electronic reprint

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Crystallographic analysis of *Eisenia* hydrolysis-enhancing protein using a long wavelength for native-SAD phasing

Xiaomei Sun, Yuxin Ye, Naofumi Sakurai, Koji Kato, Keizo Yuasa, Akihiko Tsuji and Min Yao

Acta Cryst. (2020). F**76**, 20–24

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystallographic analysis of *Eisenia* hydrolysis-enhancing protein using a long wavelength for native-SAD phasing

Xiaomei Sun,[a] Yuxin Ye,[b] Naofumi Sakurai,[b] Koji Kato,[b] Keizo Yuasa,[c] Akihiko Tsuji[c] and Min Yao[a,b]*

[a]Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan, [b]Faculty of Advanced Life Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan, and [c]Graduate School of Bioscience and Bioindustry, Tokushima University, Tokushima 770-8506, Japan. *Correspondence e-mail: yao@castor.sci.hokudai.ac.jp

Received 29 November 2019
Accepted 13 December 2019

Edited by I. Tanaka, Hokkaido University, Japan

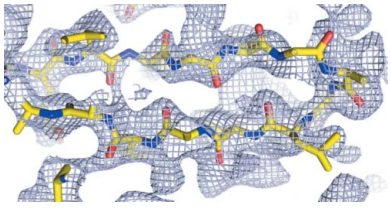
Keywords: EHEP; phlorotannin binding; solutionless crystal mount; native-SAD; biofuel.

Eisenia hydrolysis-enhancing protein (EHEP), which is a novel protein that has been identified in *Aplysia kurodai*, protects β-glucosidases from phlorotannin inhibition to facilitate the production of glucose from the laminarin abundant in brown algae. Hence, EHEP has attracted attention for its potential applications in producing biofuel from brown algae. In this study, EHEP was purified from the natural digestive fluid of *A. kurodai* and was crystallized using the sitting-drop vapor-diffusion method. Native and SAD (single-wavelength anomalous diffraction) data sets were successfully collected at resolutions of 1.20 and 2.48 Å using wavelengths of 1.0 and 2.1 Å, respectively, from crystals obtained in initial screening. The crystals belonged to space group $P2_12_12_1$ and contained one EHEP molecule in the asymmetric unit. All 20 S-atom sites in EHEP were located and the phases were determined by the SAD method using the S atoms in the natural protein as anomalous scatterers (native-SAD). After phase improvement, interpretable electron densities were obtained and 58% of the model was automatically built.

1. Introduction

Dramatic global increases in fuel demands have prompted a search for renewable energy sources. Biofuels are a promising alternative to fossil fuels because of their lower cost, renewable supply and reduced greenhouse-gas emissions (Ellabban *et al.*, 2014). Brown algae are considered to be ideal feedstocks for producing biofuels (Enquist-Newman *et al.*, 2014) because they contain large amounts of laminarin that can be digested to glucose (Wei *et al.*, 2013) and further converted to the biofuel alcohol by the fermentation process (Demirbas, 2011).

The sea hare *Aplysia kurodai* consumes brown algae as a staple food, using β-glucosidases (*aku*BGLs) as catalysts to hydrolyze the glycosidic bonds to produce glucose (Tsuji *et al.*, 2013) and making it an excellent model for investigating the biofuel production process. However, phlorotannin, which is also abundant in brown algae (La Barre *et al.*, 2010), inhibits the hydrolytic reaction of *aku*BGLs (Tsuji *et al.*, 2017). Moreover, previous studies have reported that phlorotannin also inhibits the activity of several other enzymes, such as hyaluronidase (Shibata *et al.*, 2002), α-glucosidase (Kellogg *et al.*, 2014) and α-amylase (Kellogg *et al.*, 2014). Interestingly, *Eisenia* hydrolysis-enhancing protein (EHEP), which was recently identified in the digestive fluid of *A. kurodai*, protects *aku*BGLs from this inhibition by binding and then precipitating with phlorotannin, despite not possessing a catalytic



function (Tsuji *et al.*, 2017). Furthermore, it has been shown that EHEP is a novel and unique protein, with no proteins with homologous amino-acid sequences occurring in any other organism, and it also binds to analogs of phlorotannin such as gallic acid and tannic acid (Tsuji *et al.*, 2017). Consequently, EHEP has a high potential for application in the biofuel industry. However, EHEP precipitates after binding to phlorotannin and cannot be recycled at present, limiting its potential application in industry.

