

# 1 Expression, purification, and crystallization of a plant polyketide cyclase 2 from *Cannabis sativa*

3 Xinmei Yang<sup>a</sup>, Takashi Matsui<sup>a</sup>, Takahiro Mori<sup>b</sup>, Futoshi Taura<sup>c</sup>, Hiroshi Noguchi<sup>d</sup>, Ikuro Abe<sup>b\*</sup>  
4 and Hiroyuki Morita<sup>a\*</sup>

5 <sup>a</sup>Institute of Natural Medicine, University of Toyama, 2630-Sugitani, Toyama, Toyama, 930-0194, Japan

6 <sup>b</sup>Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

7 <sup>c</sup>Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama, 2630-Sugitani, Toyama, Toyama, 930-0194, Japan

8 <sup>d</sup>School of Pharmaceutical Sciences, University of Shizuoka, 52-1 Yada, Suruga, Shizuoka, 422-8526, Japan

9 Correspondence email: [abei@mol.f.u-tokyo.ac.jp](mailto:abei@mol.f.u-tokyo.ac.jp); [hmorita@inm.u-toyama.ac.jp](mailto:hmorita@inm.u-toyama.ac.jp)

10 **Keywords:** Cannabinoids; *Cannabis sativa*; Cyclase; Olivetolic acid.

## 11 Synopsis

12 Olivetolic acid cyclase (OAC) from *Cannabis sativa* L. has been overexpressed in *E. coli*, purified, and crystal-  
13 lized. Diffraction data have been collected to a resolution of 1.40 Å.

## 14 Abstract

15 Plant polyketides are a structurally diverse family of natural products. In the biosynthesis of plant polyketides, the con-  
16 struction of the carbocyclic scaffolds is a key step for diversifying the polyketide structure. Olivetolic acid cyclase (OAC)  
17 from *Cannabis sativa* L. is the only known plant polyketide cyclase that catalyzes the C2/C7 intramolecular aldol cycliza-  
18 tion of linear pentyl tetra- $\beta$ -ketide CoA to generate olivetolic acid in the biosynthesis of cannabinoid. The enzyme is also  
19 thought to belong to the dimeric  $\alpha+\beta$  barrel (DABB) protein family. However, because of lack of the functional analysis  
20 of the other plant DABB proteins and low sequence identity with the functionally and structurally characterized bacterial  
21 DABB proteins, the catalytic mechanism of OAC has remained unclear. To clarify the intimate catalytic mechanism of  
22 OAC, the enzyme was overexpressed in *Escherichia coli* and crystallized in a vapour-diffusion method. The crystals  
23 diffracted X-rays to 1.40 Å resolutions and belonged to space group  $P3_121$  or  $P3_221$ , with unit-cell parameters  $a = b =$   
24  $47.3$  Å,  $c = 176.0$  Å. Further crystallographic analysis will provide valuable insights into the structure–function relation-  
25 ship and catalytic mechanism of OAC.

## 26 1. Introduction

27 Plant polyketides, such as flavonoids, stilbenes, and phloroglucinols, are one of the largest and most important families  
28 of natural products, with remarkable structural diversity and biological activities. In the biosynthesis of plant  
29 polyketides, the construction of the carbocyclic scaffolds is a key step for diversifying the polyketide structures. This  
30 process is generally catalyzed by a type III polyketide synthase (PKS), by carbon chain elongation and subsequent  
31 cyclization of the highly reactive poly- $\beta$ -keto intermediate (Schröder, 1999; Austin & Noel, 2003; Abe & Morita, 2010).  
32 Recent studies have suggested that the biosyntheses of plant polyketides, such as the anthranoid (Abe et al., 2005; Abdel-  
33 Rahman et al., 2013) and cannabinoid (Taura et al., 2009; Gagne et al., 2012), produced by *Aloe arborescens*  
34 and *Cannabis sativa*, respectively, require additional enzymes for proper folding cyclization of the linear poly- $\beta$ -  
35 keto intermediate to generate the final products, as in the cases of bacterial polyketide biosyntheses.

