



Structure of the plant growth-promoting factor YxaL from the rhizobacterium *Bacillus velezensis* and its application to protein engineering

Jiheon Kim,^{a‡} Ha Pham,^{b‡} Yeongjin Baek,^a Inseong Jo,^{a§} Yong-Hak Kim^{b*} and Nam-Chul Ha^{a*}

Received 22 July 2021

Accepted 4 November 2021

Edited by M. Czjzek, Station Biologique de Roscoff, France

‡ These authors contributed equally to this work.

§ Current address: KoBioLabs Inc., Seoul 08826, Republic of Korea

Keywords: plant growth-promoting factors; YxaL; β -proteins; structure-based protein engineering; *Bacillus velezensis*.

PDB references: YxaL, wild type (SeMet), 7dxn; T175W/W215G mutant, 7eq5; T175W/S213G/W215A mutant, 7evf

Supporting information: this article has supporting information at journals.iucr.org/d

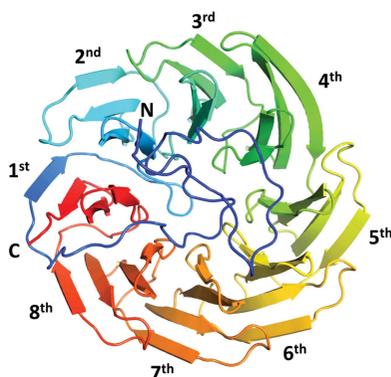
^aDepartment of Agricultural Biotechnology, Center for Food and Bioconvergence, and Research Institute of Agriculture and Life Sciences, CALS, Seoul National University, Seoul 08826, Republic of Korea, and ^bDepartment of Microbiology, Daegu Catholic University School of Medicine, Daegu 42472, Republic of Korea. *Correspondence e-mail: ykim@cu.ac.kr, hanc210@snu.ac.kr

The YxaL protein was isolated from the soil bacterium *Bacillus velezensis* and has been shown to promote the root growth of symbiotic plants. YxaL has further been suggested to act as an exogenous signaling protein to induce the growth and branching of plant roots. Amino acid sequence analysis predicted YxaL to exhibit an eight-bladed β -propeller fold stabilized by six tryptophan-docking motifs and two modified motifs. Protein engineering to improve its structural stability is needed to increase the utility of YxaL as a plant growth-promoting factor. Here, the crystal structure of YxaL from *B. velezensis* was determined at 1.8 Å resolution to explore its structural features for structure-based protein engineering. The structure showed the typical eight-bladed β -propeller fold with structural variations in the third and fourth blades, which may decrease the stability of the β -propeller fold. Engineered proteins targeting the modified motifs were subsequently created. Crystal structures of the engineered YxaL proteins showed that the typical tryptophan-docking interaction was restored in the third and fourth blades, with increased structural stability, resulting in improved root growth-promoting activity in *Arabidopsis* seeds. The work is an example of structure-based protein engineering to improve the structural stability of β -propeller fold proteins.

1. Introduction

Many soil bacteria interact with plants through commensal, symbiotic and pathogenic processes with distinct degrees of proximity to the roots and the surrounding soil (Dastogeer *et al.*, 2020). For example, plant growth-promoting rhizobacteria can enhance plant growth and protect plants from diseases through a variety of mechanisms (Kilian *et al.*, 2000; Arkhipova *et al.*, 2005; Barnawal *et al.*, 2013; Kang *et al.*, 2014; Xie *et al.*, 2014; Kuan *et al.*, 2016; Radhakrishnan & Lee, 2016; Radhakrishnan *et al.*, 2017; Rui, 2020). YxaL proteins are widely produced by plant growth-promoting *Bacillus* rhizobacteria. The N-terminal signal peptide (amino-acid residues 1–44) of YxaL suggests that it is a secreted protein in the soil that affects target plants (Hirose *et al.*, 2000).

The YxaL-producing rhizobacterium *B. velezensis* is a plant growth-promoting, aerobic, Gram-positive, endospore-forming bacterium. Plant-growth effects of *B. velezensis* YxaL have been reported in the eudicot *Arabidopsis thaliana*, the monocot *Oryza sativa* L. (rice) and *Capsicum annuum* (pepper) (Kim *et al.*, 2019). YxaL treatment of *A. thaliana* or *O. sativa* seeds effectively increased the lengths of the main roots and the number of lateral roots compared with untreated



seeds. The root morphologies of treated *A. thaliana* and *O. sativa* L. suggested that YxaL has a common effect on root development and growth across various plant types, such as eudicot and monocot plants. YxaL-treated *C. annuum* seedlings exhibited an improved tolerance to drought, showing recovery of leaf growth after rewatering. The results indicated the potential use of YxaL as a seed-treatment additive for promoting root growth and increasing tolerance to drought stress (Kim *et al.*, 2019). Further study led to the proposal that the *B. velezensis* YxaL protein provides a platform for promoting root development and plant growth, acting as an extracellular signaling protein to the root cells of diverse plants (Kim *et al.*, 2019). However, the binding receptor for YxaL in plant root cells needs to be elucidated. Since the YxaL protein functions in the extracellular environment, increasing its structural stability would increase the utility of this protein in promoting the growth of plants.

Proteins with the β -propeller fold are ubiquitous in nature and are widely used as structural scaffolds for ligand binding and enzymatic activity. The β -propeller fold consists of all- β protein architectures characterized by 4–8 β -sheets arranged toroidally around a central axis like a propeller (Neer & Smith, 2000). The β -propeller fold is mostly stabilized by interactions of the aromatic rings of hydrophobic residues between each blade (Noirot-Gros *et al.*, 2002; Kopec & Lupas, 2013). Based on stabilizing interactions, the β -propeller proteins can be categorized into several groups: WD40 (Smith, 2008), Kelch (Hudson & Cooley, 2008), YWTD (Chen *et al.*, 2011), YVTN (Chen *et al.*, 2011), NHL (Chaudhuri *et al.*, 2008), PQQ (Ghosh *et al.*, 1995) and clathrin (ter Haar *et al.*, 1998). Amino-acid sequence analysis of *B. velezensis* YxaL suggested that YxaL has an eight-bladed β -propeller fold belonging to the PQQ family of proteins that contain tryptophan-docking motifs similar to that found in the lipoprotein BamB from *Escherichia coli* (Duine & Jongejan, 1989; Ghosh *et al.*, 1995; Xia *et al.*, 1996; Kim *et al.*, 2019; Neer & Smith, 2000).

