Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, partial characterization, crystallization and preliminary X-ray diffraction of two cysteine-rich secretory proteins from *Naja atra* and *Trimeresurus stejnegeri* venoms

Jing Wang,[a,b]‡ Min Guo,[a,b]‡ Xiongying Tu,[a,b] Dinghai Zheng,[a,b] Maikun Teng,[a,b] Liwen Niu,[a,b]* Qun Liu,[c] Qingqiu Huang[c] and Quan Hao[c]

[a]Key Laboratory of Structural Biology, Chinese Academy of Sciences, 96 Jinzhai Road, Hefei, Anhui 230026, People's Republic of China, [b]Departments of Molecular and Cell Biology, School of Life Sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230026, People's Republic of China, and [c]MacCHESS, Cornell High Energy Synchrotron Source, Cornell University, NY 14853, USA

‡ These authors made equivalent contributions.

Correspondence e-mail: lwniu@ustc.edu.cn

Received 1 December 2003
Accepted 10 March 2004

Cysteine-rich secretory proteins (CRISPs) are widely distributed in mammals and snake venoms. They possess apparent homology but varying functions. The structure of CRISPs has remained elusive. Two novel members of the family, natrin and stecrisp, have been purified from *Naja atra* and *Trimeresurus stejnegeri* venoms, respectively. Their crystals diffract X-rays to resolution limits of 2.1 and 1.6 Å, repectively, and belong to the orthorhombic system with different space groups, unit-cell parameters and numbers of molecules per asymmetric unit. Their structures will contribute a structural basis for further functional studies of this family.

1. Introduction

Cysteine-rich secretory proteins (CRISPs) belong to the pathogenesis-related group-1 protein (PR-1) superfamily and have a common N-terminal domain, the structure of which has been solved (p14a; Niderman *et al.*, 1995; Ves v 5; Lu *et al.*, 1993). All CRISP-family proteins possess obvious conservation throughout the whole length of their sequence, with 16 strictly conserved cysteines, all of which are engaged in disulfide bonds (Eberspaecher *et al.*, 1995) and ten of which are clustered in the additional C-terminal third of the proteins.

CRISPs are generally expressed in mammalian tissues and can be classified into three subclasses according to their sequence homology and tissue specificity: CRISP-1 in the epididymis, CRISP-2 in spermatocytes and CRISP-3 in the salivary gland, pancreas and prostate (Kratzschmar *et al.*, 1996). CRISP-1 and CRISP-2 are associated with sperm maturation and the fertilization process (Giese *et al.*, 2002; Ellerman *et al.*, 2002; Maeda *et al.*, 1998). CRISP-3 has been found to be a defence-associated molecule and possesses a high expression level in prostate-cancer tissue (Kosari *et al.*, 2002; Udby *et al.*, 2002; Liao *et al.*, 2003). A wide distribution of CRISP-family proteins in snake venoms from several species has been documented (Yamazaki *et al.*, 2003) and many of them have been shown to have characteristic functions. Ablomin, latisemin, ophanin, triflin and piscivorin can inhibit the contraction of smooth muscle induced by high concentrations of potassium ion, whereas tigrin and catrin-1 cannot (Yamazaki, Koike *et al.*, 2002; Yamazaki *et al.*, 2003). Cyclic nucleotide-gated ion channels can be inhibited by PsTx and pseudecin, but not by triflin, latisemin, piscivorin and tigrin (Brown *et al.*, 1995, 1999; Zagotta & Siegelbaum, 1996; Yamazaki *et al.*, 2003). Helothermine (HLTx), purified from the salivary gland of *Heloderma horridum* (a Mexican beaded lizard; Nobile *et al.*, 1994), can even block voltage-gated calcium channels, potassium channels and ryanodine receptors (Morrissette *et al.*, 1995; Nobile *et al.*, 1996). Compared with most other members of the PR-1 superfamily that have a single N-terminal domain, such as the plant pathogenesis-related proteins, PR-1 proteins in yeast and fungi and insect-venom allergen proteins (Kitajima & Sato, 1999), CRISPs possess a special distribution in vertebrates, conserved additional C-terminal cysteine-rich domains and distinct physiological functions, indicating the necessity of obtaining structural information on the CRISP family in order to resolve their molecular mechanism. However, no three-dimensional structures of CRISPs from either mammalian or other species have yet been experimentally determined. The purpose of this paper is to report the purification, partial characterization, crystallization and preliminary X-ray diffraction of natrin and stecrisp, two novel cysteine-rich secretory proteins from *Naja atra* and *Trimeresurus stejnegeri* venoms, respectively.