An understanding of the ligand-binding mechanism of EHEP could provide information on how it can be recycled, and knowledge of the three-dimensional structure of EHEP will be indispensable for this. In this study, we conducted crystallographic analysis of EHEP. Because EHEP is difficult to crystallize reproducibly and is a novel and unique protein, we attempted to use the crystals obtained in initial screening to determine the structure of EHEP by single-wavelength anomalous dispersion using the S atoms in the native protein as anomalous scatterers (native-SAD) using a long wavelength.

2. Materials and methods

2.1. Preparation and characterization of EHEP

Native EHEP (25 kDa; GenBank BAV38197.1; Table 1) was purified from the natural digestive fluid of *A. kurodai* as described previously (Tsuji *et al.*, 2017). This protein was then further purified using size-exclusion chromatography (SEC) on a Superdex 200 10/300 GL column (GE Healthcare) which was equilibrated with 20 m*M* sodium acetate buffer containing 100 m*M* sodium chloride pH 6.0. The purified EHEP was examined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and was concentrated to 15–25 mg ml^{-1} using Vivaspin-4 columns (Sartorius, Göttingen, Germany) with a molecular-weight cutoff of 5.0 kDa.

A previous study predicted that native EHEP possesses a signal peptide in its N-terminus (Tsuji *et al.*, 2017). To confirm the molecular weight of mature native EHEP, MALDI-TOF mass-spectrometric analysis of purified EHEP was carried out using an Autoflex III Smartbeam mass spectrometer (Bruker). After SEC purification, the EHEP was demineralized using ZipTip (Merck) pipette tips packed with C4 resin. The sample for MALDI-TOF mass spectrometry was then prepared by mixing 1 µl EHEP with 1 µl matrix (sinapinic acid). The sample was applied onto the MALDI target plates and dried with a drier for several minutes. Spectra were then acquired and analyzed using the *Flex Analysis* software.

2.2. Crystallization

Initial crystallization screening was performed using the sitting-drop vapor-diffusion method with the JCSG Core I–IV Suites and the PEGs Suite crystallization kits (Qiagen, Hilden, Germany) at 20°C. Protein solution (1 µl) was mixed with an equal volume of reservoir solution. Initial crystals were obtained under condition No. 48 [0.2 *M* lithium sulfate, 20% polyethylene glycol (PEG) 3350] from the JCSG Core II Suite, condition No. 7 (0.2 *M* disodium hydrogen phosphate, 20% PEG 3350) from the JCSG Core III Suite and condition No. 40 (0.1 *M* HEPES pH 7.5, 25% PEG 8000) from the PEGs Suite. However, crystal formation could not be replicated under these conditions. Therefore, crystallization screening was reattempted using the Crystal Screen crystallization kit (Hampton Research, USA), which resulted in crystals appearing under condition No. 25 (1.0 *M* sodium acetate, 0.1 *M* imidazole pH 6.5) and condition No. 37 (8% PEG 4000, 0.1 *M* sodium acetate). Unfortunately, the crystallization conditions could not be further improved owing to the poor reproducibility of crystallization. The detailed crystallization conditions for the diffraction experiments are summarized in Table 2.

Table 1
EHEP production information.

