Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray crystallographic analysis of a cysteine-rich secretory protein (CRISP) from *Naja atra* venom

Yu-Ling Wang,[a,b] King-Xiang Goh,[b] Wen-guey Wu[b]* and Chun-Jung Chen[a,b,c]*

[a]Biology Group, Research Division, National Synchrotron Radiation Research Center, Hsinchu 30077, Taiwan, [b]Department of Life Sciences and Structural Biology Program, National Tsing-Hua University, Hsinchu 30077, Taiwan, and [c]Department of Physics, National Tsing-Hua University, Hsinchu 30077, Taiwan

Correspondence e-mail: wgwu@life.nthu.edu.tw, cjchen@nsrrc.org.tw

Received 26 May 2004
Accepted 10 August 2004

Cysteine-rich secretory proteins (CRISPs) play an important role in the innate immune system and are transcriptionally regulated by androgens in several tissues. The proteins are mostly found in the epididymis and granules of mammals, whilst a number of snake venoms also contain CRISP-family proteins. The natrin protein from the venom of *Naja atra* (Taiwan cobra), which belongs to a family of CRISPs and has a cysteine-rich C-terminal amino-acid sequence, has been purified using a three-stage chromatography procedure and crystals suitable for X-ray analysis have been obtained using the hanging-drop vapour-diffusion method. X-ray diffraction data were collected to 1.58 Å resolution using synchrotron radiation; the crystals belong to space group $C222_1$, with unit-cell parameters $a = 59.172$, $b = 65.038$, $c = 243.156$ Å. There are two protein molecules in the asymmetric unit and the Matthews coefficient is estimated to be 2.35 Å^3 Da^{-1}, corresponding to a solvent content of 47.60%.

1. Introduction

About 2900 species of snakes are presently known in the world and they possess various venoms composed of a number of proteins and peptides. From known snake venoms, eight snake-toxin families, Kunitz-type protease inhibitors, CRISP toxins, galactose-binding lectins, M12Bs, cystatin toxins, nerve growth-factor toxins, phospholipase A_2s and natriuretic toxins, have been classified and described (Fry & Wuster, 2004). Cysteine-rich secretory proteins (CRISPs), an emerging superfamily of secretory proteins, are characterized by 16 highly conserved cysteine residues, of which ten are concentrated in C-terminal 56 amino acids of the proteins, and share 40–80% sequence similarity, indicating that the CRISP family is a highly conserved protein family (Kratzschmar *et al.*, 1996) and suggesting the presence of a common structural motif.

The CRISP family also exists in humans and mammals and can be divided into three groups: CRISP-1, such as the sperm-coating glycoproteins DE, also termed acidic epididymal glycoproteins (AEG) and expressed mainly in epididymis (Eberspaecher *et al.*, 1995; Sivashanmugam *et al.*, 1999), CRISP-2, such as testis-specific spermatocytes protein (TPX1) and autoantigen 1 (AA1), expressed primarily in testis (Foster & Gerton, 1996), and CRISP-3, widely distributed in various cells and tissues of mammal, such as salivary glands, pancreas, prostate, B-cell neutrophils and eosinophils (Kjeldsen *et al.*, 1996; Pfisterer *et al.*, 1996; Udby *et al.*, 2002; Kosari *et al.*, 2002). CRISP-1 and CRISP-2 proteins are reported to play roles in fertilization process, sperm maturation and the adhesion of spermatogenic cells to sertoli cells (Busso *et al.*, 2003; Cohen *et al.*, 2000; Maeda *et al.*, 1998; Maeda *et al.*, 1999; Schwidetzky *et al.*, 1995), whereas the exact functions of CRISP-3 proteins are still largely unknown.