Amino-acid sequence analysis predicted that YxaL exhibits eight blades in the β -propeller fold, which are stabilized by six tryptophan-docking motifs and two modified motifs (Kim *et al.*, 2019). The template-based protein structure-prediction program also created several different structural models and did not give a consistent model around the modified motifs. This study reports the crystal structure of YxaL from *B. velezensis* and structure-based protein engineering to improve its plant growth-promoting activity by increasing its structural stability.

2. Materials and methods

2.1. Construction of plasmids and site-directed mutagenesis

The plasmid pProEX N-His-TEV-YxaL was constructed for the bacterial expression of wild-type YxaL using a PCR fragment encoding mature YxaL (residues 45–415) from *B. velezensis* strain GH1-13 (Kim *et al.*, 2019). Site-directed mutagenesis was carried out with the site-specific primer sets T175W_F, T175W_R, W215G_F, S213G/W215A_F and

W215G_R to construct the double and triple mutants T175W/W215G and T175W/S213G/W215A. All site-directed mutations were confirmed by DNA sequencing. The nucleotide sequences of the primers are listed in Supplementary Table S1.

2.2. Protein overexpression and purification of YxaL proteins

The wild-type and mutant YxaL-expressing plasmids were transformed into the *E. coli* BL21(DE3) strain for the unlabeled protein or the *E. coli* B834(DE3) strain for the selenomethionine (SeMet)-labeled protein. The unlabeled wild-type and mutant YxaL proteins were purified using Ni-NTA affinity chromatography followed by cleavage by treatment with TEV protease, as described in a previous study (Kim *et al.*, 2019). The hexahistidine-tagged cleaved proteins were subjected to size-exclusion chromatography using a HiLoad 26/600 Superdex 200 pg column (GE Healthcare) in 20 mM Tris-HCl pH 8.0 buffer containing 150 mM NaCl. The purified proteins were concentrated using Vivaspinn 20 (Sartorius, Germany) and stored at -80°C . Protein concentrations were determined based on the calculated extinction coefficients at 280 nm.

For the production of SeMet-labeled protein, *E. coli* B834(DE3) cells were cultured in M9 medium supplemented with L-SeMet. The proteins were purified using Ni-NTA agarose resin (Qiagen, Hilden, Germany), which was treated with TEV protease at a ratio of 1:100 to cleave the hexahistidine tag, as described elsewhere (Raran-Kurussi *et al.*, 2017). The TEV protease-treated proteins were then eluted from the Ni-NTA agarose column and loaded onto a HiTrap Q column (GE Healthcare, USA). A linear gradient of increasing NaCl concentration was applied to the HiTrap Q column. The fractions containing the YxaL proteins were pooled, concentrated and subjected to size-exclusion chromatography using a HiLoad 26/600 Superdex 200 pg column (GE Healthcare). The size and purity of the final protein sample were examined by SDS-PAGE.

2.3. Crystallization of the YxaL proteins

The SeMet-labeled wild-type YxaL protein was crystallized at 16°C by the sitting-drop vapor-diffusion method using a precipitant solution consisting of 0.1 M CHES-NaOH pH 7.0, 1 M sodium citrate. The crystals were flash-cooled in liquid nitrogen at -196°C . All mutant proteins were crystallized at 16°C using the hanging-drop vapor-diffusion method. The T175W/W215G mutant protein was crystallized using a precipitant solution consisting of 0.1 M HEPES-NaOH pH 7.0, 1.5 M lithium sulfate monohydrate. The T175W/S213G/W215A mutant protein was crystallized using a precipitant solution consisting of 0.2 M sodium citrate pH 5.8, 15% (w/v) PEG 4000, 20% (v/v) 2-propanol. Crystals of both the T175W/W215G and the T175W/S213G/W215A mutant proteins were flash-cooled in liquid nitrogen after dipping them into a cryoprotectant solution consisting of 12.5% (v/v) diethylene glycol, 12.5% (v/v) ethylene glycol, 12.5% (v/v) MPD, 12.5% (v/v) 1,2-propanediol, 12.5% (v/v) dimethyl sulfoxide, 12.5 mM NDSB 201.

2.4. Data collection and structural determination of the YxaL proteins

X-ray diffraction data sets were collected on the 5C and 11C beamlines at the Pohang Accelerator Laboratory and were processed with the *HKL-2000* package (Otwinowski & Minor, 1997). The SeMet-labeled wild-type YxaL protein crystals belonged to space group $P3_221$. The structure was determined by the single anomalous dispersion (SAD) method using the data set from the SeMet-labeled crystal in *Phenix* (Liebschner *et al.*, 2019). The model was built using *Coot* (Emsley *et al.*, 2010). The final structure of YxaL was refined to 1.8 Å resolution with an *R* factor of 16% and an R_{free} of 19% using *Phenix*. The structures of the mutants were determined by the molecular-replacement (MR) method using the coordinates of the wild-type protein as the search model. Table 1 gives further details of structure determination and refinement.

2.5. Treatment of *A. thaliana* seeds with YxaL and cultivation of *A. thaliana*

A. thaliana seeds were disinfected with 2% hypochlorite and 0.05% Triton-X for 10 min and washed several times with sterile water at 4°C. The disinfected seeds were then treated with 1 µg ml⁻¹ YxaL or its mutant proteins in phosphate-buffered saline (PBS; 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄ and 0.24 g KH₂PO₄ in 1 l deionized water) for 20 h at 4°C on a rotary shaker (20 rev min⁻¹). The soaked seeds were cultivated on 0.5% NuSieve GTG agarose plates (FMC Bio-Products, Rockland, USA) containing 1% sucrose in 0.5× Murashige and Skoog basal medium (Sigma, catalog No. M5519) in the dark at 25°C. The main root length and lateral root number of two-week-old seedlings after staining with 0.1% methylene blue were measured under a stereomicroscope.

2.6. Limited proteolysis using trypsin

The wild-type and mutant YxaL proteins, at a concentration of 9.5 mg ml⁻¹, were incubated with and without 0.5 mg ml⁻¹ trypsin (12:1 molar ratio of protein monomer to trypsin) in 20 mM Tris–HCl buffer pH 8.0 containing 150 mM NaCl and 2 mM β-mercaptoethanol at 60°C for 45 min. An aliquot was diluted and monitored by SDS–PAGE.