Source organism	*A. kurodai*
Complete amino-acid sequence	MVTKVLLVSLALFALGTCQVAVNLCTQYGWPNGN YPDPYDCRKYISCNGAVATVMSCALGTVFNPN TRNCDAYGNVPICQYALPSPIVVTNICNQYGW GNGNFYHPYNCAEYIGCANGLTTVNACGAGQY YDQALGRCALAGTGYCRQYVFTPPPAPVVYPD GFDTYCSANNLATGIHPDPYSCFSYVECTFGR TTHMPCPAGLSFDRSLLVCDGNRYQNCGGNVL VGK

2.3. Data collection and processing

For data collection, the crystals obtained from initial crystallization screening were soaked in cryoprotectant solution consisting of reservoir solution supplemented with 20%(*v*/*v*) glycerol and placed under a cold nitrogen-gas stream at 100 K. Native diffraction data were then collected at a wavelength of 1.0000 Å on beamline BL-5A at Photon Factory (PF), Tsukuba, Japan using a crystal from initial screening obtained under condition No. 48 from the JCSG Core II Suite. To determine the structure of EHFP, native-SAD data sets were collected using a long wavelength of 2.1000 Å on beamline BL-17A at PF using a crystal obtained under condition No. 25 of Crystal Screen by rescreening, since crystals could not be obtained again from condition No. 48 of the JCSG Core II Suite. To reduce X-ray absorption by the solution at the long wavelength, the solutionless crystal-mounting method was applied for native-SAD data collection (Kitago *et al.*, 2005) using solutionless crystal-mounting tools (solutionless loop and AFERO extraction–freezing robot; SYSCON, Japan) developed in our laboratory.

To increase the anomalous signal with high redundancy, eight native-SAD data sets [720° of rotation per data set, a total of 11 520 (8 × 720/0.5) frames] were collected by the exposure of eight positions of a single crystal. All of the data sets were then indexed, integrated, scaled and merged using *XDS*/*XSCALE* (Kabsch, 2010). The data-collection and processing statistics are summarized in Table 3.

3. Results and discussion

In this study, native EHEP was successfully obtained from *A. kurodai* with high purity (Fig. 1*a*). Sequence analysis

Table 2
Crystallization.

	Crystal for native data set	Crystal for SAD data set
Method	Sitting-drop vapor diffusion	Sitting-drop vapor diffusion
Plate type	96-well	96-well
Temperature (K)	293	293
Protein concentration (mg ml^{-1})	15–25	15–25
Buffer composition of protein solution	20 m*M* sodium acetate buffer pH 6.0, 100 m*M* NaCl	20 m*M* sodium acetate buffer pH 6.0, 100 m*M* NaCl
Composition of reservoir solution	0.2 *M* lithium sulfate, 20%(*w*/*v*) PEG 3350	1.0 *M* sodium acetate, 0.1 *M* imidazole pH 6.5
Volume and ratio of drop	2 µl, 1:1	2 µl, 1:1
Volume of reservoir (µl)	75	75

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

	Native	Native-SAD
Diffraction source	BL-5A, PF	BL-17A, PF
Wavelength (Å)	1.0000	2.1000
Temperature (K)	100	100
Detector	ADSC Quantum 315	PILATUS3 6M
Crystal-to-detector distance (mm)	90.93	193.15
Rotation range per image (°)	0.5	0.5
Total rotation range (°)	180	720 × 8
Exposure time per image (s)	0.5	0.5
Space group	$P2_12_12_1$	$P2_12_12_1$
a, *b*, *c* (Å)	42.1, 65.3, 66.4	42.2, 65.3, 66.5
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution range (Å)	50–1.20 (1.27–1.20)	46.6–2.48 (2.54–2.48)
Total No. of reflections	409883 (64023)	1295445 (61166)
No. of unique reflections	56777 (8863)	12330 (822)
Completeness (%)	97.7 (95.4)	97.7 (85.5)
Multiplicity	7.21 (7.22)	105 (74.4)
$\langle I/\sigma(I)\rangle$	17.71 (3.94)	91.88 (37.69)
$R_{\rm meas}$† (%)	7.6 (52.2)	6.9 (14.2)

† $R_{\rm meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity for a reflection and $\langle I(hkl)\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections.