36 The recently identified olivetolic acid cyclase (OAC) from *C. sativa* is a novel polyketide cyclase that is proposed  
37 to be involved in the biosynthesis of cannabinoid. The enzyme is the only known plant polyketide cyclase, and is thought  
38 to accept the linear pentyl tetra- $\beta$ -ketide CoA, produced by a type-III PKS tetraketide synthase (TKS), as the substrate,  
39 and perform the C2–C7 aldol cyclization, thioester bond cleavage, and aromatization reactions to generate olivetolic  
40 acid (OA), without requiring any co-factors (Fig. 1, Gagne et al., 2012). OAC shares 30-48% identity with  
41 functionally unidentified plant dimeric  $\alpha$ + $\beta$  barrel (DABB) proteins, such as the structurally characterized heat stable pro-  
42 tein (AtHS1) from *Arabidopsis thaliana* (48% identity, Bingman et al., 2004; Lytle et al., 2004), boiling stable protein  
43 (SP1) from *Populus tremula* (38% identity, Dgany et al., 2004), and At5g22580 from *A. thaliana* (32% identity,  
44 Cornilescu et al., 2004). OAC also shares low sequence identity (less than 20%) with bacterial DABB proteins, such as  
45 the structurally characterized bacterial polyketide cyclase, tetracenomyacin F2 cyclase (Tcm I) from *Streptomyces*  
46 *glaucescens* (17% identity, Thompson et al., 2004), and the functionally distinct, structurally characterized ActVA-Orf6  
47 monooxygenase from *S. coelicolor* (15% identity, Sciara et al., 2003) and 4-methylmuconolactone methylisomerase  
48 (MLMI) from *Pseudomonas reinekei* (13% identity, Marin et al., 2009).

49 On the basis of these findings and a comparison of the homology model of OAC with the three-dimensional  
50 structures of the plant and bacterial DABB proteins, OAC has been proposed to be a homodimeric protein consisting of  
51 12.2 kDa subunits and to possess a hydrophobic tunnel as the active site cavity in each monomer (Gagne et al., 2012), as  
52 in the cases of the other structurally characterized DABB proteins. Furthermore, site-directed mutagenesis studies have  
53 suggested that three His residues (His5, His57, and His78) play crucial roles in the OA-forming activity (Gagne et al.,  
54 2012). However, the catalytic mechanism underlying the substrate and product specificities, the aldol cyclization and  
55 aromatization reactions, and the thioester bond cleavage, as well as the catalytic role of the three His residues in the OA-  
56 forming activity, have remained unclear. Therefore, to further clarify the intimate structural details of the OAC catalyzed  
57 reaction, we expressed glutathione S-transferase (GST)-fused recombinant OAC in *Escherichia coli*, removed the GST  
58 portion, and obtained good-quality crystals of the recombinant OAC.

## 59 2. Materials and methods

### 60 2.1. Expression and purification

61 The cDNA encoding full-length OAC, with *SmaI/SalI* sites just before the initiation codon and after the stop  
62 codon, respectively, was purchased from Eurofins Genomics (Table 1). The cDNA was digested with *SmaI/SalI* and was  
63 ligated into the *SmaI/SalI* sites of the modified pQE-80L vector (QIAGEN), for expression as a fusion protein with GST  
64 at the N-terminus. A PreScission Protease (GE Healthcare) cleavage site was introduced between GST and OAC. The  
65 constructed expression plasmid was transformed into *E. coli* M15 cells (QIAGEN), and the cells harboring the plasmid  
66 were cultured to an OD<sub>600</sub> of 0.6 in Luria broth (LB) medium, containing 100  $\mu\text{g ml}^{-1}$  ampicillin, at 310 K. Isopropyl  $\beta$ -  
67 D-1-thiogalactopyranoside was added to a final concentration of 1 mM to induce protein expression, and the culture  
68 was incubated at 290 K for a further 20 h.

