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

Crystallization and data collection of the nucleotide-binding domain of Mg-ATPase

Kjell O. Håkansson* and Aida Ćurović

Department of Biology, August Krogh Building, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen Ø, Denmark

Correspondence e-mail: kohakansson@bio.ku.dk

Received 7 November 2008
Accepted 12 January 2009

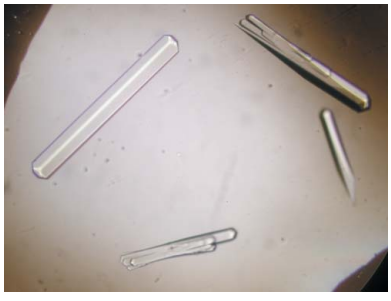
Understanding of how P-type ATPases work would greatly benefit from the elucidation of more high-resolution structures. The nucleotide-binding domain of Mg-ATPase was selected for structural studies because Mg-ATPase is closely related to eukaryotic Ca-ATPase and Na,K-ATPase while the nucleotide-binding domain itself has diverged substantially. Two fragments of Mg-ATPase were cloned in *Escherichia coli* and purified. The entire cytoplasmic loop (residues 367–673), consisting of the phosphorylation and nucleotide-binding domains, expressed well and was purified in large quantities. The smaller 19.5 kDa nucleotide-binding domain (residues 383–545) expressed less well but formed crystals that diffracted to a resolution of 1.53 Å which will be used for molecular replacement.

1. Introduction

P-type ATPases are integral multispan membrane proteins that are found in most organisms and in all kingdoms of life. They couple transmembrane transport of cations to ATP hydrolysis. They derive their name from the fact that this process involves a phosphorylated aspartic acid intermediate (Degani & Boyer, 1973). They can be grouped into subfamilies based on sequence similarity (Axelsen & Palmgren, 1998). Three of these subfamilies are represented in the *Escherichia coli* genome (Blattner *et al.*, 1997): the potassium pump Kdp belongs to subfamily 1A, the two heavy metal-ion transporters CopA and ZntA belong to subfamily 1B and the magnesium transporter MgtA belongs to subfamily 3B.

The Mg-ATPase MgtA is only found in prokaryotes, in which it is widely distributed. Some species (*e.g. Salmonella typhimurium*) have an additional Mg-ATPase called MgtB. These two proteins are closely related and contain ten transmembrane helices (Fig. 1*a*; Smith *et al.*, 1993; Maguire, 2006; Snavely *et al.*, 1991). In this respect, as well as in their amino-acid sequences, they seem to be more closely related to eukaryotic P-type ATPases such as Ca-ATPase (subfamily 2A) and Na,K-ATPase (subfamily 2C) than to the members of the prokaryotic subfamilies 1A and 1B. From the known structures of Ca-ATPase in different conformations, it is clear that ion transport is driven by the relative movements of the three cytoplasmic domains: the nucleotide-binding domain (N-domain), the phosphorylation domain (P-domain) and the actuator domain (A-domain). These movements, which are fuelled by ATP hydrolysis, are transmitted as conformational changes to the transmembrane region, in which two calcium ion-binding sites are found. The binding sites are made up of residues from transmembrane helices 4, 5, 6 and 8 and the resulting movements of these helices lead to changes in binding and occlusion of the ions. The overall result of these sequentially ordered changes is the transport of two calcium ions from the cytoplasm to the lumen (Jorgensen & Andersen, 1988; Toyoshima *et al.*, 2007; Toyoshima & Inesi, 2004).

We have cloned two fragments of *E. coli* Mg-ATPase: the nucleotide-binding domain and the entire cytoplasmic loop consisting of both the phosphorylation domain and the nucleotide-binding domain; the latter forms an insertion in the phosphorylation domain (Fig. 1*a*). They were cloned in high-expression pQE vectors which were transformed into *E. coli*. The proteins were expressed and



© 2009 International Union of Crystallography
All rights reserved

purified for the purpose of structural studies of the nucleotide-binding site of this hitherto uncharacterized protein. Since the nucleotide-binding domain is the part of the P-type ATPases in which the sequences have diverged most, such a structure would shed further light on how these proteins bind ATP. Crystals of the nucleotide-binding domain that diffracted to high resolution and were suitable for structure elucidation have been obtained.