CRISP-related proteins are also found in the venoms of snakes and lizards, the sequences of which show a high similarity to the mammalian CRISP family. Helothermine (HLX) isolated from the saliva of *Heloderma horridum*, the first venomous CRISP-related protein, has been reported to block the voltage-gate calcium channel, voltage-gate potassium channel and skeletal muscle ryanodine receptor (Mochca-Morales *et al.*, 1990; Nobile *et al.*, 1994, 1996; Morrissette *et al.*, 1995). Recently, a number of CRISP-related venom proteins have been isolated from Viperidae and Elapidae snakes, indicating that CRISPs might be widely distributed in snake venoms. Pseudechetoxin, a CRISP-related protein in the venom of Australian King Brown snake was reported to inhibit current through cyclic nucleotide-gated (CNG) channels (Brown *et al.*, 1999, 2003; Yamazaki, Brown *et al.*, 2002). Others CRISP-related snake-venom proteins, such as ablomin (*A. blomhoffi*), triflin (*T. flavoviridis*) and latisemin (*L. semifasciata*) have been shown to inhibit the contraction of smooth muscle (Yamazaki, Koike *et al.*, 2002). Several structural analyses of other cysteine-rich proteins from venoms have recently been

reported, including disintegrins from *Agkistrodon* (Fujii *et al.*, 2002, 2003; Moiseeva *et al.*, 2002) and *Echis carinatus* viper (Tomar *et al.*, 2001), applaggin (*A. piscivorus piscivorus* platelet-aggregation inhibitor) from North American water moccasin (Arni *et al.*, 1999) and EMS16, an antagonist of collagen receptor from *Echis multisquamatus* (Horii *et al.*, 2003).

Many CRISP-family members from snake venoms have been found and isolated, such as those from the eastern cottonmouth, king cobra, western diamondback rattlesnake, Taiwan habu *etc*., but the actual functions of these proteins remain puzzling (Yamazaki *et al.*, 2003; Chang *et al.*, 1997). NA-CRVP1, also called natrin, is a CRISP-related protein isolated from the venom of the Taiwan cobra (*Naja atra*), which displays a high amino-acid sequence similarity with ophanin from *Ophiophagus hannah*. Previous study has shown that the protein exhibited no lethality, myotoxicity or enzymatic activities (Jin *et al.*, 2003). We have identified the natrin protein in crude cobra venom and established a successful protein-purification procedure that produces large amounts of protein for crystallization and structure study. The protein is composed of a single polypeptide chain of 239 amino acids containing a total of 16 conserved cysteine residues typical of the CRISP family. To further understand the biological significance of natrin in the venom, it is important and necessary to understand its detailed three-dimensional structure–function relationship. Furthermore, this may also provide common structural folding information for other proteins of the highly conserved CRISP family. In this paper, we report the purification, crystallization and preliminary X-ray characterization of natrin. To our knowledge, it is the first crystal of a CRISP-family protein with a full-length sequence.

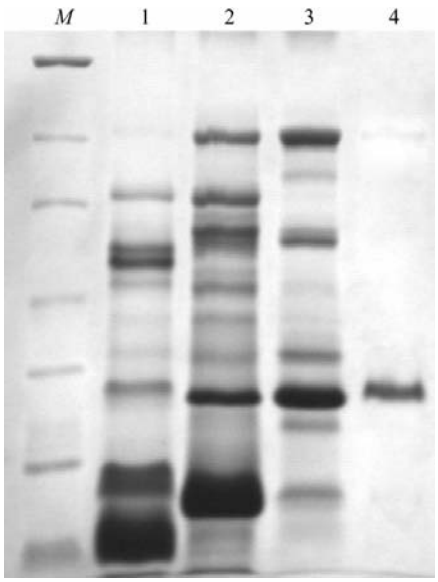

Figure 1
Coomassie-blue stained 10% SDS gel of pools from natrin purification. Lane *M*, molecular-weight markers in kDa; lane 1, pool from crude venom; lane 2, eluted from SP Sephadex C-25 column; lane 3, eluted from DEAE Sepharose fast flow column; lane 4, purified natrin from CM Sepharose fast flow column.