69 All of the following procedures were performed at 277 K. The *E. coli* cells were harvested by centrifugation at 5,000  
70 g for 20 min, and then resuspended in 50 mM Tris-HCl buffer, pH 8.0, containing 200 mM NaCl, 5%(v/v) glycerol and  
71 2 mM DTT (buffer A). The cells were disrupted by sonication, and the lysate was centrifuged at 6,000 g for 10 min. The  
72 supernatant was loaded on a COSMOGEL GST-Accept Resin column (Nacalai Tesque) equilibrated with buffer A. After  
73 the column was washed with 20 mM HEPES-NaOH buffer, pH 7.5, containing 100 mM NaCl, 5%(v/v) glycerol and 2  
74 mM DTT (buffer B), the GST-tag was cleaved on the column by PreScission Protease (GE Healthcare) overnight, and  
75 OAC was eluted with buffer B. The resultant protein thus contains three additional N-terminal residues (GPG), derived

76 from the PreScission Protease recognition sequence. The protein solution was diluted ten-fold with 20 mM HEPES-  
77 NaOH buffer, pH 7.5, containing 2 mM DTT (buffer C), and then applied to a Resource Q column (GE Healthcare). The  
78 column was washed with buffer C containing 10 mM NaCl, and the protein was subsequently eluted at 50 mM NaCl,  
79 using a linear gradient of 10-200 mM NaCl. The protein was further purified by size-exclusion chromatography on a  
80 HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare), and was concentrated to 17 mg ml<sup>-1</sup> in 20 mM HEPES-  
81 NaOH buffer, pH 7.5, containing 25 mM NaCl.

82 A dynamic light-scattering (DLS) analysis was performed, using a DynaPro-MSXTC molecular-sizing  
83 instrument (Protein Solutions). After centrifugation through a 0.22 μm Ultrafree-MC filter (Millipore) to remove partic-  
84 ulate material from the protein solution, the solution properties of the purified protein were monitored. Data were ac-  
85 quired as 50 scattering measurements at 293 K. Data from five sets of measurements were analyzed using the  
86 DYNAMICS software package (Protein Solutions) and averaged.

## 87 2.2. Crystallization

88 Initial crystallization conditions were found by screening with the commercial JCSG+ Suite screening kit  
89 (QIAGEN). The crystallizations were performed at 278 K, using the sitting-drop vapour-diffusion method with a 96-well  
90 plate. All crystallization drops were prepared by mixing 1 μl of protein solution (17 mg ml<sup>-1</sup>) with an equal volume of  
91 reservoir solution, and were equilibrated against 50 μl reservoir solution. Clusters of crystals were observed three days  
92 later, in the crystallization condition consisting of 100 mM Bicine, pH 8.5, and 20%(w/v) PEG 6000. Further  
93 crystallization was attempted using Additive Screen HT (Hampton Research) at various pH values, together with the use  
94 of 20%–30%(w/v) PEG 6000 as a precipitant. Diffraction-quality crystals were finally obtained at 278 K, in 100  
95 mM Tris-HCl, pH 8.8, 25%(w/v) PEG 6000, and 100 mM sodium malonate, using the sitting-drop vapour-diffusion  
96 method (Table 2).

## 97 2.3. Data collection and processing

98 Single crystals were transferred into a crystallization solution containing 10%(v/v) glycerol as a cryoprotectant, picked up  
99 with a nylon loop, and then flash-cooled at 100 K in a nitrogen-gas stream. X-ray diffraction data sets were collected on  
100 Beamline NW-12A at the Photon Factory (PF; wavelength 1.0000 Å) using an ADSC Q210r detector, with a distance of  
101 117.1 mm between the crystal and the detector. A total of 360 frames were recorded, with a 0.5° oscillation angle and 1 s  
102 exposure time. The data were indexed, integrated, and scaled with the *XDS* program package (Kabsch, 2010).  
103 The Matthews volume ( $V_M$ ) and the solvent content were calculated with *Xtriage* (Adams et al., 2010).

