Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of the magnesium transporter CorA

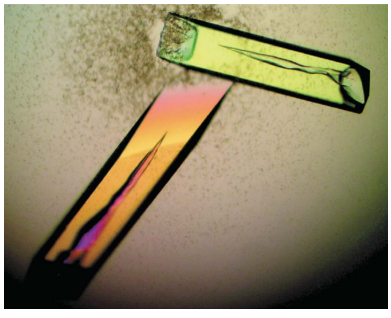
Jian Payandeh[a,b]*‡ and Emil F. Pai[a,b,c]*‡

[a]Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 2M9, Canada, [b]Division of Cancer Genomics and Proteomics, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario M5G 2M9, Canada, and [c]Departments of Biochemistry and Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5G 2M9, Canada

‡ Present address: Division of Cancer Genomics and Proteomics, Ontario Cancer Institute, MaRS Centre, Toronto Medical Discovery Tower, 101 College Street, Toronto, Ontario M5G 1L7, Canada.

Correspondence e-mail: payandeh@uhnres.utoronto.ca, pai@hera.med.utoronto.ca

Received 8 December 2005
Accepted 10 January 2006



The full-length integral membrane protein CorA from *Thermotoga maritima* ($TmCorA_{1–351}$) has been expressed in *Escherichia coli* and purified without membrane isolation. $TmCorA_{1–351}$ crystallized in the monoclinic space group $C2$, with unit-cell parameters a = 214.25, b = 86.30, c = 181.53 Å, β = 112.23°. Native crystals diffracted to 3.7 Å using synchrotron radiation, but selenomethionine-substituted crystals rarely diffracted to better than 5.0 Å. All full-length protein crystals were highly mosaic and produced anisotropic diffraction patterns. To aid in crystallographic phasing, soluble domain constructs were screened and the periplasmic domain of CorA from *Archaeoglobus fulgidus* ($AfCorA_{1–263}$) was crystallized in the hexagonal space group $P6_122$, with unit-cell parameters a = b = 101.17, c = 142.87 Å. Native and SeMet-substituted $AfCorA_{1–263}$ crystals diffracted to ∼3.0 Å using synchrotron radiation.

1. Introduction

At least three distinct types of transport systems mediate Mg^{2+} homeostasis across the bacterial cell membrane (Kehres & Maguire, 2002). The nearly ubiquitous CorA is an integral membrane protein that is constitutively expressed and represents the major pathway for magnesium import (Kehres & Maguire, 2002; Gardner, 2003). In fact, CorA is essential for the viability of the gastric pathogen *Helicobacter pylori* (Pfeiffer *et al.*, 2002). CorA proteins are also found in *Archaea*, where the functional homology has been directly established (Smith *et al.*, 1998). Despite overall weak sequence similarity, some eukaryotic homologues are also known to be involved in magnesium transport and can functionally substitute for each other or their bacterial counterparts (Gardner, 2003). These proteins share a universally conserved GMN tripeptide motif near their predicted transmembrane-spanning regions (Gardner, 2003; Knoop *et al.*, 2005). A distant CorA homologue from *Salmonella typhimurium* (named ZntB) is unable to transport Mg^{2+}, but has instead been characterized as a zinc transporter (Worlock & Smith, 2002). Although mutations in the CorA GMN motif are known to abolish Mg^{2+} transport, it is thought that naturally occurring variants of this motif may be associated with the transport of other divalent cations (*e.g.* Zn^{2+} and Cd^{2+}; Knoop *et al.*, 2005). The *corA* locus was named after the cobalt-resistant mutants in which it was originally identified. This phenotype is explained by the fact that CorA can also transport Co^{2+} and Ni^{2+}, although Mg^{2+} is its predominant substrate (Kehres & Maguire, 2002; Gardner, 2003; Knoop *et al.*, 2005).