2. Materials and methods

2.1. Protein purification

The crude venom (15 ml) was collected directly from about 100 *N. atra* cobras at a snake farm in Tainan City, Taiwan and lyophilized before further use. The natrin protein was purified using a three-step chromatography procedure consisting of Sephadex C25, DEAE Sepharose fast flow and CM Sepharose fast flow (Amersham-Pharmacia Biotech) ion-exchange chromatography on an FPLC system (ÄKTA purifier, Amersham-Pharmacia Biotech) at 277 K. Lyophilized *N. atra* venom (1 g) was first dissolved in 10 ml 50 m*M* sodium phosphate buffer pH 7.4. After centrifugation (6000 rev min^{-1}) and filtration (0.22 mm pore filter) to remove insoluble materials, the venom was applied onto the first SP-Sephadex C-25 column, which was pre-equilibrated with the same sodium phosphate buffer. The venom was eluted at a flow rate of 1 ml min^{-1} using an NaCl gradient (0–0.5 *M*) and collected with a volume of 12 ml per tube. All following purification steps were monitored by 10% SDS–PAGE. The fractions containing natrin protein were then passed through the second DEAE Sepharose fast flow column equilibrated with 20 m*M* Tris buffer pH 7.4 and eluted at a flow rate of 1 ml min^{-1} using an NaCl gradient (0–0.5 *M*). The natrin fractions (15 ml per tube) were then desalted and concentrated (4000 rev min^{-1}) using Centricon Plus centrifugal filter devices (Millipore). Finally, the purified protein was eluted from the third CM Sepharose fast flow column equilibrated with 10 m*M* sodium phosphate buffer pH 7.4 at a flow rate of 1 ml min^{-1} using an NaCl gradient (0–0.5 *M*) and collected with a volume of 10 ml per tube. The natrin protein were desalted and concentrated (4000 rev min^{-1}) using Centricon Plus centrifugal filter devices. The yield of the protein was approximately 5 mg and the purity was better than 95% as analysized by 10% SDS–PAGE (Fig. 1) and UV–visible spectra.

2.2. Crystallization

The protein sample was crystallized using the hanging-drop vapour-diffusion method at 291 K. The initial crystallization condition containing 20%(*w*/*v*) PEG 6000 as the major precipitant was obtained using Crystal Screens I and II from Hampton Research (Jancarik & Kim, 1991) and Wizard Screen from BioStructure Co. This condition was further refined to produce larger and better crystals. 2 µl drops of purified protein (∼8 mg ml^{-1}) in 0.1 *M* Tris buffer pH 7.4 were mixed with equal volumes of reservoir solution containing 20%(*w*/*v*) PEG 6000 and 60 m*M* ammonium citrate in 0.1 *M* HEPES buffer pH 7.0 and equilibrated against 0.5 ml reservoir solution in a 24-well ADX plate (Hampton Research). Crystals appeared after two weeks and reached maximum dimensions in one month. The diffraction-quality crystals were used for X-ray data collection (Fig. 2).

2.3. X-ray data collection and processing

The crystal was transferred into a cryoprotectant solution containing 20%(*v*/*v*) glycerol, mounted on a 0.1–0.2 mm glass-fibre loop (Hampton Research) and flash-cooled in liquid nitrogen at 100 K. X-ray diffraction data collection from frozen crystals was carried out using a synchrotron-radiation X-ray source at beamline BL12B2 equipped with a Quantum 4R CCD (ADSC) detector at SPring-8 in Japan. For complete high-resolution data collection, two 360° sets of data were measured with 1° oscillations using an X-ray wavelength of 1.00 Å at 110 K in a nitrogen stream produced by an X-Stream cryosystem (Rigaku/MSC). The first rotation measured the medium- to high-resolution diffraction (8–1.58 Å) with an exposure time of 40 s and a crystal-to-detector distance of 250 mm. The CCD detector was offset 20° in order to avoid diffraction-spot overlaps arising from the long *c* axis (243 Å). Some of the low-

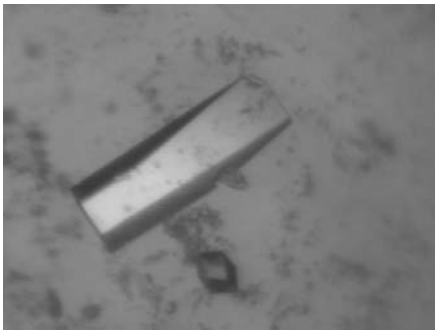

Figure 2
Single crystals of natrin grown by the hanging-drop method.

resolution (>8 Å) reflections saturated the detector electronics during the recording at 1.58 Å and were not measured. The second rotation was to measure the low- to medium-resolution diffraction (25–2.0 Å) with a crystal-to-detector distance of 180 mm, a detector offset angle of 0° and the shorter exposure time of 20 s, which allowed the measurement of the few reflections that remained saturated in the 8–2 Å resolution range of the first data set (Fig. 3). Reflections that were duplicated between the first and second data set were merged for later use.