## 104 3. Results and discussion

105 The recombinant OAC was overexpressed in *E. coli*, as a fusion protein with a GST-tag at the N-terminus. After  
106 cleavage of the GST-tag, the purified OAC migrated as a single band with a molecular weight of 10 kDa upon 20%(w/v)  
107 SDS-PAGE, in agreement with the calculated molecular weight of 12.2 kDa (Fig. 2a). In contrast, the gel-filtration  
108 experiment by size exclusion chromatography indicated a molecular weight of 25 kDa (Fig. 2b). In addition to the gel-  
109 filtration analysis, the DLS analysis after the size-exclusion chromatography revealed a monomodal distribution, with a  
110 polydispersity value of 9.5% and a molecular weight estimate of 23 kDa, which allowed us to confirm that the  
111 recombinant OAC is a homodimeric protein, and also suggested that OAC is a member of the DABB protein family. The  
112 typical yield of the purified recombinant OAC was about 1.5 mg per liter of culture.

113 The OAC crystals appeared reproducibly in the optimized crystallization solution within a few days, with various  
114 sizes up to approximate dimensions of 0.15 × 0.15 × 0.12 mm. Among them, only the crystals with approximate dimen-  
115 sions of 0.07 × 0.07 × 0.06 mm gave good quality diffraction up to 1.40 Å resolution (Fig. 3). The detailed data collection

116 statistics are summarized in Table 3. The preliminary crystallographic analysis indicated that the crystals belonged to the  
 117 space group  $P3_121$  or  $P3_221$ , with unit-cell parameters  $a = b = 47.3 \text{ \AA}$ ,  $c = 176.0 \text{ \AA}$ . With two monomers in the asymmet-  
 118 ric unit, the Matthews volume ( $V_M$ ) was calculated to be  $2.5 \text{ \AA}^3 \text{ Da}^{-1}$  and the estimated solvent content was 50.7%, which  
 119 is in the range normally observed for protein crystals. Structure determination by the molecular-replacement method is  
 120 currently under way, using the *Phaser* (McCoy et al., 2007) and *Molrep* (Vagin & Teplyakov, 2010) programs with the  
 121 crystal structures of the plant DABB proteins, AtHS1 (PDB entry 1q53), At5g22580 (PDB entry 1rjj), and SP1 (PDB  
 122 entry 1si9), and the other bacterial homologues, such as TcmI (PDB entry 1tuw), ActVA-Orf6 (PDB entry 1lq9), and  
 123 MLMI (PDB entry 2ifx) as search models. Simultaneously, we are also attempting to crystallize OAC complexed with its  
 124 product and product analogues, together with the expression, purification, and crystallization of selenomethionine-labeled  
 125 OAC. These structural analyses will provide valuable insights into not only the structure–function relationship and  
 126 catalytic mechanism of OAC, but also the functional diversity of the DABB proteins.

127 **Table 1**

128 Macromolecule production information

129	Source organism	<i>Cannabis sativa</i> L.
130	DNA source	Chemical synthesis
131	Cloning vector	pEX-A
132	Expression vector	pQE-80L
133	Expression host	M15 (QIAGEN)
134	Complete amino acid sequence of the construct produced†	<i>MRGSHHHHHGSMSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYE RDEGDKWRNKKFELGL              EFPNLPYYIDGDVKTQSMAIIRYIADKHN MLGGCPKERAIEISMLEGAVLDIRYGVSR IAYSKDF              ETLKVDFLSKLP EMLKMFEDRLCHKTYLNGDHVTHPDFMFLYDALDVVLYMDP MCLD AFPKL              VCFKKRIE AIPQIDKYLKSSKYLA WPLQGWQATFGGGDHPPK              SDLEVL FQGP GMAVKHLIVLKFKDEITEAQKEEFFKTYVNLVNIIPAMKDVYWGKDVTQK              NKEEGYTHIVEVTFESVETIQDYIIHPAHVGF GDVYRSFWEKLLIFDYTPRK</i>

135 †The amino-acid sequence of glutathione S-transferase (GST) tag is shown in italics. PreScission protease recognition site (LEVL FQGP) is underlined. The  
 136 peptide bond between the Q and G residues is cleaved by PreScission protease.