2. Materials and methods

2.1. Materials

Dried crude *N. atra* and *T. stejnegeri* venoms were obtained from the Tunxi Snakebite Institute (Anhui, China). High-performance Q-Sepharose, high-performance phenyl-

Sepharose, high-performance SP-Sepharose, CM-Sepharose and Superdex 75 HR 10/30 columns were purchased from Pharmacia (Uppsala, Sweden). Other reagents and chemicals were of analytical grade from commercial sources.

2.2. Sample purification and partial characterization

The isolation and purification of natrin and stecrisp were carried out by two-step chromatography at room temperature (~293 K, see Fig. 1) and by four-step chromatography at ~277 K (see Fig. 2), respectively.

SDS–PAGE was performed according to the Mini Protein II dual-gel method (Bio-Rad); the gel was stained with Coomassie Blue G-250. The mass spectra of the protein samples were acquired by MALDI–TOF mass spectrometry (Bruker Co.). Amino-acid sequence determinations were performed by Edman degradation using Applied Biosystems protein sequencers (476A and 496A).

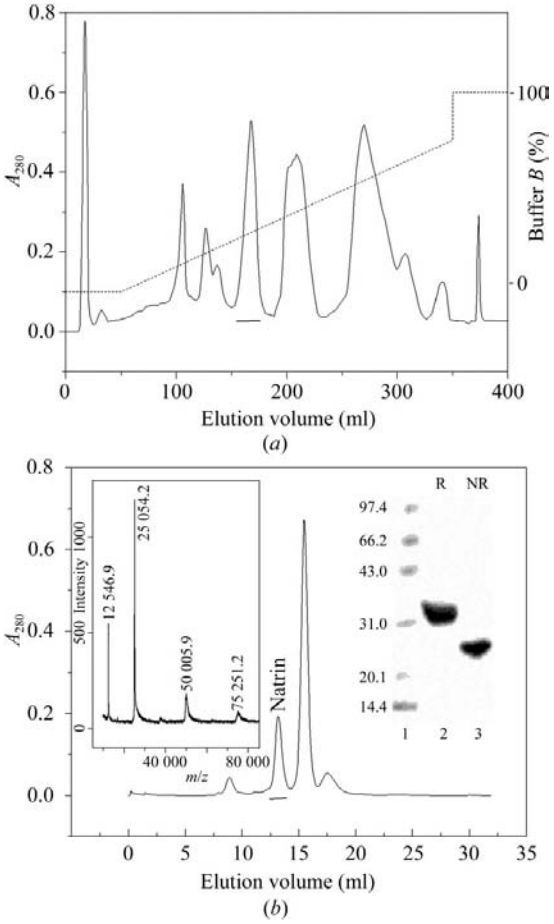

Figure 1
Purification of natrin from *N. atra* venom. (*a*) 500 mg of dried crude venom was dissolved in 10 ml buffer *A* (0.02 *M* NaH_2PO_4/Na_2HPO_4 pH 6.0) and centrifuged at 10 000 rev min^{-1} for 20 min to remove insoluble materials. Using a CM-Sepharose column (1.6 × 20 cm), elution was performed at a flow rate of 0.5 ml min^{-1} with a combined pH gradient (0.02 *M* NaH_2PO_4/Na_2HPO_4 from pH 6.0 to 8.0) and salt-concentration gradient (0–0.5 *M* NaCl, indicated by a dotted line); the elution gradient was produced by mixing 200 ml of buffer *A* with 200 ml of buffer *B* (0.02 *M* NaH_2PO_4/Na_2HPO_4 pH 8.0 containing 0.5 *M* NaCl). The fractions containing natrin (indicated by a solid bar) were pooled and desalted for further isolation. (*b*) The fractions from the previous column were applied onto a Superdex 75 column (2 × 30 cm) and eluted with buffer (0.05 *M* Tris–HCl pH 8.0 containing 0.15 *M* NaCl) at a flow rate of 60.0 ml h^{-1}. The fractions containing natrin (indicated by a solid bar) were pooled, desalted and lyophilized. Left inset: mass spectrum of natrin. Right inset: the homogeneity of natrin was also demonstrated by SDS–PAGE under reducing (lane 2) and non-reducing (lane 3) conditions; lane 1 contains protein standards (in kDa) for molecular-weight estimation.