(http://www.cbs.dtu.dk/services/SignalP/) predicted that the 17 N-terminal amino acids represent a signal peptide. To confirm this, purified native EHEP was analyzed by MALDI-TOF mass spectrometry. The result showed that mature native EHEP has a molecular weight of 22.582 kDa (Fig. 1*b*). This value is consistent with the estimated molecular weight of EHEP without the 19 or 20 N-terminal amino acids (22.667 and 22.568 kDa, respectively), indicating that these residues are cleaved off during the maturation process. To clarify the cleavage site, N-terminal sequencing was performed by Edman degradation. However, the analysis failed because the N-terminus seemed to be modified. Considering the effect of chemical modification, the truncation of 20 residues is most likely, although the modification has not yet been verified.

In initial crystallization screening, we obtained crystals under five different conditions; however, only two conditions (No. 48 of the JCSG Core II Suite and No. 25 of Crystal Screen) produced high-quality crystals for the diffraction experiment. Moreover, it proved to be difficult to reproduce these crystals even when the same crystallization kits were used or the conditions were expanded around the growth conditions outlined in Section 2.2 by adjusting the pH and the concentrations of buffer and precipitants (salt and organic solvent). After obtaining a native data set at a resolution of 1.2 Å using a crystal obtained during initial screening (Figs. 2*a* and 3*a*), we attempted to solve the structure of EHEP using the molecular-replacement method. Since no homologous protein or domain structure was available (identities of <28% for a fragment of about half of full-length EHEP), we undertook modeling using *Phyre*2 (http://www.sbg.bio.ic.ac.uk/servers/phyre2) to provide a search model. However, this failed to provide any answers. EHEP is a cysteine-rich protein, and the diffraction ratio ($\langle\Delta F\rangle/\langle F\rangle$) of the anomalous scat-

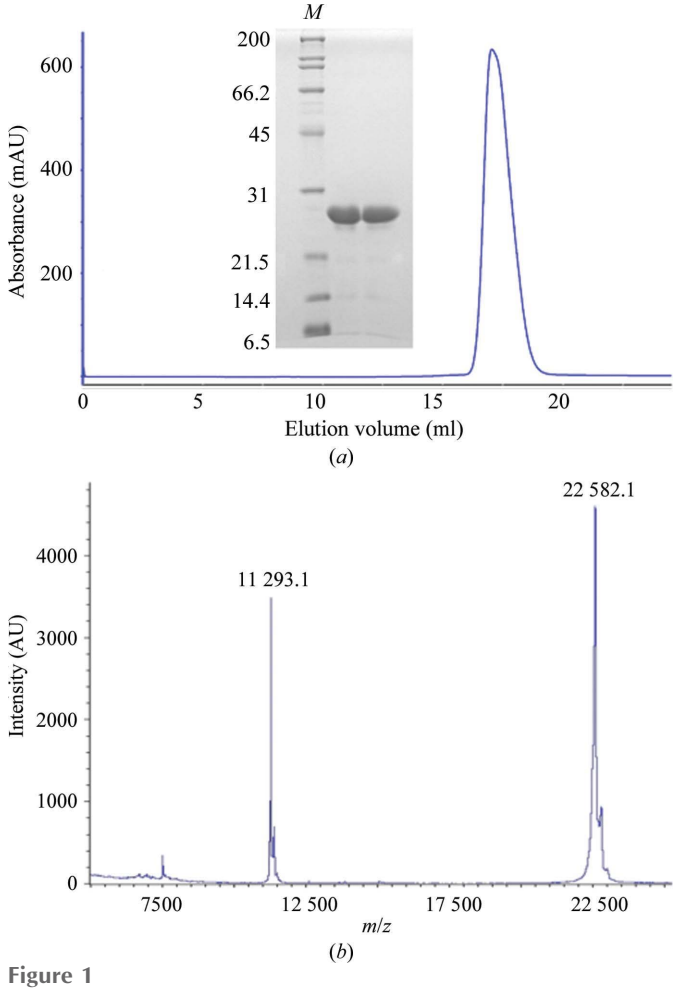


Figure 1
Purification and characterization of EHEP. (*a*) Size-exclusion chromatography of SEC. Purified EHEP was checked by SDS–PAGE. Lane *M* contains molecular-weight markers (labeled in kDa). (*b*) MALDI-TOF mass spectrometry of EHEP.