Mg^{2+} is the most abundant divalent cation within cells (Maguire & Cowan, 2002). Accordingly, Mg^{2+} enjoys diverse and important roles in biology. For example, Mg^{2+} participates in and is essential to many enzymatic reactions and it is the central atom of the chlorophyll molecule (Knoop *et al.*, 2005). By bridging together neighbouring lipid, protein and sugar components through ionic interactions, Mg^{2+} also helps to provide membrane stability (Schindler & Osborn, 1979; Bara & Guiet-Bara, 1984). The Mg^{2+} ion has some unique properties. Compared with K^+, Na^+ and Ca^{2+}, Mg^{2+} has a substantially smaller ionic radius but a significantly larger hydrated radius. Hydrated Mg^{2+} is approximately 400 times larger than its dehydrated ionic form; in contrast, there is a 25-fold change for Na^+ and Ca^{2+} and only a fourfold change for K^+. Compared with the promiscuous six-, seven-,

eight- and nine-coordinate arrangements that are known for Ca^{2+}, Mg^{2+} is almost invariably hexacoordinate. Although the octahedral coordination sphere for hexacoordinate Ca^{2+} is highly flexible, for Mg^{2+} it is much more constrained (Maguire & Cowan, 2002). The high charge density and relative structural rigidity of the Mg^{2+} ion probably reflects its very distinct roles in basic biochemical processes. Accordingly, Mg^{2+} transporters are thought to constitute a unique class of integral membrane proteins (Kehres & Maguire, 2002; Gardner, 2003).

The CorA protein has been fairly well characterized biochemically. A polypeptide of ~40 kDa, CorA has a relatively large N-terminal periplasmic domain (~30 kDa) and contains no detectable signal sequence. Because the C-terminal ~80 amino acids of CorA are predicted to contain only three transmembrane-spanning α-helices, it is expected that the functional transporter would be a homo-oligomer (Kehres & Maguire, 2002; Warren *et al.*, 2004). Here, we present the crystallization and preliminary diffraction analysis of the full-length CorA from *Thermotoga maritima* and the periplasmic domain of *Archaeoglobus fulgidus*.

2. Materials and methods

2.1. Cloning, expression and purification of CorA

The full-length ($TmCorA_{1–351}$) and periplasmic domain ($TmCorA_{1–263}$) of CorA from *T. maritima* (Genebank accession No. AE001731) and the periplasmic domain of CorA ($AfCorA_{1–263}$) from *A. fulgidus* (Genebank accession No. AE001050) were cloned from genomic DNA (ATCC 43589D and 49558D, respectively) into the *Nde*I and *Bam*HI sites of the pET-15b vector (Novagen). Native proteins were expressed in *Escherichia coli* BL21-CodonPlus(DE3)-RIL (Stratagene) grown in LB media containing ampicillin (100 µg ml^{-1}) at 310 K with shaking at 200 rev min^{-1}; cells were induced at an OD_{600} of ~0.8 with 0.1 m*M* IPTG for 8–12 h and the final OD_{600} was ~2.3. Selenomethionyl (SeMet) protein was expressed similarly, but in *E. coli* B834(DE3) and in M9 minimal media supplemented with 40 mg l^{-1} SeMet. After harvesting, cells from 2 l cultures were resuspended in 80 ml buffer *A* (50 m*M* Tris pH 8.0, 100 m*M* NaCl, 5 m*M* imidazole, 10 m*M* β-mercaptoethanol) supplemented with 16 µg ml^{-1} benzamidine, 10 µg ml^{-1} aprotinin, 10 µg ml^{-1} leupeptin, 1 m*M* phenylmethylsulfonyl fluoride and 7 µg ml^{-1} DNAse. Cells were lysed by three passages through an Emulsiflex-C5 (Avestin) run at 69 MPa.

For the full-length transmembrane protein $TmCorA_{1–351}$, the crude lysate was mixed with an equal volume of solubilization buffer (buffer *A* + 2% *n*-dodecyl-β-D-maltoside; Anatrace) and gently agitated on a platform shaker for 1 h at room temperature. After centrifugation at 17 000 rev min^{-1} for 1 h, the supernatant was loaded by gravity directly onto 3 ml Ni–NTA agarose (Qiagen) pre-equilibrated in buffer *B* (buffer *A* + 0.026% *n*-dodecyl-β-D-maltoside; DDM) and washed extensively with the same buffer. The protein was eluted in buffer *B* containing 500 m*M* imidazole and the N-terminal His_6 tag was removed by thrombin digestion (Calbiochem) for ~12 h in the presence of 3 m*M* $CaCl_2$. $TmCorA_{1–351}$ was passed over a Superdex 200 column (Amersham) in buffer *B* without imidazole and peak fractions containing the protein (~250 kDa) were pooled. This sample was passed over a second Ni column and the flowthrough was collected and concentrated using a centrifugal filter device (Millipore, 30 kDa molecular-weight cutoff). The buffer was exchanged to 20 m*M* Tris pH 8.0, 100 m*M* NaCl, 0.026% DDM, 1 m*M* Tris(2-carboxyethyl)phosphine hydrochloride by three rounds of concentration and dilution.