137 **Table 2**

138 Crystallization

139	Method	sitting-drop vapour-diffusion method
140	Plate type	96 well CrystalQuick, 288 square wells, flat bottom (greiner bio-one)
141	Temperature (K)	278
142	Protein concentration	17 mg ml <sup>-1</sup>
143	Buffer composition of protein solution	20 mM HEPES-NaOH buffer, pH 7.5, and 25 mM NaCl
144	Composition of reservoir solution	100 mM Tris-HCl, pH 8.8, 25%(w/v) PEG 6000, and 100 mM sodium malonate
145	Volume and ratio of drop ( $\mu$ l)	1:1
146	Volume of reservoir ( $\mu$ l)	50

147 **Table 3**

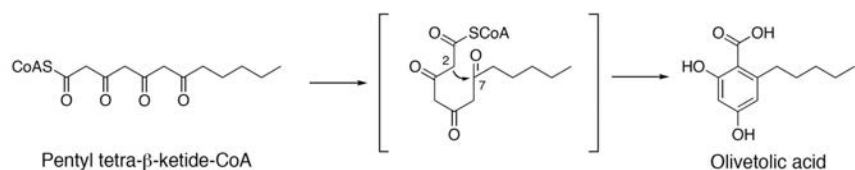
148 Data collection and processing

149 Values for the outer shell are given in parentheses.

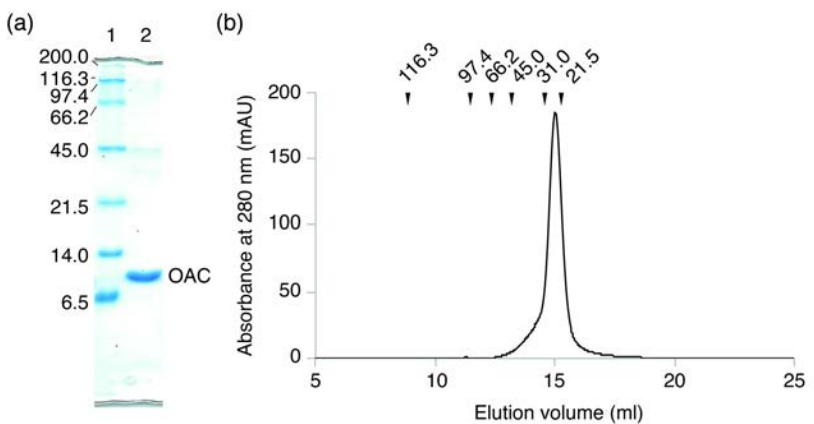
150	Diffraction source	NW-12A, PF
151	Wavelength ( $\text{\AA}$ )	1.0000
152	Temperature (K)	100
153	Detector	ADSC Q210r

154	Crystal-detector distance (mm)	117.1
155	Rotation range per image (°)	0.5
156	Total rotation range (°)	180
157	Exposure time per image (s)	1
158	Space group	$P3_121$ or $P3_221$
159	a, b, c (Å)	47.3, 47.3, 176.0
160	Mosaicity (°)	0.193
161	Resolution range (Å)	50.0 - 1.40 (1.48 - 1.40)
162	Total No. of reflections	475,525 (74,586)
163	No. of unique reflections	45,736 (7,287)
164	Completeness (%)	98.9 (98.9)
165	Redundancy	10.4 (10.2)
166	$\langle I/\sigma(I) \rangle$	24.2 (12.3)
167	$R_{merge}$	0.074 (0.155)
168	Overall $B$ factor from Wilson plot (Å <sup>2</sup> )	10.19
169	No. molecules per AU	2
170	$V_M$ (Å <sup>3</sup> Da <sup>-1</sup> )	2.5
171	$V_{solv}$ (%)	50.7

172  $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ , where  $I(hkl)$  is the intensity of reflection  $hkl$ ,  $\sum_{hkl}$  is the sum over all  
 173 reflections, and  $\sum_i$  is the sum over  $i$  measurements of reflection  $hkl$ .