2.3. Crystallization and preliminary X-ray diffraction analysis

Lyophilized natrin was dissolved in double-distilled water, adjusted to a concentration of 25 mg ml^{-1} and then centrifuged to remove insoluble components. The crystallization was performed at room temperature using conventional hanging-drop vapour diffusion (McPherson, 1982). 2 µl protein solution was mixed with 2 µl precipitant solution (0.1 *M* Tris–HCl pH 8.9 containing 2.5 *M* NaCl and 10% ethanol) and then equilibrated against precipitant solution. Crystals with dimensions 0.15 × 0.1 × 0.1 mm were obtained one week later (Fig. 3*a*).

Purified stecrisp was concentrated to 5 mg ml^{-1} in an Amicon stirred-cell concentrator. The crystallization of stecrisp was also carried out at room temperature using hanging-drop vapour diffusion. 2 µl protein solution was added to 1 µl precipitant solution (44% PEG 4000, 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 7.75) and then equilibrated against precipitant solution. Crystals grew to dimensions of 0.2 × 0.05 × 0.05 mm within two weeks (Fig. 3*b*).

For data collection, the natrin crystal was soaked for about 10 s in cryoprotection solution (10% ethanol, 2.0 *M* NaCl, 20% glycerol, 0.1 *M* Tris–HCl pH 8.9); the stecrisp crystal was soaked for more than one week in a solution consisting of 60% PEG 4000, 0.2 *M* sodium acetate, 0.1 *M* Tris–HCl pH 7.75. X-ray diffraction data were then collected at MacCHESS beamline A1 (Cornell University, USA) with an incident X-ray wavelength of 0.9363 Å at 100 K using an ADSC Q210 CCD detector. The crystal-to-detector distance, oscillation angle and exposure time for each image frame were set to 180 mm, 1° and 10 s, respectively. The diffraction data were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

Natrin and stecrisp, two homologous proteins purified from *N. atra* and *T. stejnegeri* venoms, respectively (see Figs. 1 and 2), show an apparently single band on SDS–PAGE (about 25 and 24 kDa under non-reducing conditions, with higher bands at 32 and 28 kDa under reducing conditions, respectively, supporting the existence of disulfide bonds in the two proteins; Eberspaecher *et al.*, 1995). Both of the mature proteins consist of 221 amino-acid residues and belong to the CRISP family, as indicated by the N-terminal residue sequences of natrin (NVDFNSESTRRKKKQKEIVDLHNSLRRRVS), stecrisp (NVDFDSESPRKPEIQNEIVDLHNSLRRSVNP) and two trypsin-cleaved fragments of stecrisp (TATPYTSGTPCGDCPSDCDNGLCTNPCTR and CIESHSSAESR), as well as their cDNA sequences (Genbank accession Nos. AY324325 and AY423708, respectively). The average molecular weights of natrin (25 054 Da; Fig. 1*b*) and stecrisp (25 135 Da; Fig. 2*c*) determined by mass spectra are in accord with the calculated molecular weights of 24 954 Da (natrin) and 25 129 Da (stecrisp), implying that there was no further post-translational modification. On alignment (Fig. 4), the sequences of natrin and stecrisp are found to be homologous with those of other cysteine-rich secretory proteins, such as human testis-specific protein (TPX-1; Kasahara *et al.*, 1989; Kratzschmar *et al.*, 1996), elapid-venom pseudechetoxin (PsTx; Yamazaki, Brown *et al.*, 2002) and helothermine (HLTX; Morrissette *et al.*, 1995). Recently, a further three CRISP proteins from *T. jerdonii* and *N. atra* venoms (Jin *et al.*, 2003) with 90–60% identity to stecrisp and natrin have also been purified, indicating the obvious homogeneity in the CRISP family, especially among the snake-venom CRISPs. The sequences of natrin and stecrisp also resemble those of two structurally known proteins, p14a (Niderman *et al.*, 1995) and Ves v 5 (Lu *et al.*,