For the periplasmic domains $TmCorA_{1–263}$ and $AfCorA_{1–263}$, purifications omitted the solubilization buffer step but otherwise followed the same scheme as described for the full-length protein, using buffer *A* instead of buffer *B*. After elution from Ni–NTA agarose and removal of the N-terminal His_6 tag, gel filtration on a Superdex 200 column resulted in peak fractions containing an apparent $TmCorA_{1–263}$ monomer (~30 kDa) or an $AfCorA_{1–263}$ oligomer (120–150 kDa). These proteins were passed over a second Ni column and the flowthrough was collected and concentrated using a centrifugal filter device (Millipore, 10 kDa molecular-weight cutoff). The buffer was exchanged to 20 m*M* Tris pH 8.0, 100 m*M* NaCl, 1 m*M* Tris(2-carboxyethyl)phosphine hydrochloride by three rounds of concentration and dilution. For the full-length and truncated CorA proteins, these procedures reproducibly yielded ~15 mg

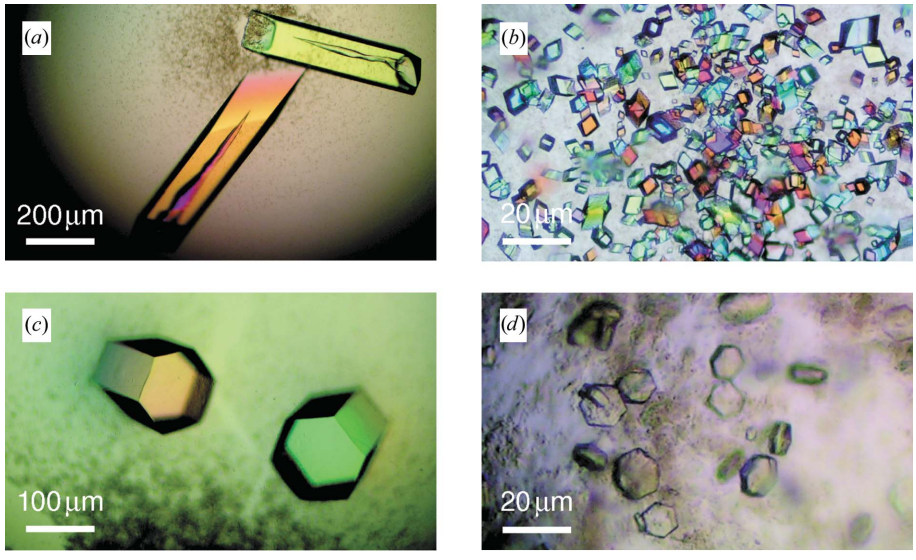


Figure 1
(*a*) Crystals of $TmCorA_{1–351}$ grown in the presence of Ca^{2+}. (*b*) Crystals of $TmCorA_{1–351}$ grown in the presence of Ni^{2+}. (*c*) Crystals of the $AfCorA_{1–263}$ periplasmic domain. (*d*) Crystals of the $TmCorA_{1–263}$ periplasmic domain.

from 2 l bacterial cultures. All proteins could be concentrated to >50 mg ml^{-1}.