Taking advantage of the large number of cysteines (16 S atoms) and the three methionines (three S atoms) in natrin, the sulfur anomalous dispersion signal was collected for use in SAD phasing to determine the protein structure at beamline BL17B2 equipped with a Q210 CCD detector (ADSC) at the National Synchrotron Radiation Research Center (NSRRC) in Taiwan. Thus, another 360° data set (1° degree oscillation frames, each exposed for 40 s) at 2.6 Å resolution was collected with a crystal-to-detector distance of 120 mm at 110 K using the long wavelength of 1.743 Å. The data sets were indexed, integrated, scaled and merged using the programs *DENZO* and *HKL*2000 (Otwinowski & Minor, 1997).

3. Results and discussion

Analysis of the diffraction pattern indicated that the crystals exhibited orthorhombic symmetry and systematic absences suggested that the space group was $C222_1$. The merged data set was 95.3% complete with an internal agreement (R_{sym}) of 7.4% for the resolution range 25.0–1.58 Å. The locked self-rotation function of crystals using data in the 15–4 Å resolution range and the program *GLRF* (Tong & Rossmann, 1990) showed a twofold non-crystallographic symmetry (NCS) axis at the spherical polar angle $\varphi = 0$, $\psi = 43.5$, $\kappa = 180°$ (Fig. 4). Assuming the presence of two molecules in the asymmetric unit, the Matthews coefficient is estimated to be 2.35 Å^3 Da^{-1}, corresponding to a solvent content of 47.60% (Matthews, 1968), which is within the normal range for protein crystals. Details of the data statistics are given in Table 1.

Owing to the presence of 16 cysteine and three methionine residues in natrin, initial

Table 1
Crystal diffraction statistics of natrin.

Wavelength (Å)	1.00 (BL12B2, SPring-8)	1.743 (BL17B2, NSRRC)
Temperature (K)	110	110
Resolution range (Å)	25.0–1.58	30.0–2.60
Space group	$C222_1$	$C222_1$
Unique reflections	61378	14344
Completeness (%)	95.6 (83.7)†	95.8 (92.2)‡
$\langle I/\sigma(I)\rangle$	17.6 (2.8)†	22.1 (8.0)‡
Average redundancy	6.86	10.38
R_{sym}§ (%)	7.3 (21.2)†	5.6 (13.8)‡
Mosaicity (°)	0.481	0.473
Unit-cell parameters (Å)		
a	59.172	59.132
b	65.038	64.973
c	243.156	242.921
No. molecules per AU	2	2

† Values in parentheses are for the highest resolution shell (1.64–1.58 Å). ‡ Values in parentheses are for the highest resolution shell (2.69–2.60 Å). § $R_{sym} = \sum_h \sum_i [|I_i(h) - \langle I(h)\rangle|/\sum_h \sum_i I_i(h)]$, where I_i is the ith measurement and $\langle I(h)\rangle$ is the weighted mean of all measurements of $I(h)$.

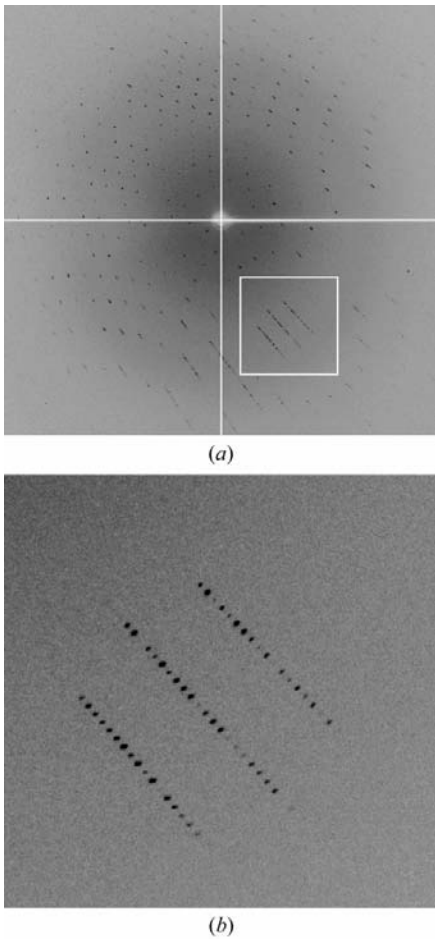

Figure 3
X-ray diffraction pattern of the data collection without detector offset. The box shows the close diffraction spots arising from the long *c* axis of about 243 Å.