2.7. T_m measurement by thermal shift assay

Thermal shift assays were performed in a QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific, USA) using a Protein Thermal Shift Dye Kit (Thermo Fisher Scientific, USA). Melt curves were constructed with 50 µM each of wild-type, T175W/W215G and T175W/S213G/W215A mutant YxaL proteins in 20 mM Tris pH 8.0 buffer containing 150 mM NaCl from 13 to 99°C in intervals of 0.1°C.

2.8. Statistics

Experimental data from at least three independent replicates were reported as the mean and standard deviation of the mean. Statistics were analyzed by performing the Mann–

Table 1
X-ray diffraction data-collection and refinement statistics.

	Wild-type YxaL (SeMet)	T175W/W215G mutant YxaL	T175W/S213G/ W215A mutant YxaL
X-ray diffraction data			
Beamline	5C, PAL	11C, PAL	5C, PAL
Wavelength (Å)	0.97934	0.97942	1.00003
Space group	$P3_221$	$P6_5$	$R3$
<i>a</i> , <i>b</i> , <i>c</i> (Å)	81.6, 81.6, 195.9	176.446, 176.446, 63.596	152.987, 152.987, 61.694
α , β , γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	50.0–1.80 (1.83–1.80)	50.0–2.60 (2.64–2.60)	50.0–1.50 (1.53–1.50)
$R_{\text{p.i.m.}}$	0.018 (0.098)	0.029 (0.074)	0.020 (0.116)
R_{merge} ($I/\sigma(I)$)	0.070 (0.336)	0.125 (0.296)	0.060 (0.320)
Completeness (%)	99.9 (100.0)	100.0 (100.0)	98.1 (95.2)
Multiplicity	14.3 (10.9)	17.2 (14.8)	8.3 (6.5)
Refinement			
Resolution (Å)	40.76–1.80	48.89–2.60	38.25–1.50
No. of reflections	70892	35089	83953
$R_{\text{work}}/R_{\text{free}}$	0.1531/0.1858	0.1796/0.2133	0.1482/0.1648
Total No. of atoms	6148	5601	3124
Wilson <i>B</i> factor (Å ²)	18.18	37.50	12.67
R.m.s. deviations			
Bond lengths (Å)	0.007	0.008	0.014
Bond angles (°)	0.856	0.76	1.30
Ramachandran plot			
Favored (%)	97.28	97.27	96.50
Allowed (%)	2.72	2.73	3.50
Outliers (%)	0.00	0.00	0.00
PDB code	7dxn	7eq5	7evf

Whitney *U* test and the *t*-test, with *p* values of less than 0.05 considered to be significantly different.

3. Results

3.1. Crystal structure of wild-type YxaL protein with an eight-bladed β-propeller fold

To obtain a mature form of the wild-type YxaL protein (residues 45–415), we overexpressed the YxaL protein without the signal sequence (residues 1–44) in the *E. coli* cytosol. A large amount of the protein was successfully purified to homogeneity. Size-exclusion chromatography combined with multi-angle light scattering (SEC–MALS) suggested that the mature YxaL protein is monomeric in solution (Supplementary Fig. S1a). The crystal structure of YxaL was determined at 1.8 Å resolution by single anomalous dispersion (SAD) using SeMet-labeled crystals (Table 1). The asymmetric unit contained two molecules with brief molecular contact in the crystal, suggesting that YxaL is a monomeric protein, consistent with the SEC–MALS result.

The overall structure was an approximate torus shape with a width of ~50 Å and a depth of ~30 Å. The YxaL structure exhibited an eight-bladed β-propeller fold domain (residues 87–415), apart from the N-terminal loop region (residues 45–86). Each blade in the β-propeller fold domain was formed by a twisted β-sheet consisting of four antiparallel β-strands radiating from the central axis. The structure of YxaL exhibited a Velcro-like closure, as found in typical β-propeller fold proteins. The outermost β-strand of the first blade was the N-terminal β-strand, which stabilized the circular toroidal

arrangement between the first and second blades in the Velcro-like closure (Fig. 1*a* and Supplementary Fig. S2).

The top region is the N-terminal region and the bottom region is the C-terminal tail region in the standard orientation. The loop connecting the first and second blades was inserted into the bottom region of the torus, resulting in a flat surface on the bottom of the torus (Fig. 1*a*). The central axis of the top region is covered by the N-terminal loop region like a plug (Fig. 1*b*). A mutant protein in which the N-terminal loop region was deleted was not correctly expressed in the *E. coli* system, indicating that the N-terminal region contributes to the structural stability of the β -propeller fold of the YxaL protein.

3.2. Conserved GXXXW motifs in the tryptophan-docking interaction

Each blade consists of four antiparallel β -strands, named A–D, where β -strand A is closest to the central axis and β -strand D is on the external surface. Unlike the other blades, which consist of continuous residues, the first blade is a combination of the N- and C-terminal regions. β -Strands A–C are from the C-terminal region, whereas β -strand D is from

the N-terminal region (Supplementary Fig. S2). The GXXXW motif (the invariant Gly and Trp residues are in the G and W sites with any amino acid in the X sites) is conserved in the outer β -strand D of each blade in the PQQ family of β -propeller fold proteins. The conserved GXXXW motifs play a crucial role in stabilizing the β -propeller structure by forming a tryptophan-docking interaction between adjacent blades (Anthony & Ghosh, 1998; Jansen *et al.*, 2012). Unlike typical PQQ family β -propeller fold proteins, YxaL has variations in the GXXXW motifs in the third and fourth blades (Supplementary Fig. S3).

The tryptophan-docking interaction between the adjacent GXXXW motifs is made between the tryptophan residue in the GXXXW motif and the glycine residue in the GXXXW motif from the next blade in a circularly permuted manner (Fig. 2*a*). In the tryptophan-docking interaction, the indole ring of tryptophan is fitted into a hydrophobic pocket in the latter blade. The aromatic indole ring of the tryptophan residue makes a further π interaction with the planar peptide bond of the glycine residue of the GXXXW motif in the next blade (Fig. 2*b*). The tryptophan-docking interaction mediated by the GXXXW motifs thus appears to be important for stabilization of the blades in the β -propeller fold.