2.2. CorA crystallization

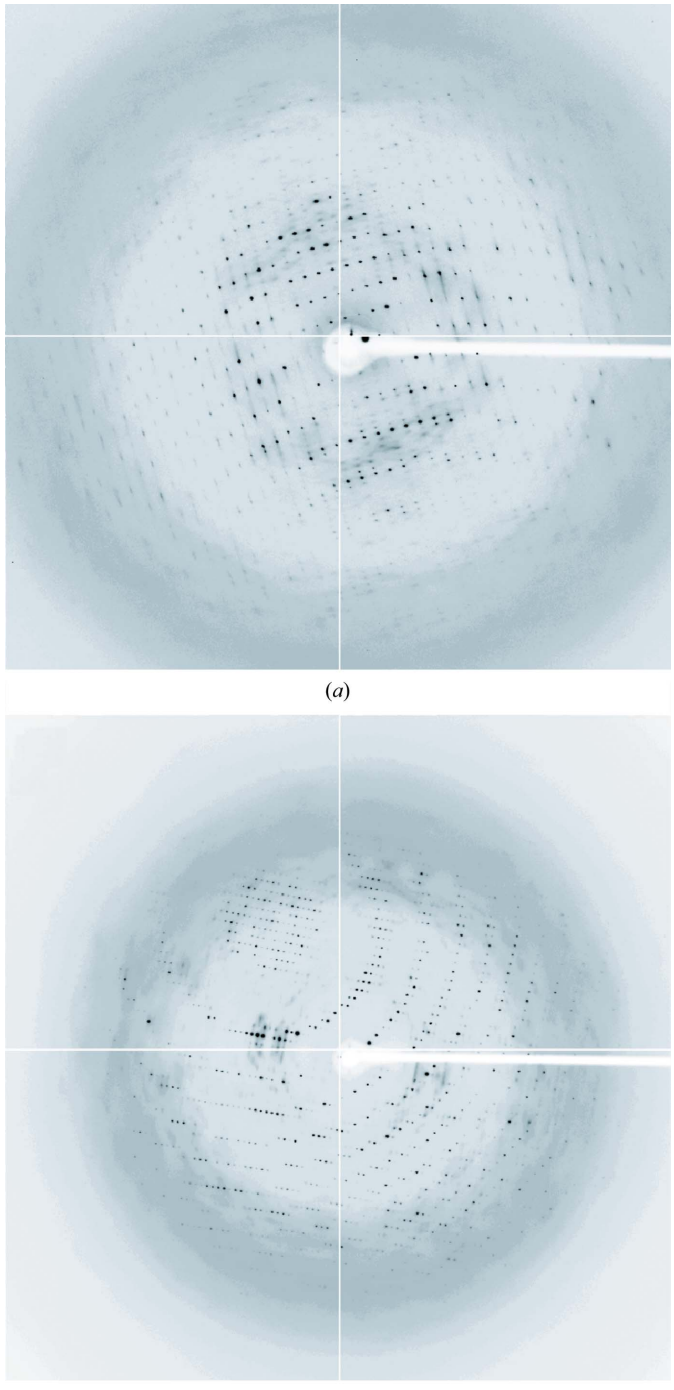
Initial crystallization conditions were obtained with Crystal Screen and Crystal Screen II (Hampton Research) using the hanging-drop vapour-diffusion method. The full-length $TmCorA_{1-351}$ protein was crystallized at room temperature by mixing equal volumes (2 + 2 µl) of protein solution (15 mg ml^{-1}) with a reservoir solution containing 20% PEG 400, 0.2 *M* $CaCl_2$, 0.1 *M* HEPES pH 7.0. Crystals appeared overnight and grew to maximum dimensions of ∼200 × 200 × 600 µm within a week (Fig. 1*a*). In addition to varying the protein and precipitant concentrations, the pH and temperature (277 and 295 K), extensive experimentation with the Additive Screen (96 compounds) and Detergent Screens 1, 2 and 3 (72 detergents) from Hampton Research failed to produce significant improvements in crystal quality as judged by X-ray diffraction. While crystallization without the inclusion of a divalent cation was possible (occurring over 1–2 weeks), providing Ca^{2+} proved to be the most reliable way to obtain crystals. Mg^{2+} did not improve crystal quality, Co^{2+} resulted in crystal twinning and crystals grown in the presence of Ni^{2+} did not exceed ∼20 µm in any dimension (Fig. 1*b*). Most of the detergents screened actually prevented crystallization, but many of those containing a maltoside or glucoside head group consistently allowed it. In this respect, detergents with shorter acyl chains increased the tendency for crystal twinning and this trend was only further exacerbated when using thio-substituted equivalents. The crystallization behaviour of the SeMet-substituted protein was essentially identical to that of the native protein. Crystals were transferred into mother liquor containing 30–35% PEG 400 with 0.026% DDM and then flash-frozen in a stream of boiling nitrogen.