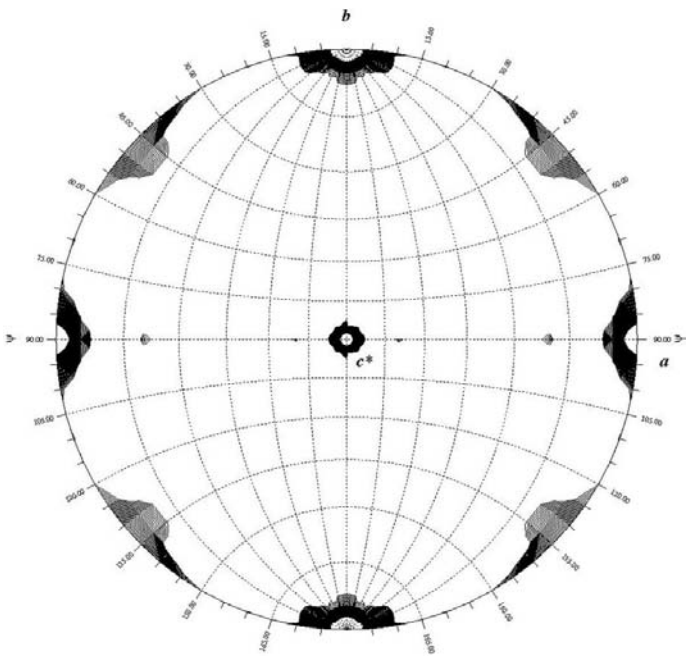

Figure 4
Projection of the $\kappa = 180°$ section of the self-rotation function of crystals calculated with *GLRF* using 15–4 Å resolution data, showing the presence of the non-crystallographic twofold axis at polar angles $\varphi = 0$, $\psi = 43.5$, $\kappa = 180°$.

attempts to solve the crystal structure were performed using the sulfur single-wavelength anomalous scattering method, with additional data collected to 2.6 Å resolution at the long wavelength of 1.743 Å (data statistics shown in Table 1). The Harker sections of the anomalous difference Patterson map show clear solutions for S atoms that can be used in phase calculation using the sulfur-SAD method. The sulfur sites were determined and input to the program *SOLVE*2.06 (Terwilliger, 1999) and gave an initial figure of merit (FOM) of 0.37 to 3.0 Å resolution. Completion of the structure determination is currently in progress and will be described in a separate paper. The structure and its classification will not only provide a step towards understanding the clear physiological role of natrin in venoms, but will also contribute to the understanding of the biological functions of CRISP-family proteins with highly conserved cysteine residues in other species, including humans.

We are grateful to our colleagues Yu-Cheng Jean, Yu-San Huang and Chun-Shiun Chao for valuable technical assistance and discussions during synchrotron-radiation X-ray facility data collection at BL17B2 of NSRRC, Taiwan and Mau-Tsu Tang and Jey-Jau Lee at BL12B2 of SPring-8, Japan. This study was supported in part by the National Synchrotron Radiation Research Center grants 924RSB02, 934RSB02 and National Science Council grant NSC-92-2321-B-213-001, Taiwan to C-JC and NSC-92-2311-B-007-014 to WW.

References

Arni, R. K., Padmanabhan, K. P. & Tulinsky, A. (1999). *Acta Cryst.* D**55**, 1468–1470.

Brown, R. L., Haley, T. L., West, K. A. & Crabb, J. W. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 754–759.

Brown, R. L., Lynch, L. L., Haley, T. L. & Arsanjani, R. (2003). *J. Gen. Physiol.* **122**, 749–760.

Busso, D., Cohen, D. J., Da Ros, V., Fissore, R. & Cuasnicu, P. S. (2003). *Cell Mol. Biol.* **49**, 407–412.

Chang, T. Y., Mao, S. H. & Guo, Y. W. (1997). *Toxicon*, **35**, 879–888.

Cohen, D. J., Ellerman, D. A. & Cuasnicu, P. S. (2000). *Biol. Reprod. Dev.* **63**, 462–468.