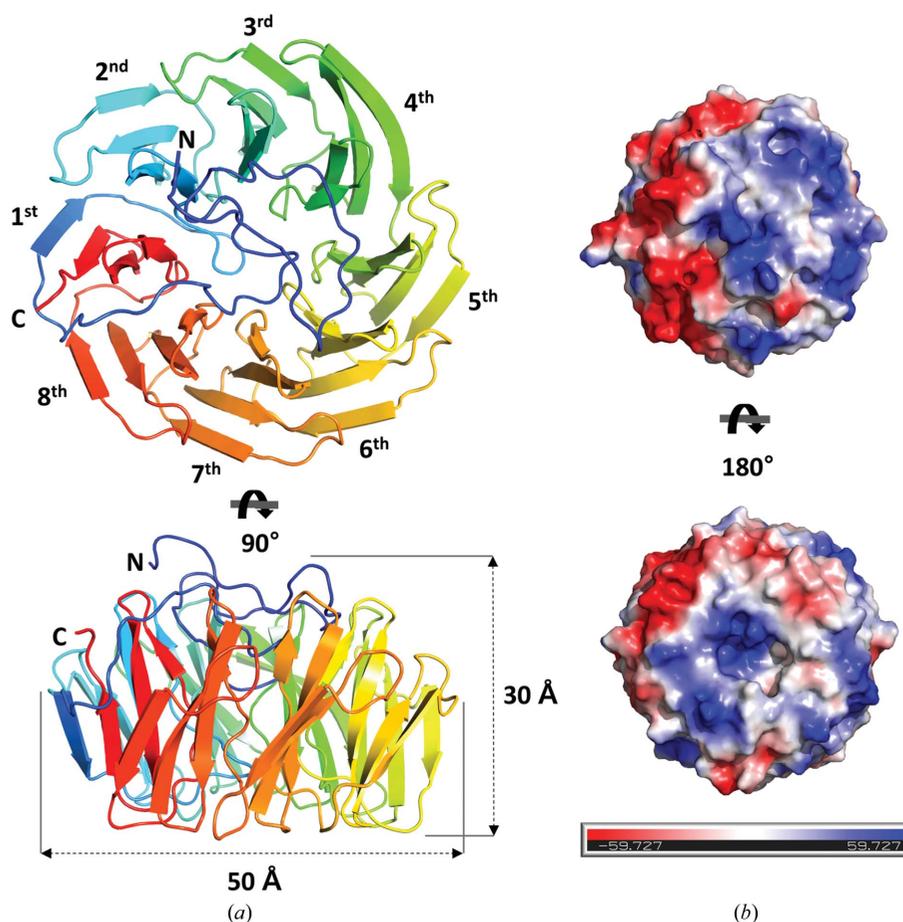


Figure 1

Crystal structure of the wild-type YxaL protein. (a) The structure of wild-type YxaL is displayed in a ribbon representation in orthogonal views (the top view is at the top and the side view is at the bottom). The β -propeller fold of YxaL is color-ramped from blue (N-terminus) to red (C-terminus). The N-terminus (N), C-terminus (C) and order of the blades are annotated. The width and depth are shown by dashed arrows at the bottom. (b) Surface representations of YxaL colored by electrostatic surface potential. The top view is at the top and the bottom view is at the bottom.

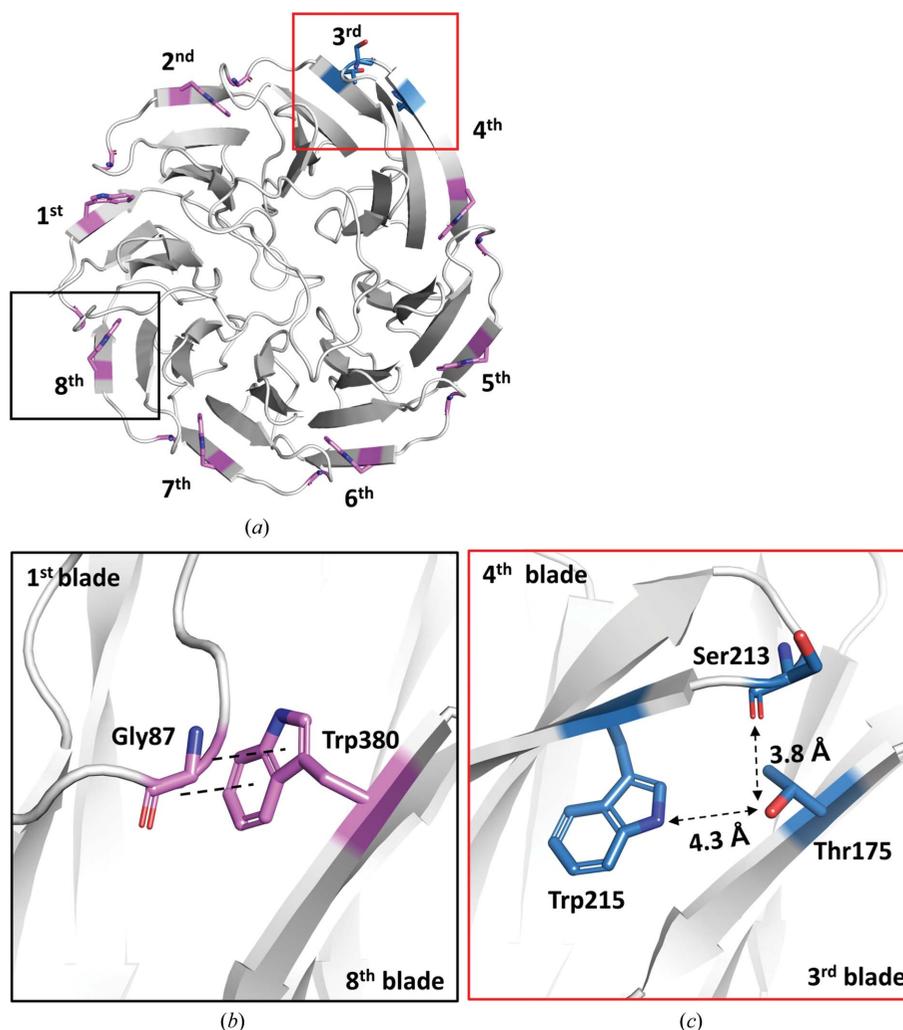


Figure 2
 Structural comparison of the conserved GXXXW motifs and their modified motifs in the third and fourth blades of the wild-type YxaL protein. (a) All tryptophan and glycine residues in the conserved GXXXW motifs are emphasized as violet sticks. Thr175, Ser213 and Trp215 in the modified motifs are shown as sky-blue stick models. (b) Close-up view of the conserved GXXXW motifs framed in the black box in (a). The π interaction between Trp380 and Gly87 is shown as dotted lines. (c) Close-up view of the modified motifs in the third and fourth blades framed in the red box in (a). The distances between Thr175 and Ser213 and between Thr175 and Trp215 are shown as dashed arrows.