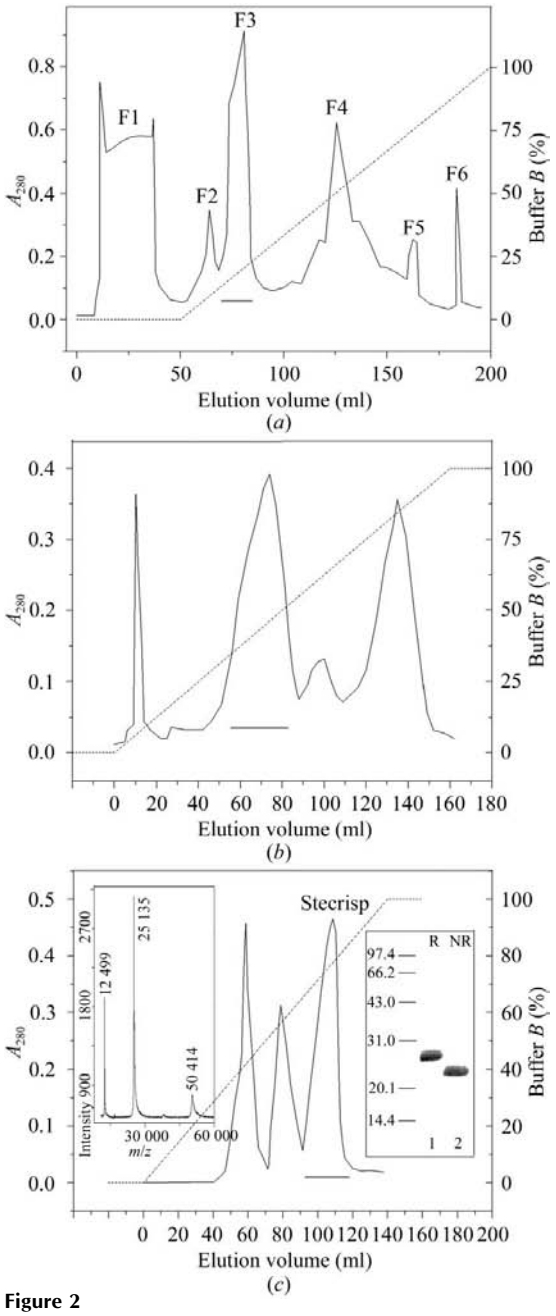

Figure 2
Purification profile of stecrisp from *T. stejnegeri* venom. (*a*) 300 mg of dried crude venom was applied onto a high-performance Q-Sepharose column (1.6 × 15 cm) and eluted at a flow rate of 60.0 ml h^{-1} with a combined pH gradient (0.02 *M* Tris–HCl from pH 9.4 to 7.8) and salt-concentration gradient (0–0.32 *M* NaCl, indicated by dotted line). Fraction 3 (indicated by a bar) was collected, concentrated and precipitated with 1.0 *M* ammonium sulfate. The supernatant was applied to a high-performance phenyl-Sepharose column (1.6 × 15 cm); the desired fraction was eluted with a stepwise gradient of 0.7–0.4 *M* ammonium sulfate in 0.02 *M* Tris–HCl pH 8.0. (*b*) The fraction from the previous step was loaded onto a high-performance SP-Sepharose column (1.6 × 15 cm) and eluted with a gradient of NaCl from 0.0 to 0.2 *M* (represented by a dashed line) in 0.04 *M* acetic acid/sodium acetate pH 4.0. The second peak (indicated by a bar) was collected. (*c*) The sample was applied onto another high-performance Q-Sepharose column and eluted with a NaCl gradient from 0.03 to 0.15 *M* (dotted line) in 0.02 *M* Tris–HCl pH 8.7. The fraction containing stecrisp (indicated by a bar) was pooled for further research. Left inset: mass spectrum of stecrisp. Right inset: the homogeneity of stecrisp was detected by SDS–PAGE under reducing (lane 1) and non-reducing (lane 2) conditions.

Table 1
Statistics of data collection and reduction.

Values in parentheses are for the highest resolution shell.