2. Material and methods

2.1. Cloning

All polymerase chain reactions were performed with Pwo polymerase (Roche, Basel, Switzerland). The gene for full-length MgtA was amplified from the *E. coli* chromosome with the primers ACACACGAATTCATTAAAGAGGAGAAATTAACTATGTTTAAAGAAATTTTTACCCGG (forward) and ACACACAAGCTTCTATTAGTGATGGTGATGGTGATGTTGCCAGCCGTAACGACGG (reverse). The product was obtained after 40 cycles with an annealing temperature of 332 K and an elongation time of 4.5 min. The product was digested and inserted into a pQE-40 vector between *Eco*RI and *Hin*dIII restriction sites. This plasmid was subsequently used as a template for amplification of the parts of the gene corresponding to the cytoplasmic domains. The beginnings and the ends of these domains were determined by sequence alignment and analysis of the Ca-ATPase structure (Toyoshima *et al.*, 2000). The nucleotide-binding domain gene fragment and the entire cytoplasmic loop (phosphorylation domain plus nucleotide-binding domain) fragment were amplified using the following primers: ACACACGAATTCATTAAAGAGGAGAAATTAACTATGATTGTGCTGGAGATACATACC (N-dom, forward), ACACACAAGCTTTTAGTGATGGTGATGGTGATGATCAAGAAAAGCAATATATCCTTC (N-dom, reverse), ACACACGAATTCATTAAAGAGGAGAAATTAACTATGGATATTCTGTGCACTGATA (PN-dom, forward) and ACACACAAGCTTTTAGTGATGGTGATGGTGATGCAGGATGATATCAGCCGCTT (PN-dom, reverse). The forward primers were designed to contain a Shine–Dalgarno sequence and all reverse primers code for a C-terminal hexahistidine tag.

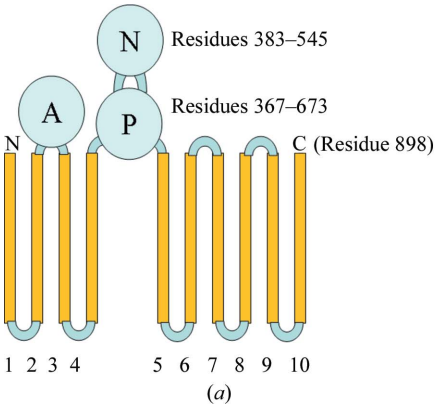
Initial experiments using 45 cycles and an annealing temperature of 332 K with elongation times of 30 and 70 s gave relatively poor yields. For the gene fragment corresponding to the nucleotide-binding domain, the number of cycles was increased to 47. Amplification of the fragment corresponding to the entire cytoplasmic loop (phosphorylation domain plus nucleotide-binding domain) was repeated with the initial PCR product as a template and an annealing temperature of 338 K. The amino-acid sequences of the gene products, which were verified by nucleic acid sequencing (Eurofins MWG Operon, Ebersberg, Germany), are shown in Fig. 1(*b*). The plasmids were heat-shock transformed into *E. coli* XL1-blue cells, which were stored at 193 K in 24% glycerol.

2.2. Purification

For expression, cells were grown in one to four Erlenmeyer flasks each containing 1 l LB medium with ampicillin (50 µg l^{-1}) and tetracycline (8 µg l^{-1}) under vigorous shaking at 310 K. Cells were induced with 1 m*M* IPTG when the OD_{600} reached 0.5, harvested by centrifugation 3 h later and stored at 253 K. Cells were thawed, resuspended in 10 ml 50 m*M* Tris–HCl, 50 m*M* NaCl, 20 m*M* imidazole pH 7.5 and sonicated on ice. The insoluble fraction was pelleted and the supernatant was loaded on an Ni–NTA Superflow column (5 ml; Qiagen, Germantown, Maryland, USA) equilibrated with the same buffer and eluted with an imidazole gradient (20–250 m*M*) using an ÄKTA purifier (GE Healthcare, Chalfont St Giles, England). Fractions containing the protein were pooled and loaded on a pre-packed 10/100 MonoQ column (GE Healthcare), washed with 50 m*M* Tris–HCl pH 7.5 and eluted with an NaCl gradient (0–1 *M* over five column volumes). The protein was concentrated to 10 mg ml^{-1} using Centricon ultrafilters (Millipore Corporation, Bedford, Massachusetts, USA). The protein concentration was determined from an $E_{280,\,0.1\%}$ estimate of 0.87 for the nucleotide-binding domain, which is equivalent to an extinction coefficient of 17 000 M^{-1} cm^{-1}.