3.3. The modified motifs in the third and fourth blades

Structural prediction using the amino-acid sequence of YxaL failed to identify GXXXW motifs in the third and fourth blades. A GXXXW sequence was not found explicitly in the sequence between the second and fifth explicit GXXXW motifs. The crystal structure of YxaL revealed that Gly171 in the third blade provides the tryptophan-docking interaction with Trp133 in the GXXXW motif in the second blade, resulting in a GXXXT sequence as a variant of the GXXXW motif in the third blade. The structure further revealed that Trp219 in the fourth blade is involved in a tryptophan-docking interaction with Gly256 of the GXXXW motif in the fifth blade, resulting in a variant WXXXW sequence in the fourth blade (Fig. 2c, Supplementary Figs. S2 and S3).

Thr175 in the GXXXT sequence of the third blade was not involved in any tryptophan-docking interaction (Fig. 2c). The indole ring of Trp215 in the WXXXW sequence of the fourth

blade seemed to be fixed by interaction with the hydrophobic residues Phe165, Val177, Pro178 and Ile207, similar to the conserved GXXXW motif. However, the π interaction mediated by Trp215 was not observed because of the absence of the π electron partner. The structural stability of YxaL appeared to decrease due to the absence of this interaction. Thr175 and Trp215 in the variant GXXXW sequences occupied a lower part of the β -propeller than the canonical tryptophan and glycine residues (Figs. 2b and 2c). The modified motifs also affected the upper region, extending β -strands C4 and D4 noticeably to the top of the β -propeller (Figs. 2a and 2c).

3.4. Structure-based protein design to facilitate the tryptophan-docking interaction

We noted that the presence of Thr175 in the GXXXT sequence of the third blade improves the structural stability of YxaL. Suppose that the typical tryptophan-docking interac-

tion is restored in the third and fourth blades. In this case, the Trp residue would occupy the Thr175 site and would make a π interaction with the Gly residue. Thus, we decided to change Thr175 to tryptophan to restore the typical tryptophan-docking interaction. The similar distances of Thr175 to Ser213 and Trp215 represented two cases for selecting a residue that would be changed to glycine.

We first created a T175W/W215G double mutant to restore the conserved GXXXW sequences in the modified GXXXW motifs. In the conserved GXXXW motif, the Thr175 site in the GXXXW sequence and the Trp215 site in the WXXXW sequence become a tryptophan and a glycine, respectively. We

expected a tryptophan-docking interaction to be made between the mutated Trp175 of the GXXXW sequence in the third blade and the mutated Gly215 of the GXXXW sequence in the fourth blade. The resulting sequences in the third and fourth blades of the T175W/W215G mutant mimicked the conserved GXXXW motif.

We next designed a T175W/S213G/W215A triple mutant based on the structural features of the third and fourth blades. The crystal structure of YxaL showed that Thr175 in the third blade and Ser213 in the fourth blade were within a proximity of 4 Å (Fig. 2c). The changes of Thr175 to tryptophan and Ser213 to glycine would mimic the putative tryptophan-

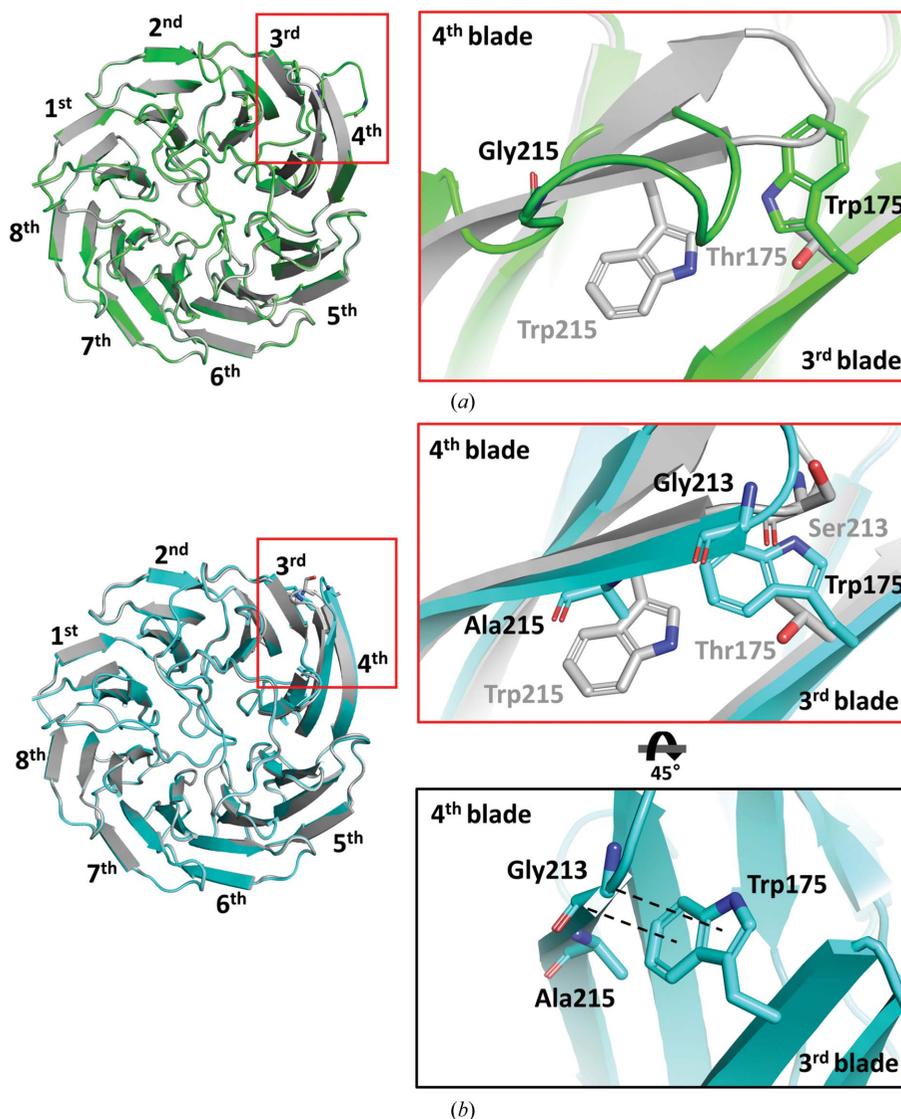


Figure 3

Crystal structures of the mutant YxaL proteins. (a) Superimposed structures of T175W/W215G mutant YxaL (green) and wild-type YxaL (gray) in ribbon representation. The orders of the blades are annotated in the left panel. The boxed region in the left panel is enlarged in the right panel. The mutated Trp175 in the restored $^{171}\text{GXXXW}^{175}$ motif in the third blade and the mutated Gly215 in the restored $^{215}\text{GXXXW}^{219}$ motif in the fourth blade are shown in stick representation. (b) Superimposed structures of T175W/S213G/W215A mutant YxaL (cyan) and wild-type YxaL (gray) in ribbon representation. The boxed region in the left panel is enlarged in the right panels with 45° rotated views. In the bottom right panel, only the mutant structure is shown for clarity. The mutated Trp175 in the restored $^{171}\text{GXXXW}^{175}$ motif in the third blade is shown in stick representation. The mutated Gly213 and Ala215 in the modified $^{213}\text{GXAXXXW}^{219}$ motif in the fourth blade are shown as stick models. The π interaction between the mutated Trp175 and the mutated Gly213 is shown with dotted lines.