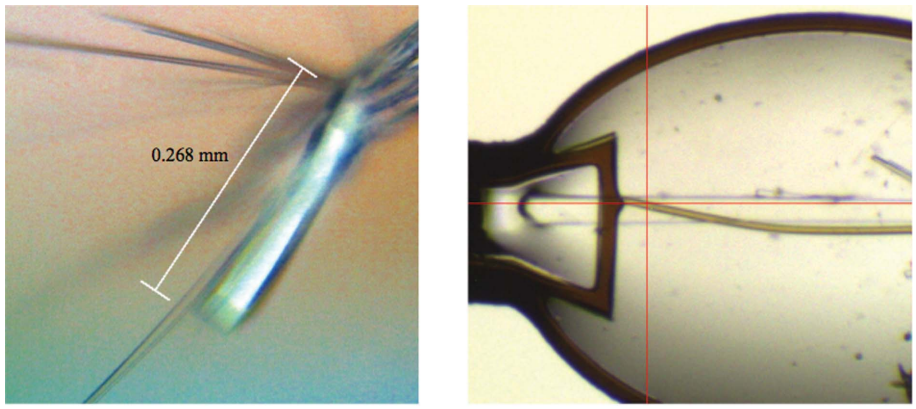


Figure 2
Photographs of EHEP crystals. (*a*) EHEP crystal for the collection of native data. (*b*) EHEP crystal for the collection of native-SAD data cooled using a solutionless loop.

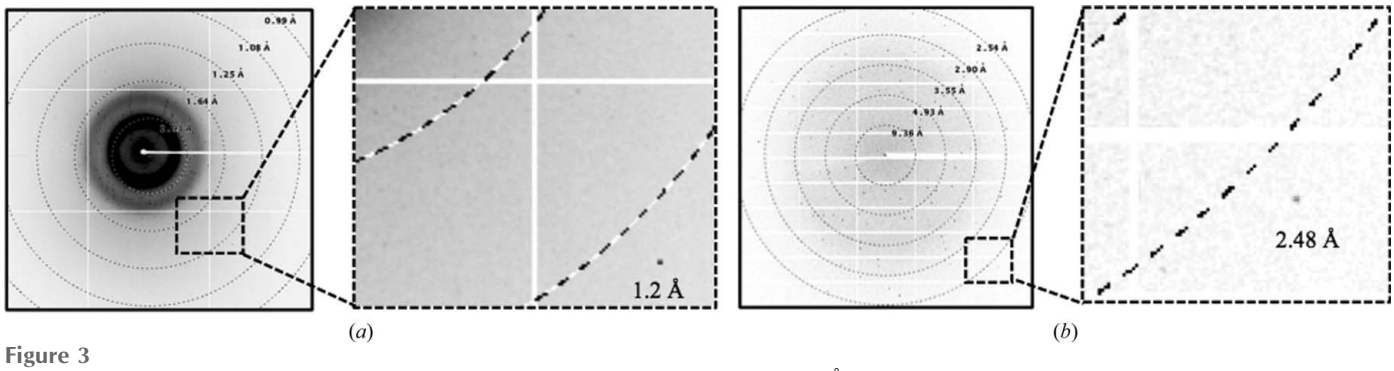


Figure 3
X-ray diffraction patterns of EHEP crystals. (*a*) Native data (using a wavelength of 1.0000 Å and a crystal-to-detector distance of 90.93 mm). (*b*) Native-SAD data (using a wavelength of 2.1000 Å and a crystal-to-detector distance of 193.15 mm).

tering of S atoms at a wavelength of 2.1 Å is calculated to be 2.28% (0.57% at 1.0 Å wavelength) (Hendrickson & Teeter, 1981), which is considered sufficient for native-SAD phasing. Therefore, we tried to use the native-SAD method for phasing. Native-SAD data were successfully collected using a wavelength of 2.1000 Å (Figs. 2*b* and 3*b*).