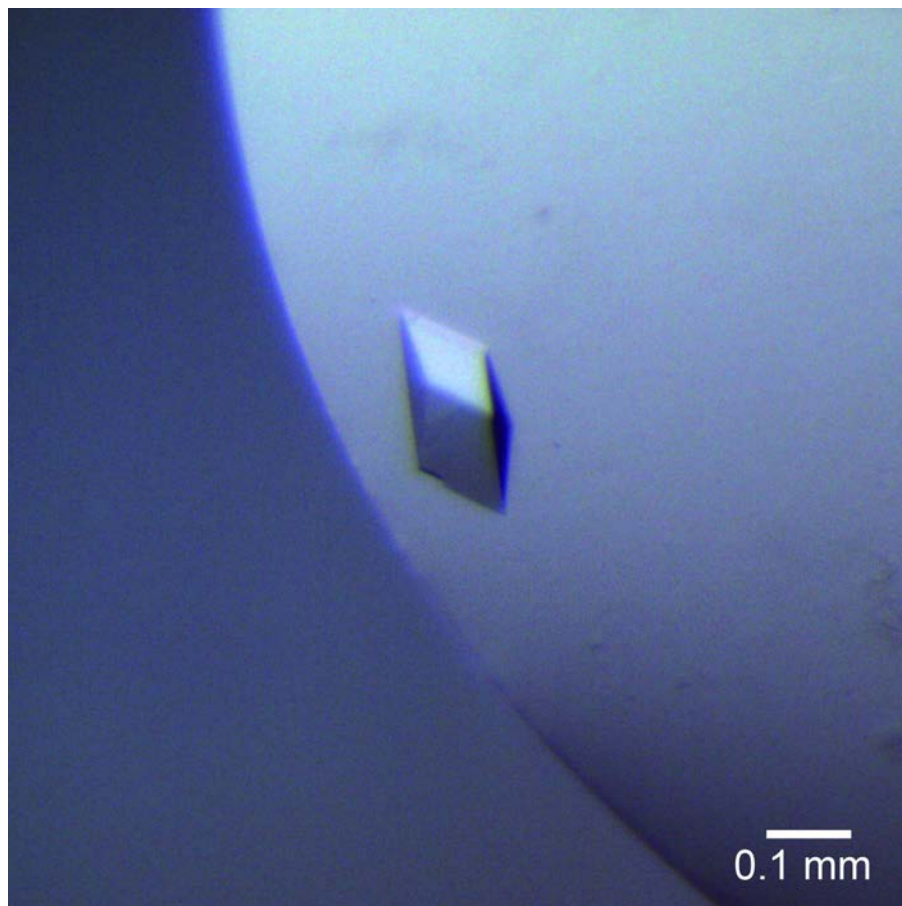


174 **Figure 1**  
 175 Proposed mechanism for the formation of olivetolic acid by OAC.



176 **Figure 2**  
 177 Analysis of molecular weight of OAC. (a) SDS-PAGE gel is shown. Lane 1, molecular weight marker; lane 2, peak  
 178 fraction. Molecular weight markers are labelled in kDa. (b) Chromatogram of size-exclusion chromatography is shown.  
 179 Peak represents OAC. Molecular weights (kDa) of marker proteins are indicated on the top of the chromatogram.

7\_original.jpg



180 **Figure 3** ↻

181 Crystals of OAC. The dimensions of the crystals were approximately  $0.07 \times 0.07 \times 0.06$  mm.

### 182 Acknowledgements

183 This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education,  
184 Culture, Sports, Science and Technology, Japan (to H.N., I.A., and H.M.), by a grant from Japan Foundation for  
185 Applied Enzymology (to H.M.), and by a grant from NAGASE Science Technology Foundation (to H.M.). The authors  
186 declare that they have no competing financial interests. We thank the beamline staff at the Photon Factory for their  
187 assistance with X-ray diffraction data collection.