The $AfCorA_{1-263}$ periplasmic domain was crystallized at room temperature by mixing equal volumes (2 + 2 µl) of protein solution (40 mg ml^{-1}) with a reservoir solution containing 1.3 *M* ammonium sulfate, 10 m*M* $CoCl_2$, 0.1 *M* HEPES pH 7.4. Crystals appeared over 1–3 d and grew to maximum dimensions of ∼100 × 100 × 300 µm (Fig. 1*c*). Interestingly, crystallization was found to be dependent upon the presence of a divalent cation, specifically Co^{2+}, and soaking these crystals in significant concentrations of other divalent cations (*e.g.* >1 m*M* Cd^{2+} or >20 m*M* Ni^{2+}) led to crystal cracking and/or serious decay of their diffraction pattern. The SeMet-substituted $AfCorA_{1-263}$ behaved similarly to the native protein. Crystals were transferred into mother liquor supplemented with 20% glycerol and then flash-frozen in a stream of boiling nitrogen.

The $TmCorA_{1-263}$ periplasmic domain was crystallized at room temperature by mixing equal volumes (2 + 2 µl) of protein solution (2.5 mg ml^{-1}) with a reservoir solution containing 4 *M* NaCl, 10 m*M* $CoCl_2$ and 0.1 *M* sodium citrate pH 5.6. Crystals appeared over 2–3 d but never exceeded a maximum size of ∼20 × 20 × 5 µm (Fig. 1*d*) unless they became twinned. Although crystallization without the inclusion of a divalent cation was possible, all further attempts to optimize or screen for useful additives failed. The SeMet-substituted $TmCorA_{1-263}$ behaved similarly to the native protein. Crystals were transferred into mother liquor containing 3.5 *M* NaCl supplemented with 20% glycerol and then flash-frozen in a stream of boiling nitrogen.

2.3. Data collection and processing

All data collections were performed at 100 K under a cold nitrogen stream. Numerous data sets were collected at the NSLS beamline X8C from full-length $TmCorA_{1-351}$ crystals using an ADSC Quantum 4 CCD detector. Our best $TmCorA_{1-351}$ native data were collected at a wavelength of 0.9799 Å (Table 1 and Fig. 2*a*). A three-wavelength MAD data set was collected from an SeMet-$TmCorA_{1-351}$ crystal at the Se absorption edge (peak at 0.9789 Å, edge at 0.9791 Å and remote at 0.9700 Å) on the APS SBC 19-BM beamline using an SBC-3 CCD (ANL-ECT) detector (Table 1). 360° of data were collected using 1 and 0.5° oscillations for the native and MAD $TmCorA_{1-351}$ data sets, respectively.



(*a*)

(*b*)

Figure 2
(*a*) Diffraction image of native $TmCorA_{1-351}$. The resolution is ∼3.5 Å at the side and ∼2.5 Å in the corners. Some reflections are observed beyond 3.5 Å. (*b*) Diffraction image of native $AfCorA_{1-263}$. The resolution is ∼2.5 Å at the side and ∼1.9 Å in the corners. Some reflections are observed beyond 2.7 Å.

Table 1
Data-collection statistics for $TmCorA_{1–351}$.

Values in parentheses are for the highest resolution shells: 3.83–3.70 and 4.76–4.60 Å for native and MAD data, respectively.

		MAD SeMet derivative		
	Native	Peak	Edge	Remote
Beamline	X8C, NSLS	19BM, APS		
Wavelength (Å)	0.9799	0.9789	0.9791	0.9700
Space group	*C*2	*C*2		
Unit-cell parameters (Å, °)	a = 214.25, b = 86.30, c = 181.53, β = 112.23, $\alpha = \gamma$ = 90	a = 214.40, b = 86.02, c = 181.93, β = 112.61, $\alpha = \gamma$ = 90		
Resolution (Å)	3.7	4.6	4.6	4.6
Measured reflections	424453	53126	50427	43488
Unique reflections	31195	15841	15307	13751
Completeness (%)	86.1 (73.8)	86.6 (46.0)	81.1 (37.9)	69.4 (17)
$\langle I/\sigma(I)\rangle$	19.7 (2.3)	10.9 (2.6)	12.0 (2.3)	11.0 (2.0)
R_{merge}† (%)	6.1 (46.5)	8.9 (21.3)	7.1 (20.7)	6.9 (24.1)

† $R_{merge} = \sum|I_i - \langle I_i\rangle|/\sum\langle I_i\rangle$, where I_i is the observed intensity and $\langle I_i\rangle$ is the average intensity over symmetry-equivalent measurements.