docking interaction. Trp215 would be too close to the mutated Trp175 in the modeled structure of T175W/S213G. Thus, we further changed Trp215 to alanine to prevent a possible clash with the mutated Trp175. The T175W/S213G/W215A triple mutant was created to mimic the typical tryptophan-docking interaction between Trp175 and Gly213.

Size-exclusion chromatography combined with multi-angle light scattering (SEC-MALS) suggested that the two YxaL mutants were monomeric in solution, similar to the wild-type protein (Supplementary Figs. S1b and S1c). Thus, the mutant proteins were expected to be correctly folded, exhibiting a similar overall structure to wild-type YxaL. We determined the two crystal structures at a high resolution to examine how

the mutations affected the 3D structures at the molecular level in subsequent studies (Table 1).

3.5. Structural analysis of the T175W/W215G mutant YxaL protein

In the crystal structure of the T175W/W215G mutant, two molecules were present in the asymmetric unit, showing different conformations in a loop connecting β -strand C4 to β -strand D4 (residues 211–214). In one molecule of the asymmetric unit the connecting loop was not well defined (Supplementary Fig. S4a). In the other molecule, Pro212 of the loop in the fourth blade made a π interaction with the mutated Trp175 in the third blade (Supplementary Fig. S4b). The

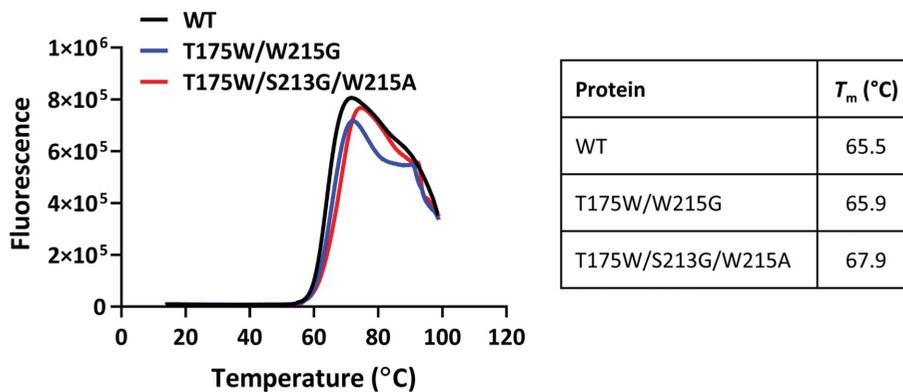


Figure 4 Comparison of the structural stability of wild-type and engineered YxaL proteins. The T_m values of the wild-type (WT; black line), T175W/W215G mutant (blue line) and T175W/S213G/W215A mutant (red line) YxaL proteins were measured by thermal shift assay. The line graphs are shown in the left panel and the mean T_m values of the wild-type and mutant YxaL proteins are shown in the table on the right. The individual raw data points of fluorescence displayed for each temperature represent the average of three end-point readings.

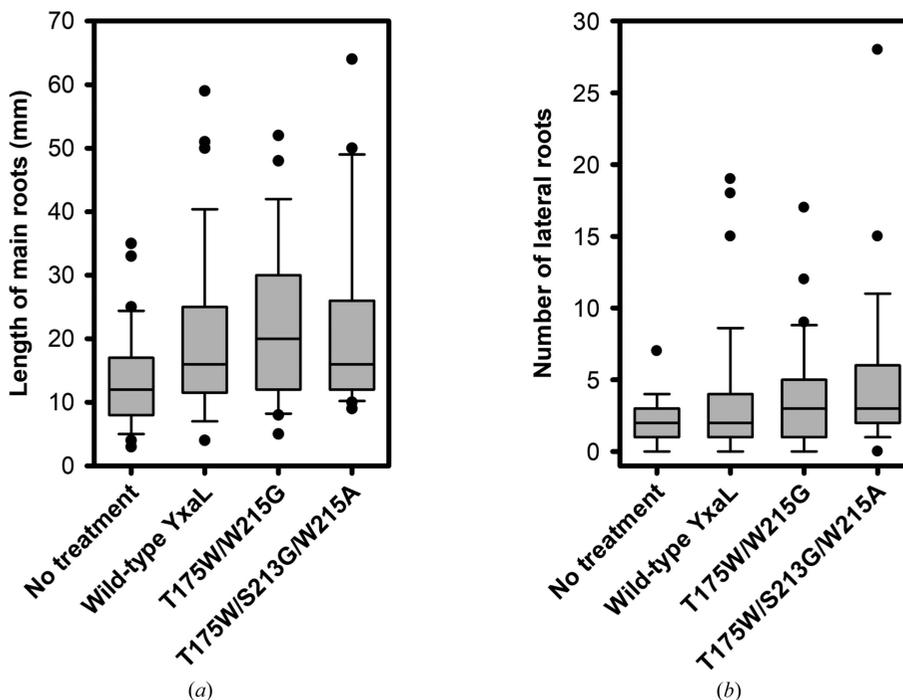


Figure 5 Effects of wild-type or mutant YxaL proteins on the main root length and lateral root number of two-week-old *A. thaliana* seedlings. Box plots are shown of the main root length (a) and lateral root number (b) of two-week-old seedlings planted after 2 h treatment with or without the addition of $1 \mu\text{g ml}^{-1}$ wild-type, T175W/W215G or T175W/S213G/W215A mutant protein to the soaking solution. Statistically significant differences between the data from untreated control and treated seed groups were determined by two-tailed Mann–Whitney U tests, with p values of less than 0.05.

difference between the two molecules suggested that the loop is flexible and is possibly exposed out of the structure. The indole ring of the mutated Trp175 did not make π interactions with the mutated Gly215 in either molecule (Fig. 3*a*). The intended tryptophan-docking interaction between the third and fourth blades was not observed in the T175W/W215G mutant. Interestingly, the loop connecting β -strand C4 to β -strand D4 in the well defined molecule of the asymmetric unit protruded from the β -propeller structure (Fig. 3*a*). The residues of this loop in one protomer seemed to interact with the bottom of the other protomer in the symmetry mate (Supplementary Fig. S4*b*).