2.3. Crystallization

Crystallization screens were set up using vapour equilibration in hanging drops at 293 K. A small amount of protein solution (1–3 µl, depending on the availability of protein) was mixed with an equal amount of precipitant and vapour-equilibrated against 0.5 ml precipitant.



(*a*)

```
MDILCTDKTGTLTQDKIVLE NHTDISGKTSERVLHSAWLN SHYQTGLKNLLDTAVLEGTD
367               383
EESARSLASRWQKIDEIPFD FERRRMSVVVAENTEHHQLV CKGALQEILNVCSQVRHNGE

IVPLDDIMLRKIKRVTDTLN RQGLRVVAVATKYLPAREGD YQRADESDLILEGYIAFLDP
                                                              345
PKETTAPALKALKASGITVK ILTGDSELVAAKVCHEVGLD AGEVVIGSDIETLSDDELAN

LAQRTTLFARLTPMHKERIV TLLKREGHVVGFMGDGINDA PALRAADIGISVDGAVDIAR

EAADIILHHHHHH
       674
```

(*b*)

Figure 1
(*a*) The overall topology of Mg-ATPase with ten transmembrane helices (yellow bars) and three cytoplasmic domains: the actuator domain (A), the phosphorylation domain (P), which contains the phosphorylation residue Asp373, and the nucleotide-binding domain, which is an insertion in the phosphorylation domain. (*b*) The sequence of the cloned nucleotide-binding domain (Ile383–Asp345, underlined sequence) and the fragment consisting of both the phosphorylation and nucleotide-binding domains (Met367–Leu673, entire sequence). The N-terminal methionine and the terminal hexahistidyl tag have been underlined in order to illustrate that they were also added to the termini of the smaller nucleotide-binding domain. The molecular masses of the two fragments are 19.5 and 34.5 kDa, respectively. The full-length protein consists of 898 amino acids.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	MAX-lab beam station I911-2 (Lund, Sweden)
Wavelength (Å)	1.0379
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 99.6$, $c = 46.2$
Resolution (Å)	27.63–1.53 (1.61–1.53)
Completeness (%)	94.7 (70.0)
$R_{\rm merge}$ (%)†	4.2 (20.6)
$\langle I/\sigma(I)\rangle$	28.1 (10.1)
No. of unique observations	33678 (3949)
Redundancy	7.5 (6.0)
Wilson B factor (Å^2)	19.8

† Observed R factor from *XDS* (Kabsch, 1993): $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$.

2.4. Data collection

Crystals were transferred to fresh mother liquor made up with 50%(*w*/*v*) sucrose and cryocooled in a stream of nitrogen gas at 100 K. In-house data were collected using a Rigaku RU-300 X-ray generator and a MAR345 image-plate detector and were processed with *MOSFLM* (Leslie, 1992) and *SCALA* (Collaborative Computational Project, Number 4, 1994). Synchrotron data were collected using a CCD detector on MAX-lab beam station I911-2 (Lund, Sweden) and were processed with *XDS* (Kabsch, 1993).

3. Results and discussion

The expression of the entire cytoplasmic loop (phosphorylation domain plus nucleotide-binding domain) was sufficient for extensive but fruitless crystallization screenings. Expression of the nucleotide-binding domain, on the other hand, yielded only 0.7 mg purified protein from 4 l culture, but fortunately sizable and well diffracting crystals were obtained in the first attempt with Crystal Screen (Hampton Research, Aliso Viejo, California, USA) from condition No. 39 (0.1 *M* HEPES pH 7.5, 2% PEG 400, 2 *M* ammonium sulfate). The crystals, shown in Fig. 2, appeared after 1–2 d and displayed a rod-like tetragonal morphology. The largest crystals were approximately 0.8 mm in length and grew only marginally upon prolonged incubation. HEPES was subsequently replaced by Tris–HCl buffer, but small variations around these conditions did not result in any apparent improvement in crystal size. Crystallization was inhibited by the dual presence of 100 m*M* ATP and Mg^{2+}. Crystals were cryocooled in mother liquor made up with 50%(*w*/*v*) sucrose prior to data collection. A data set collected in-house was used to determine the space group as either $P4_12_12$ or $P4_32_12$ and data were subsequently collected to 1.53 Å resolution at MAX-lab (Lund, Sweden) beam station I911-2. The unit-cell parameters were $a = b = 99.6$, $c = 46.2$ Å. Assuming the presence of one molecule per asymmetric unit gives a Matthews coefficient of 2.9 Å^3 Da^{-1} and a solvent content of 58%. Data-collection statistics are displayed in Table 1.