Protein	Natrin	Stecrisp
Space group	$P2_12_12_1$	$I222$ or $I2_12_12_1$
Unit-cell parameters		
a (Å)	101.68	61.38
b (Å)	90.75	78.42
c (Å)	91.55	106.61
Estimated No. molecules per AU	3	1
V_M (Å^3 Da^{-1})	2.82	2.04
Resolution limits (Å)	50–2.10 (2.14–2.10)	33–1.60 (1.63–1.60)
$I/\sigma(I)$	47.1 (9.0)	29.2 (3.0)
No. observations	345304	183333
No. independent reflections	50103	34038
R_{merge}† (%)	7.5 (38.3)	8.3 (34.8)
Completeness‡ (%)	99.9 (99.9)	99.2 (93.5)

† $R_{merge} = \sum_h \sum_j |I(h)_j - \langle I(h)\rangle| / \sum_h \sum_i I(h)_j$, where $I(h)_j$ is the observed intensity of reflection and $\langle I(h)\rangle$ is the mean intensity of reflection *h*. ‡ The ratio of the number of reflections to the number of possible reflections.

1993), in their N-terminal domains. They represent a compact $\alpha\beta\alpha$-sandwich core fold with high thermal stability, which is suggested to be conserved in the PR-1 superfamily from the sequence homology (Fernandez *et al.*, 1997; Henriksen *et al.*, 2001). All of the N-terminal domains share a common core stretch (HF/YTQI/V/MVW) which has been postulated to be the 'folding nucleus' (Fernandez *et al.*, 1997). The alignment also shows the conservation of four externally exposed amino acids (His60, Glu74, Glu96 and His115 numbering from the N-terminus of the mature proteins), which are believed to be important in recognition and binding in the possible active site (Szyperski *et al.*, 1998). In addition to the N-terminal portion, natrin and stecrisp are also similar in their C-terminal motif, which is a characteristic of the CRISP family. The C-terminal CRISP motif is homologous along an approximately 60-amino-acid segment and contains ten strictly conserved cysteine residues, which may imply a relatively conserved and pivotal structural unit for the family (Eberspaecher *et al.*, 1995; Orengo *et al.*, 1994). Several residues clustered at the beginning of this motif are also of relatively high sequence similarity (Fig. 4). Compared with the single N-terminal domain proteins, such as plant PR-1 proteins, venom allergen 5 (insect venom) and GliPR (human brain, with C-terminal transmembrane segment), the members of the CRISP family display a much greater range of physiological functions and a particular histological distribution. These charactistics of the CRISP family suggest that their important biological roles are related to the C-terminal CRISP motif or to the interaction of the two domains. The structures of natrin and stecrisp will offer templates for modelling other CRISPs and will provide a molecular basis for understanding the physiological functions of this family.

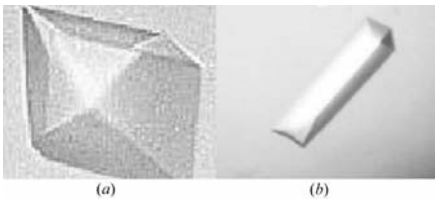

Figure 3
Crystal photomicrographs of natrin (*a*) and stecrisp (*b*).

The crystals of natrin and stecrisp (Figs. 3*a* and 3*b*) diffracted X-rays to resolution limits of 2.1 and 1.6 Å, respectively. They belong to the orthorhombic system with different space groups, unit-cell parameters and numbers of molecules per asymmetric unit (Table 1). The structural determinations of natrin and stecrisp are being carried out in our laboratory.

Financial support for this project to LN and MT was provided by research grants from the National Natural Science Foundation of China (Grant Nos. 39870108, 30025012 and 30121001), the 973 Plan of the Ministry of Science and Technology of China (Grant No. G1999075603) and the Chinese Academy of Sciences (Grant Nos. STZ-2-07, STZ98-2-12 and STZ01-29).