Table 2
Data-collection statistics for $AfCorA_{1–263}$.

Values in parentheses are for the highest resolution shells: 3.00–2.90 and 3.31–3.20 Å for native and MAD data, respectively.

		MAD SeMet derivative		
	Native	Peak	Edge	Remote
Beamline	X8C, NSLS	X8C, NSLS		
Wavelength (Å)	1.100	0.9784	0.9800	0.9600
Space group	$P6_122$	$P6_122$		
Unit-cell parameters (Å, °)	$a = b$ = 101.17, c = 142.87, $\alpha = \beta$ = 90, γ = 120	$a = b$ = 101.33, c = 142.71, $\alpha = \beta$ = 90, γ = 120		
Resolution (Å)	2.9	3.2	3.2	3.2
Measured reflections	267234	214598	216574	214044
Unique reflections	10141	7618	7658	7652
Completeness (%)	99.9 (99.9)	99.7 (100)	99.6 (100)	99.7 (100)
$\langle I/\sigma(I)\rangle$	63.5 (7.0)	37.3 (4.6)	38.4 (4.7)	42.0 (4.8)
R_{merge}† (%)	6.7 (43.7)	8.7 (42.3)	7.5 (39.3)	7.3 (41.1)

† $R_{merge} = \sum|I_i - \langle I_i\rangle|/\sum\langle I_i\rangle$, where I_i is the observed intensity and $\langle I_i\rangle$ is the average intensity over symmetry-equivalent measurements.

A three-wavelength MAD data set was collected from an SeMet-$AfCorA_{1–263}$ crystal at the Se absorption edge (peak at 0.9794 Å, edge at 0.9800 Å and remote at 0.9600 Å) at the NSLS beamline X8C (Table 2). A native $AfCorA_{1–263}$ data set was similarly collected, but using a wavelength of 1.100 Å (Table 2 and Fig. 2*b*). 200° of data were collected in 1° oscillations for all $AfCorA_{1–263}$ data sets. Data were indexed, integrated and scaled using *DENZO*/*SCALEPACK* or the *HKL*2000 suite (Otwinowski & Minor, 1997).

3. Discussion

We note that our purification procedure omits the isolation and subsequent solubilization of the membrane fraction, a time-consuming and labour-intensive step that is frequently encountered in most integral membrane-protein purification protocols. $TmCorA_{1–351}$ yielded crystals of indistinguishable quality independent of the inclusion of this membrane isolation and solubilization step. This shortened procedure should find practical application in the context of structural genomics and/or detergent-screening protocols.

Obtaining large crystals of the full-length magnesium transporter $TmCorA_{1–351}$ proved to be relatively straightforward (Fig. 1*a*); however, collecting high-quality diffraction data posed some problems. For instance, all $TmCorA_{1–351}$ crystals diffracted anisotropically; their mosaic spread typically ranged from ∼0.5 to 1.3° within a single crystal, where the length of the *a* axis varied by ±1 Å. Efforts to optimize crystal quality by extensive screening, *e.g.* the addition of small molecules, amphiphiles, detergents or organic compounds and variations in temperature, were all unsuccessful. Although the $TmCorA_{1–351}$ crystals proved to be extremely resilient, attempts to anneal or dehydrate the samples also failed to improve the resolution or quality of data. We feel that the limited resolution and quality of the $TmCorA_{1–351}$ crystals is an inherent property of this specific protein.

Analysis of the packing density within the $TmCorA_{1–351}$ crystals reveals that 4–6 molecules per asymmetric unit would yield reasonable values for the solvent content of an integral membrane-protein crystal (73.7–60.6%; V_M = 4.7–3.1 Å^3 Da^{-1}). If we assume that CorA forms tetramers, as reported in the literature (Warren *et al.*, 2004), the presence of a single transporter within the asymmetric unit seems likely.

To circumvent the limited data quality of the $TmCorA_{1–351}$ crystals, we first cloned CorA homologues from *A. fulgidus*, *Methanococcus jannaschii* and *Pyrococcus furiosus*. In contrast to our experience with $TmCorA_{1–351}$ and a report on *E. coli* CorA (Chen *et al.*, 2003), expression of these other full-length transporters failed to produce promising results. We then decided to pursue structural studies based on the possibility of producing stable soluble periplasmic domains. To this end, we cloned CorA periplasmic domains from *A. fulgidus*, *E. coli*, *M. jannaschii*, *T. maritima* and *P. furiosus* following a previous report to help determine domain boundaries (Warren *et al.*, 2004).

