteins (Hbs1 and eRF3; PDB entries 3p26 and 1r5b; van den Elzen *et al.*, 2010; Kong *et al.*, 2004). However, the MR solution is so far ambiguous. Therefore, further attempts to solve the structure of Ski7C using MAD phasing are needed.

We are grateful to the beamline staff at beamline 5C of the Pohang Accelerator Laboratory, Pohang, Republic of Korea for their help during data collection. This study was supported by National Research Foundation of Korea (NRF) grants funded by the Korean government (NRF-2013R1A1A2020177).

References

- Anderson, J. S. & Parker, R. P. (1998). *EMBO J.* **17**, 1497–1506.
- Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S. & Katada, T. (2001). *EMBO J.* **20**, 4684–4693.
- Atkinson, G. C., Baldauf, S. L. & Hauryliuk, V. (2008). *BMC Evol. Biol.* **8**, 290.
- Becker, T., Armache, J.-P., Jarasch, A., Anger, A. M., Villa, E., Sieber, H., Motaal, B. A., Mielke, T., Berninghausen, O. & Beckmann, R. (2011). *Nature Struct. Mol. Biol.* **18**, 715–720.

- Boyle, A., Chapman, A., Mulrone, S., McPherson, R. & Bailey, S. (1989). *Kidney Int.* **35**, 222.
- Bunkóczi, G., Echols, N., McCoy, A. J., Oeffner, R. D., Adams, P. D. & Read, R. J. (2013). *Acta Cryst.* **D69**, 2276–2286.
- Butler, J. S. (2002). *Trends Cell Biol.* **12**, 90–96.
- Chen, L., Muhrad, D., Haurlyuk, V., Cheng, Z., Lim, M. K., Shyp, V., Parker, R. & Song, H. (2010). *Nature Struct. Mol. Biol.* **17**, 1233–1240.
- Cho, S.-J., Lee, H., Dutta, S., Seog, D.-H. & Moon, I. S. (2012). *BMB Rep.* **45**, 227–232.
- Cudney, B. & Patel, S. (1994). *Am. Biotechnol. Lab.* **12**, 42.
- Ebihara, K. & Nakamura, Y. (1999). *RNA*, **5**, 739–750.
- Elzen, A. M. van den, Henri, J., Lazar, N., Gas, M. E., Durand, D., Lacroute, F., Nicaise, M., van Tilbeurgh, H., Séraphin, B. & Graille, M. (2010). *Nature Struct. Mol. Biol.* **17**, 1446–1452.
- Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrero, A. L., Parker, R. & Dietz, H. C. (2002). *Science*, **295**, 2258–2261.
- Halbach, F., Reichelt, P., Rode, M. & Conti, E. (2013). *Cell*, **154**, 814–826.
- Hilleren, P. & Parker, R. (1999). *RNA*, **5**, 711–719.
- Isken, O. & Maquat, L. E. (2007). *Genes Dev.* **21**, 1833–1856.
- Jungwirth, H., Bergler, H. & Högenauer, G. (2001). *J. Biol. Chem.* **276**, 36419–36424.
- Klauser, A. A. & van Hoof, A. (2012). *Wiley Interdiscip. Rev. RNA*, **3**, 649–660.
- Kong, C., Ito, K., Walsh, M. A., Wada, M., Liu, Y., Kumar, S., Barford, D., Nakamura, Y. & Song, H. (2004). *Mol. Cell*, **14**, 233–245.
- Lee, H. H., Jang, J. Y., Yoon, H.-J., Kim, S. J. & Suh, S. W. (2010). *Biochem. Biophys. Res. Commun.* **399**, 600–606.
- Lee, H. H. et al. (2007). *Mol. Cell*, **27**, 938–950.
- Maquat, L. E. (2002). *Science*, **295**, 2221–2222.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 499–501.
- Merkulova, T. I., Frolova, L. Y., Lazar, M., Camonis, J. & Kisselev, L. L. (1999). *FEBS Lett.* **443**, 41–47.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Paushkin, S. V., Kushnirov, V. V., Smirnov, V. N. & Ter-Avanesyan, M. D. (1997). *Mol. Cell. Biol.* **17**, 2798–2805.
- Pittman, Y. R., Kandl, K., Lewis, M., Valente, L. & Kinzy, T. G. (2009). *J. Biol. Chem.* **284**, 4739–4747.
- Taylor, D., Unbehaun, A., Li, W., Das, S., Lei, J., Liao, H. Y., Grassucci, R. A., Pestova, T. V. & Frank, J. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 18413–18418.
- Toh-E, A., Guerry, P. & Wickner, R. B. (1978). *J. Bacteriol.* **136**, 1002–1007.
- Vanwetswinkel, S., Kriek, J., Andersen, G. R., Güntert, P., Dijk, J., Canters, G. W. & Siegal, G. (2003). *J. Biol. Chem.* **278**, 43443–43451.