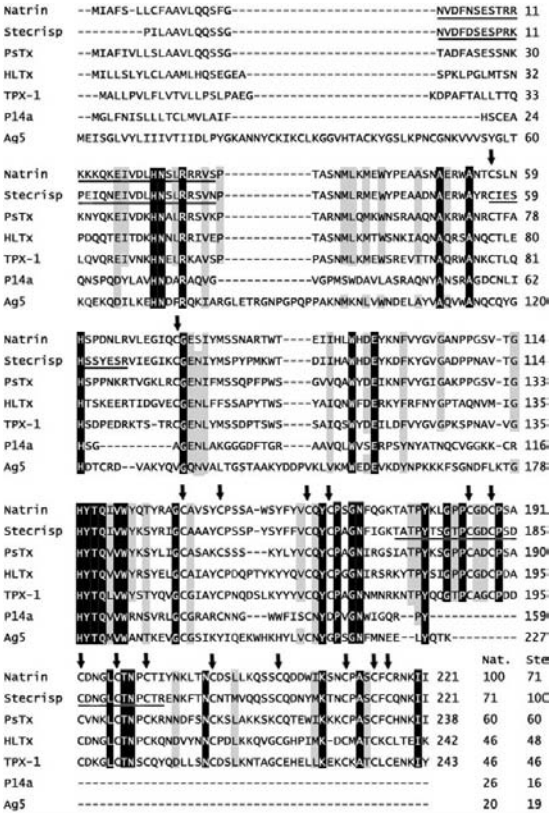

Figure 4
Sequence alignment of natrin and stecrisp with relatives from the CRISP family. Residues from direct sequencing of native proteins and trypsin-digested *S*-pyridylethylated peptides are underlined. The abbreviations are: PsTx, pseudechetoxin; TPX-1, testis-specific protein; HLTx, helothermine; Ag5, yellow jacket venom allergen 5; P14a, tomato pathogenesis-related protein P14a. GenBank accession Nos.: natrin, AY324325; opharin, AY181984.1; PsTx, AY072695; TPX-1, P16562; HLTX, Q91055; Ag5, Q05110; P14a, P04284. Identical amino-acid residues are highlighted in black. Relatively conservative amino-acid exchanges are highlighted in grey. The highly conserved cysteines are indicated by arrows. The relative degrees of identity are shown at the lower right.

References

Brown, R. L., Gramling, R., Bert, R. J. & Karpen, J. W. (1995). *Biochemistry*, **34**, 8365–8370.

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

Ellerman, D. A., Da Ros, V. G., Cohen, D. J., Busso, D., Morgenfeld, M. M. & Cuasnicu, P. S. (2002). *Biol. Reprod.* **67**, 1225–1231.

Eberspaecher, U., Roosterman, J., 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.

Fernandez, C., Szyperski, T., Bruyere, T., Ramage, P., Mosinger, E. & Wuthrich, K. (1997). *J. Mol. Biol.* **266**, 576–593.

Giese, A., Jude, R., Kuiper, H., Raudsepp, T., Piumi, F., Schambony, A., Guerin, G., Chowdhary, B. P., Distl, O., Topfer-Petersen, E. & Leeb, T. (2002). *Gene*, **299**, 101–109.

Henriksen, A., King, T. P., Mirza, O., Monsalve, R. I., Meno, K., Ipsen, H., Larsen, J. N., Gajhede, M. & Spangfort, M. D. (2001). *Proteins*, **45**, 438–448.

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

Kasahara, M., Gutknecht, J., Brew, K., Spurr, N. & Goodfellow, P. N. (1989). *Genomics*, **5**, 527–534.

Kitajima, S. & Sato, F. (1999). *J. Biochem.* **125**, 1–8.

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.

Liao, Q., Kleeff, J., Xiao, Y., Guweidhi, A., Schambony, A., Topfer-Petersen, E., Zimmermann, A., Buchler, M. W. & Friess, H. (2003). *Histol. Histopathol.* **18**, 425–433.

Lu, G., Villalba, M., Coscia, M. R., Hoffman, D. R. & King, T. P. (1993). *J. Immunol.* **150**, 2823–2830.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. New York: Wiley.

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

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. & Possani, L. D. (1995). *Biophys. J.* **68**, 2280–2288.

Niderman, T., Genetet, I., Bruyere, T., Gees, R., Stintzi, A., Legrand, M., Fritig, B. & Mosinger, E. (1995). *Plant Physiol.* **108**, 17–27.

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.

Orengo, C. A., Jones, D. T. & Thornton, J. M. (1994). *Nature (London)*, **372**, 631–634.

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

Szyperski, T., Fernandez, C., Mumenthaler, C. & Wuthrich, K. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2262–2266.

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.

Zagotta, W. N. & Siegelbaum, S. A. (1996). *Annu. Rev. Neurosci.* **19**, 235–263.