3.6. Structural analysis of the T175W/S213G/W215A YxaL mutant protein

In the crystal structure of the T175W/S213G/W215A mutant, we found a putative tryptophan-docking interaction between Trp175 and Gly213 (Fig. 3*b*). The mutated residue Trp175 in the third blade was also surrounded by the hydrophobic residues Phe165, Val177 and Ala215 in the third and fourth blades. As we intended, Trp175 and Gly213 in the fourth blade resulted in the π interaction being observed in the conserved GXXXW motif. Interestingly, the removed indole ring at Trp215 made space for the mutated Trp175 in the tryptophan-docking interaction (Fig. 3*b*).

Differences between the restored and typical tryptophan-docking interactions were found in the side-chain conformation of the tryptophan residue. The indole ring of the tryptophan residue in the GXXXW motif of the preceding blade was buried between the former and the latter blades. However, the indole ring of the mutated Trp175 pointed out of the β -propeller structure compared with the GXXXW motif. The result seemed to be attributable to the side-chain conformation of Phe165, which appeared to fix Trp175 via hydrophobic interactions. Furthermore, the outermost β -strand D4 of the fourth blade was extended due to the π interaction between Trp175 and Gly213 (Fig. 3*b*). Despite the difference in the side-chain conformation of Trp175, the T175W/S213G/W215A mutant restored the tryptophan-docking interaction in the modified GXXXW motifs that was not found in the T175W/W215G mutant.

3.7. The engineered triple-mutant protein improved the structural stability, with better root-growth activity

We compared the melting temperature (T_m) values of the proteins. The T175W/S213G/W215A mutant YxaL protein exhibited a higher T_m value than the wild-type and T175W/W215G mutant YxaL proteins (Fig. 4). The triple mutant was also effective in the development of roots compared with the others that we tested. Consistently, the triple mutant T175W/S213G/W215A showed less sensitivity to proteolysis than the wild-type YxaL protein when the purified proteins were treated with trypsin (*t*-test, $p < 0.05$; Supplementary Fig. S5).

Finally, we investigated the effects on germination and root growth of treating seeds with wild-type, double-mutant or triple-mutant YxaL proteins prior to planting, using

A. thaliana as the model. No significant differences were observed between untreated control seeds and all wild-type and mutant YxaL-treated seeds in germination. Notably, all seed groups soaked with $1 \mu\text{g ml}^{-1}$ of the wild-type, double-mutant or triple-mutant protein for 2 h before planting showed significantly increased main root lengths of two-week-old seedlings compared with the untreated seed group (Fig. 5*a* and Supplementary Figs. S6*a* and S6*b*). Furthermore, treatment with the T175W/S213G/W215A mutant significantly increased the number of lateral roots compared with untreated seeds (Fig. 5*b* and Supplementary Fig. S6*b*). These results indicated that structure-based protein engineering could improve the root growth-promoting activity of YxaL, probably through increased structural stability.

4. Discussion

Protein engineering facilitates the production of valuable proteins from naturally occurring proteins, often by modifying their amino-acid sequences (Rahman *et al.*, 2015). Recently, the number of solved protein structures has increased exponentially with the development of protein structure-determination methods. The accumulating information on protein structures facilitates increasingly accurate structure-based protein engineering, applying structural knowledge to alter the structures and functions of proteins (Rahman *et al.*, 2015; Engqvist & Rabe, 2019). In this study, we determined the crystal structure of YxaL from *B. velezensis*, a plant growth-promoting rhizobacterium. The crystal structure of YxaL revealed an eight-bladed β -propeller fold stabilized by conserved GXXXW motifs in the outermost β -strand of each blade. The GXXXW motifs are responsible for a tryptophan-docking interaction formed between the tryptophan residue in the preceding blade and the glycine residue in the succeeding blade. The crystal structure of YxaL showed seven typical tryptophan-docking interactions and one atypical tryptophan-docking interaction between the third and fourth blades, which did not make π interactions due to their variant motifs. The tryptophan-docking interaction is evolutionarily conserved among PQQ family proteins to connect adjacent blades, indicating that the tryptophan-docking interaction optimally stabilizes the β -propeller fold (Ghosh *et al.*, 1995; Xia *et al.*, 1996; Neer & Smith, 2000; Jansen *et al.*, 2012). Thus, our findings suggested that the YxaL proteins have room for improvement in terms of structural stability by changing the atypical tryptophan-docking interaction.

YxaL is a secreted protein that can be affected by numerous factors in the soil, such as pH, temperature, inorganic ions and proteases, before it reaches its target cells on plant roots (Khare & Arora, 2015; Al-Kaisi *et al.*, 2017). Since the extracellular environment contains many proteases produced by soil bacteria, resistance to proteases may increase the chance of YxaL reaching plant root cells (Reardon *et al.*, 2018; Gotsmy *et al.*, 2021). We engineered the YxaL protein based on the crystal structure by restoring the typical tryptophan-docking interaction between the third and fourth blades, as confirmed by structural analysis. As a result, the engineered triple-mutant YxaL protein showed higher resistance to the

protease trypsin, with a higher T_m value, resulting in improved root-promoting activity on the model plant *A. thaliana*.

The β -propeller fold proteins are categorized into several groups depending on the conserved motifs that enhance the interactions between adjacent blades that stabilize the β -propeller structure (Chen *et al.*, 2011). However, the motifs are not always conserved in all of the blades. Many β -propeller fold proteins have variations in their conserved motifs like YxaL. For example, galactose oxidase has only three conserved Tyr-Trp-Thr-Asp (YWTD) motifs among its seven blades. The YWTD motif contributes to structural stability by an interblade hydrophobic interaction between the Tyr and Trp residues (Ito *et al.*, 1994). If the protein-engineering strategy used in YxaL is applied to this protein, we could expect an improvement in structural stability by restoring the hydrophobic interaction between the Tyr and Trp residues. Thus, our strategy would be widely applicable to other β -propeller fold proteins that naturally occur in diverse factors and enzymes.