In contrast to their homologues, the CorA periplasmic domains $TmCorA_{1–263}$ and $AfCorA_{1–263}$ readily produced promising gel-filtration profiles. $TmCorA_{1–263}$ eluted at the molecular weight expected for a monomeric protein (∼30 kDa), with little or no signs of any higher order oligomer. In contrast, $AfCorA_{1–263}$ produced two prominent peaks; the first presumably represents an $AfCorA_{1–263}$ oligomer (∼120–150 kDa) and the second is an apparent $AfCorA_{1–263}$ monomer (∼30 kDa). When these two $AfCorA_{1–263}$ populations were independently pooled, concentrated and reapplied onto the gel-filtration column, their respective oligomeric states essentially remained unchanged. Crystallization trials were successful for the monomeric $TmCorA_{1–263}$ and the oligomeric $AfCorA_{1–263}$ proteins, but failed for the $AfCorA_{1–263}$ monomer. Although the space group of the $TmCorA_{1–263}$ periplasmic domain crystals remains undetermined owing to their limited diffraction power, they appear to have a similar morphology to the $AfCorA_{1–263}$ crystals (Figs. 1*c* and 1*d*). Analysis of the packing density within the $AfCorA_{1–263}$ crystals suggests that one molecule per asymmetric unit yields a reasonable solvent-content fraction for a soluble protein crystal (64.8%; V_M = 3.5 Å^3 Da^{-1}).

While this manuscript was in preparation, Lunin and coworkers (Structural Genomics Consortium) deposited the coordinates of a soluble domain (PDB code 2bbh) and the full-length TmCorA (PDB code 2bbj) in the Protein Data Bank. Structural analyses of $TmCorA_{1–351}$ at improved resolution in an alternate space group and of the homologous $AfCorA_{1–263}$ will hopefully shed more light on the selectivity and mechanism of magnesium transport.

We thank staff at the NSLS beamline X8C and the APS SBC beamline 19-BM for their time commitments and expert help. A joint grant from the Canadian Institutes of Health Research and National Sciences and Engineering Research Council of Canada enabled the use of beamline X8C at the National Synchrotron Light Source,

Brookhaven National Laboratory. Use of the Argonne National Laboratory Structural Biology Center beamline at the Advanced Photon Source was supported by the US Department of Energy, Office of Energy Research under Contract No. W-31-109-ENG-38. We acknowledge the gift of clones initially provided by Drs Aled Edwards and Alexei Savchenko, Ontario Centre for Structural Proteomics. JP would like to thank Ms Emily Cowan and Mrs Wanda Gillon for their continued support. This work was supported by the Canada Research Chairs Program and the National Sciences and Engineering Research Council of Canada (EFP).

References

Bara, M. & Guiet-Bara, A. (1984). *Magnesium*, **3**, 215–225.
Chen, Y., Song, J., Sui, S. F. & Wang, D. N. (2003). *Protein Expr. Purif.* **32**, 221–231.
Gardner, R. C. (2003). *Curr. Opin. Plant Biol.* **6**, 263–267.
Kehres, D. G. & Maguire, M. E. (2002). *Biometals*, **15**, 261–270.
Knoop, V., Groth-Malonek, M., Gebert, M., Eifler, K. & Weyand, K. (2005). *Mol. Genet. Genomics*, **274**, 205–216.
Maguire, M. E. & Cowan, J. A. (2002). *Biometals*, **15**, 203–210.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pfeiffer, J., Guhl, J., Waidner, B., Kist, M. & Bereswill, S. (2002). *Infect. Immun.* **70**, 3930–3934.
Schindler, M. & Osborn, M. J. (1979). *Biochemistry*, **18**, 4425–4430.
Smith, R. L., Gottlieb, E., Kucharski, L. M. & Maguire, M. E. (1998). *J. Bacteriol.* **180**, 2788–2791.
Warren, M. A., Kucharski, L. M., Veenstra, A., Shi, L., Grulich, P. F. & Maguire, M. E. (2004). *J. Bacteriol.* **186**, 4605–4612.
Worlock, A. J. & Smith, R. L. (2002). *J. Bacteriol.* **184**, 4369–4373.