Eberspaecher, U., Roosterman, D., Kratzschmar, J., Haendler, B., Habenicht, U. F., Becker, A., Quensel, C., Petri, T., Schleuning, W. D. & Donner, P. (1995). *Mol. Reprod. Dev.* **42**, 157–172.

Foster, J. A. & Gerton, G. L. (1996). *Mol. Reprod. Dev.* **44**, 221–229.

Fry, B. G. & Wuster, W. (2004). *Mol. Biol. Evol.* **21**, 870–883.

Fujii, Y., Okuda, D., Fujimoto, Z., Horii, K., Morita, T. & Mizuno, H. (2003). *J. Mol. Biol.* **332**, 1115–1122.

Fujii, Y., Okuda, D., Fujimoto, Z., Morita, T. & Mizuno, H. (2002). *Acta Cryst.* D**58**, 145–147.

Horii, K., Okuda, D., Morita, T. & Mizuno, H. (2003). *Biochemistry*, **42**, 12497–12502.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Jin, Y., Lu, Q., Zhou, X., Zhu, S., Li, R., Wang, W. & Xiong, Y. (2003). *Toxicon*, **42**, 539–547.

Kjeldsen, L., Cowland, J. B., Johnsen, A. H. & Borregaard, N. (1996). *FEBS Lett.* **380**, 246–250.

Kosari, F., Asmann, Y. W., Cheville, J. C. & Vasmatzis, G. (2002). *Cancer Epidemiol. Biomarkers Prev.* **11**, 1419–1426.

Kratzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P. & Schleuning, W. D. (1996). *Eur. J. Biochem.* **236**, 827–836.

Maeda, T., Nishida, J. & Nakanishi, Y. (1999). *Dev. Growth Differ.* **41**, 715–722.

Maeda, T., Sakashita, M., Ohba, Y. & Nakanishi, Y. (1998). *Biochem. Biophys. Res. Commun.* **248**, 140–146.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mochca-Morales, J., Martin, B. M. & Possani, L. D. (1990). *Toxicon*, **28**, 299–309.

Moiseeva, N., Swenson, S. D., Markland, F. S. Jr & Bau, R. (2002). *Acta Cryst.* D**58**, 2122–2124.

Morrissette, J., Kratzschmar, J., Haendler, B., El-Hayek, R., Mochca-Morales, J., Martin, B. M., Patel, J. R., Moss, R. L., Schleuning, W. D. & Coronado, R. (1995). *Biophys. J.* **68**, 2280–2288.

Nobile, M., Magnelli, V., Lagostena, L., Mochca-Morales, J., Possani, L. D. & Prestipino, G. (1994). *J. Membr. Biol.* **139**, 49–55.

Nobile, M., Noceti, F., Prestipino, G. & Possani, L. D. (1996). *Exp. Brain Res.* **110**, 15–20.

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

Pfisterer, P., Konig, H., Hess, J., Lipowsky, G., Haendler, B., Schleuning, W. D. & Wirth, T. (1996). *Mol. Cell Biol.* **16**, 6160–6168.

Schwidetzky, U., Haendler, B. & Schleuning, W. D. (1995). *Biochem. J.* **309**, 831–836.

Sivashanmugam, P., Richardson, R. T., Hall, S., Hamil, K. G., French, F. S. & O'Rand, M. G. (1999). *J. Androl.* **20**, 384–393.

Terwilliger, T. C. (1999). *Acta Cryst.* D**55**, 1863–1871.

Tomar, S., Yadav, S., Chandra, V., Kumar, P. & Singh, T. P. (2001). *Acta Cryst.* D**57**, 1669–1670.

Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.

Udby, L., Calafat, J., Sorensen, O. E., Borregaard, N. & Kjeldsen, L. (2002). *J. Leukoc. Biol.* **72**, 462–469.

Yamazaki, Y., Brown, R. L. & Morita, T. (2002). *Biochemistry*, **41**, 11331–11337.

Yamazaki, Y., Hyodo, F. & Morita, T. (2003). *Arch. Biochem. Biophys.* **412**, 133–141.

Yamazaki, Y., Koike, H., Sugiyama, Y., Motoyoshi, K., Wada, T., Hishinuma, S., Mita, M. & Morita, T. (2002). *Eur. J. Biochem.* **269**, 2708–2715.