It is necessary to identify the plant-derived protein that binds to YxaL to allow better engineering to promote plant growth, because tighter binding of YxaL to the plant protein would give more potent plant-growth activity. In general, the β -propeller fold proteins have ligand-binding sites in the top or the bottom regions, consisting of connecting loops between the β -strands. Since this study changed the side regions of the β -propeller fold, it is expected that the binding ability of YxaL to its binding-partner proteins would not be affected. Thus, the top or bottom regions should be engineered to make better proteins based on structural information on the binding surfaces. Currently, we are investigating the binding-partner proteins of YxaL to reveal the binding interface with the plant receptor proteins for further engineering of YxaL in the future. The improved activity of the engineered YxaL protein will help with crop growth, facilitating high availability in the agricultural industry, particularly in the treatment of seeds in smart hydroponics farming systems.

Acknowledgements

We made use of beamlines 5C and 11C at the Pohang Accelerator Laboratory, Pohang, Republic of Korea and the MALS facility at the Korea Basic Science Institute, Ochang, Republic of Korea.

Funding information

This study was supported by grants from the National Research Foundation of Korea (NRF-2017M3A9F6029755 to N-CH and NRF-2018R1A5A1025077 to Y-HK).

References

Al-Kaisi, M. M., Lal, R., Olson, K. R. & Lowery, B. (2017). *Soil Health and Intensification of Agroecosystems*, edited by M. M. Al-Kaisi & B. Lowery, pp. 1–23. London: Academic Press.

Anthony, C. & Ghosh, M. (1998). *Prog. Biophys. Mol. Biol.* **69**, 1–21.

Arkhipova, T., Veselov, S., Melentiev, A., Martynenko, E. & Kudoyarova, G. (2005). *Plant Soil*, **272**, 201–209.

Barnawal, D., Maji, D., Bharti, N., Chanotiya, C. S. & Kalra, A. (2013). *J. Plant Growth Regul.* **32**, 809–822.

Chaudhuri, I., Söding, J. & Lupas, A. N. (2008). *Proteins*, **71**, 795–803.

Chen, C. K.-M., Chan, N.-L. & Wang, A. H.-J. (2011). *Trends Biochem. Sci.* **36**, 553–561.

Dastogeer, K. M., Tumpa, F. H., Sultana, A., Akter, M. A. & Chakraborty, A. (2020). *Curr. Plant Biol.* **23**, 100161.

Duine, J. A. & Jongejan, J. A. (1989). *Annu. Rev. Biochem.* **58**, 403–426.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.

Engqvist, M. K. M. & Rabe, K. S. (2019). *Plant Physiol.* **179**, 907–917.

Ghosh, M., Anthony, C., Harlos, K., Goodwin, M. G. & Blake, C. (1995). *Structure*, **3**, 177–187.

Gotsmy, M., Escalona, Y., Oostenbrink, C. & Petrov, D. (2021). *Proteins*, **89**, 925–936.

Haar, E. ter, Musacchio, A., Harrison, S. C. & Kirchhausen, T. (1998). *Cell*, **95**, 563–573.

Hirose, I., Sano, K., Shioda, I., Kumano, M., Nakamura, K. & Yamane, K. (2000). *Microbiology*, **146**, 65–75.

Hudson, A. M. & Cooley, L. (2008). *The Coronin Family of Proteins*, edited by C. S. Clemen, L. Eichinger & V. Rybakina, pp. 6–19. Austin: Landes Bioscience.

Ito, N., Phillips, S. E., Yadav, K. D. & Knowles, P. F. (1994). *J. Mol. Biol.* **238**, 794–814.

Jansen, K. B., Baker, S. L. & Sousa, M. C. (2012). *PLoS One*, **7**, e49749.

Kang, S.-M., Radhakrishnan, R., You, Y.-H., Joo, G.-J., Lee, I.-J., Lee, K.-E. & Kim, J.-H. (2014). *Indian J. Microbiol.* **54**, 427–433.

Khare, E. & Arora, N. K. (2015). *Plant Microbes Symbiosis: Applied Facets*, edited by N. K. Arora, pp. 353–381. New Delhi: Springer India.

Kilian, M., Steiner, U., Krebs, B., Junge, H., Schmiedeknecht, G. & Hain, R. (2000). *Pflanzenschutz-Nachr. Bayer*, **53**, 72–93.

Kim, Y.-H., Choi, Y., Oh, Y. Y., Ha, N.-C. & Song, J. (2019). *PLoS One*, **14**, e0207968.

Kopec, K. O. & Lupas, A. N. (2013). *PLoS One*, **8**, e77074.

Kuan, K. B., Othman, R., Abdul Rahim, K. & Shamsuddin, Z. H. (2016). *PLoS One*, **11**, e0152478.

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.* **D75**, 861–877.

Neer, E. J. & Smith, T. F. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 960–962.

Noirot-Gros, M. F., Soultanas, P., Wigley, D., Ehrlich, S., Noirot, P. & Petit, M. A. (2002). *Mol. Genet. Genomics*, **267**, 391–400.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Radhakrishnan, R., Hashem, A. & Abd Allah, E. F. (2017). *Front. Physiol.* **8**, 667.

Radhakrishnan, R. & Lee, I.-J. (2016). *Plant Physiol. Biochem.* **109**, 181–189.

Rahman, M. M., Sarker, I., Elmenoufy, A. H., Bahlol, H. S. & Zaman, S. (2015). *Sch. J. Appl. Med. Sci.* **3**, 1224–1237.

Raran-Kurussi, S., Cherry, S., Zhang, D. & Waugh, D. S. (2017). *Methods Mol. Biol.* **1586**, 221–230.

Reardon, P. N., Walter, E. D., Marean-Reardon, C. L., Lawrence, C. W., Kleber, M. & Washton, N. M. (2018). *Sci. Rep.* **8**, 813.

Ruij, L. (2020). *Agronomy*, **10**, 861.

Smith, T. F. (2008). *The Coronin Family of Proteins*, edited by C. S. Clemen, L. Eichinger & V. Rybakina, pp. 20–30. Austin: Landes Bioscience.

Xia, Z., Dai, W., Zhang, Y., White, S. A., Boyd, G. D. & Mathews, S. F. (1996). *J. Mol. Biol.* **259**, 480–501.

Xie, S.-S., Wu, H.-J., Zang, H.-Y., Wu, L.-M., Zhu, Q.-Q. & Gao, X.-W. (2014). *Mol. Plant Microbe Interact.* **27**, 655–663.