### 188 References

- 189 Abe, I. & Morita, H. (2010). *Nat. Prod. Rep.* 27, 809–838.
- 190 Abe, I., Oruro, S., Utsumi, Y., Sano, Y. & Noguchi, H. (2005). *J. Am. Chem. Soc.* 127, 12709–12716.
- 191 Abdel-Rahman, I.A.M., Beuerle, T., Ernst, L., Abdel-Baky, A. M., Desoky, E. E-D. K., Ahmed, A. S. & Beerhues,  
192 L. (2013). *Phytochemistry* 88, 15–24.
- 193 Adams, O. D., Afonine, P. V., Bunkóczi, Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G.  
194 J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J.

- 195 S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Crystallogr. D. Biol. Crystallogr.* 66, 213–221.
- 196 Austin, M. B. & Noel, J. P. (2003). *Nat. Prod. Rep.* 20, 79–110.
- 197 Bingman, C. A., Johnson, K. A., Peterson, F. C., Frederick, R. O., Zhao, Q., Thao, S., Fox, B. G., Volkman, B. F., Jeon,  
198 W. B., Smith, D. W., Newman, C. S., Ulrich, E. L., Hereman, A., Sussman, M. R., Markley, J. L. & Phillips, G. N.  
199 (2004). *Proteins* 57, 218–220.
- 200 Cornilescu, G., Cornilescu, C. C., Zhao, Q., Frederick, R. O., Peterson, F. C., Thao, S. & Markley, J. L. (2004). *J.*  
201 *Biomol. NMR.* 29, 387–390.
- 202 Dgany, O., Gonzalez, A., Sofer, O., Wang, W., Zolotnitsky, G., Wolf, A., Shoham, Y., Altman, A., Wolf, S. G.,  
203 Shoseyov, O. & Almog, O. (2004). *J. Biol. Chem.* 279, 51516–51523.
- 204 Gagne, S. J., Stout, J. M., Liu, E., Boubakir, Z., Clark, S. M. & Page, J. E. (2012). *Proc. Natl. Acad. Sci. U. S.*  
205 *A.* 109, 12811–12816.
- 206 Kabsch, W. (2010). *Acta Crystallogr. D. Biol. Crystallogr.* 66, 125–132.
- 207 Lytle, B., Peterson, F., Kjer, K., Frederick, R., Zhao, Q., Thao, S., Bingman, C., Johnson, K., Phillips Jr., G. &  
208 Volkman, B. (2004). *J. Biomol. NMR* 28, 397–400.
- 209 Marin, M., Heinz, D. W., Pieper, D. H. & Klink, B. U. (2009). *J. Biol. Chem.* 284, 32709–32716.
- 210 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J.*  
211 *Appl. Crystallogr.* 40, 658–674.
- 212 Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., Schimperna, G., Savino, C. &  
213 Vallone, B. (2003). *EMBO J.* 22, 205–215.
- 214 Schröder, J. (1999). *Comprehensive Natural Products Chemistry*, Vol. 2, edited by D. E. Cane, pp. 749–771.  
215 Oxford: Elsevier.
- 216 Taura, F., Tanaka, S., Taguchi, C., Fukamizu, T., Tanaka, H., Shoyama, Y. & Morimoto, S. (2009). *FEBS Lett.* 583, 2061–  
217 2066.
- 218 Thompson, T. B., Katayama, K., Watanabe, K., Hutchinson, C. R. & Rayment, I. (2004). *J. Biol. Chem.* 279, 37956–  
219 37963.
- 220 Vagin, A. & Teplyakov, A. (2010). *Acta Crystallogr. D. Biol. Crystallogr.* 66, 22–25.