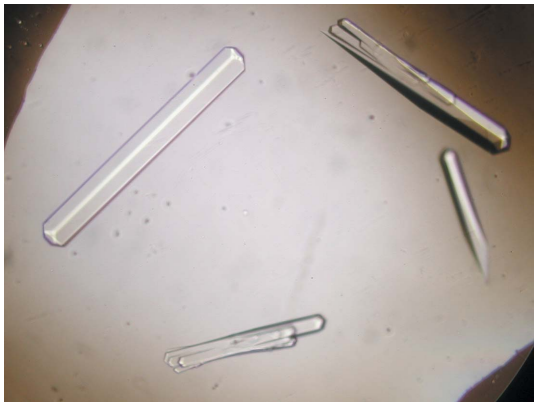


Figure 2
Crystals of Mg-ATPase nucleotide-binding domain. The longest crystal is approximately 0.8 mm in length.

For molecular replacement, the structure of the Na,K-ATPase nucleotide-binding domain (PDB code 13qi; Hakansson, 2003; Haue *et al.*, 2003) was modified by removing all those parts in which the backbone was not superimposable on the Ca-ATPase nucleotide-binding domain (Toyoshima *et al.*, 2000). Using default parameters, the *CCP*4 version of *AMoRe* (Navaza, 1994; Collaborative Computational Project, Number 4, 1994) resulted in no outstanding solutions in either space group $P4_12_12$ or $P4_32_12$. However, the apparent best solution ($R_{\rm cryst}$ = 54% for the 20.0–3 Å resolution range) had a $D_{\rm min}$ (the shortest distance between the centres of mass of symmetry-equivalent molecules) of 33 Å. Furthermore, it displayed a realistic intermolecular packing with neither backbone overlap of neighbouring molecules nor excessively wide gaps between molecules. This solution in space group $P4_12_12$ was refined using the *AMoRe fitfun* procedure to a crystallographic R value of 50.5%. Rigid-body refinement, atomic positional refinement and individual B-factor refinement with *CNS* (Brünger *et al.*, 1998) yielded an $R_{\rm work}$ and an $R_{\rm free}$ of 45% and 47%, respectively, for the entire resolution range. Together with the quality of the electron-density maps, this indicates that the solution is correct. The structure will be rebuilt according to the Mg-ATPase sequence and refined.

The purified *E. coli* genome was a gift from Professor Lars Hederstedt, Lund. Dr Johan Unge is thanked for his help during data collection and Professors Sine Larsen and Leila Lo Leggio are thanked for sharing synchrotron beam time. This work was supported by the Danish Research Council FNU and the Carlsberg Foundation.

References

Axelsen, K. B. & Palmgren, M. G. (1998). *J. Mol. Evol.* **46**, 84–101.
Blattner, F. R. *et al.* (1997). *Science.* **277**, 1453–1474.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Degani, C. & Boyer, P. D. (1973). *J. Biol. Chem.* **248**, 8222–8226.
Håkansson, K. O. (2003). *J. Mol. Biol.* **332**, 1175–1182.
Haue, L., Pedersen, P. A., Jorgensen, P. L. & Håkansson, K. O. (2003). *Acta Cryst.* D**59**, 1259–1261.
Jorgensen, P. L. & Andersen, J. P. (1988). *J. Membr. Biol.* **103**, 95–120.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Leslie, A. G. W. (1992). *Jnt CCP4–ESF/EACBM Newsl. Protein Crystallogr.* **26**.
Maguire, M. E. (2006). *Front. Biosci.* **11**, 3149–3163.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Smith, D. L., Tao, T. & Maguire, M. E. (1993). *J. Biol. Chem.* **268**, 22469–22479.
Snavely, M. D., Miller, C. G. & Maguire, M. E. (1991). *J. Biol. Chem.* **266**, 815–823.
Toyoshima, C. & Inesi, G. (2004). *Annu. Rev. Biochem.* **73**, 269–292.
Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. (2000). *Nature (London)*, **405**, 647–655.
Toyoshima, C., Norimatsu, Y., Iwasawa, S., Tsuda, T. & Ogawa, H. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 19831–19836.