The preliminary X-ray diffraction pattern showed that the crystals obtained from two conditions (from different kits) were isomorphous and belonged to space group $P2_12_12_1$. Assuming the presence of one molecule per asymmetric unit, the Matthews coefficient (Matthews, 1968) was estimated to be 1.83 Å^3 Da^{-1}, corresponding to a solvent content of 32.7%. Using native-SAD data, all 20 S-atom sites were located using *SHELXC*/*D* (Sheldrick, 2010) and model building was performed using *phenix.autosol* in the *Phenix* software suite (Liebschner *et al.*, 2019) with rephasing and phase improvement. This resulted in 58% of the structure of EHEP being built automatically (Fig. 4). Full model building and refinement are currently in progress.

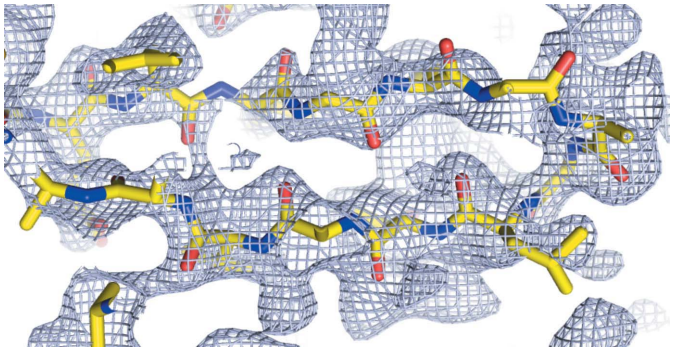


Figure 4
An initial electron-density map showing fragments of the EHEP molecule. The fragments (stick representation) were built by *phenix.autosol* based on an electron-density map contoured at the 1.5σ level (gray).

Acknowledgements

We are grateful to the PF synchrotron for beam time and the beamline staff for their assistance.

Funding information

This work was supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research; BINDS) from the Japan Agency for Medical Research and Development (AMED) under grant No. JP18am0101071.

References

Demirbas, M. F. (2011). *Appl. Energy*, **88**, 3473–3480.

Ellabban, O., Abu-Rub, H. & Blaabjerg, F. (2014). *Renew. Sustain. Energy Rev.* **39**, 748–764.
Enquist-Newman, M., Faust, A. M., Bravo, D. D., Santos, C. N., Raisner, R. M., Hanel, A., Sarvabhowman, P., Le, C., Regitsky, D. D., Cooper, S. R., Peereboom, L., Clark, A., Martinez, Y., Goldsmith, J., Cho, M. Y., Donohoue, P. D., Luo, L., Lamberson, B., Tamrakar, P., Kim, E. J., Villari, J. L., Gill, A., Tripathi, S. A., Karamchedu, P., Paredes, C. J., Rajgarhia, V., Kotlar, H. K., Bailey, R. B., Miller, D. J., Ohler, N. L., Swimmer, C. & Yoshikuni, Y. (2014). *Nature*, **505**, 239–243.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature*, **290**, 107–113.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kellogg, J., Grace, M. H. & Lila, M. A. (2014). *Mar. Drugs*, **12**, 5277–5294.
Kitago, Y., Watanabe, N. & Tanaka, I. (2005). *Acta Cryst.* D**61**, 1013–1021.
La Barre, S., Potin, P., Leblanc, C. & Delage, L. (2010). *Mar. Drugs*, **8**, 988–1010.
Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Shibata, T., Fujimoto, K., Nagayama, K., Yamaguchi, K. & Nakamura, T. (2002). *Int. J. Food Sci. Technol.* **37**, 703–709.
Tsuji, A., Kuwamura, S., Shirai, A. & Yuasa, K. (2017). *PLoS One*, **12**, e0170669.
Tsuji, A., Tominaga, K., Nishiyama, N. & Yuasa, K. (2013). *PLoS One*, **8**, e65418.
Wei, N., Quarterman, J. & Jin, Y.-S. (2013). *Trends Biotechnol.* **31**, 70–77.