Acta Cryst. (1998). D**54**, 1460–1463

Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the receiver domain of PhoB

Maria Solà,[a] F.-Xavier Gomis-Rüth,[a] Alicia Guasch,[a] Luis Serrano[b] and Miquel Coll[a]* at [a]*Departament de Biologia Molecular i Cellular, Centre d'Investigació i Desenvolupament-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain, and* [b]*European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. E-mail: mcccri@cid.csic.es*

(*Received 24 March 1998; accepted 27 April 1998*)

Abstract

PhoB is the response regulator of the *E. coli* two-component signal transduction system for phosphate regulation. It is a transcription factor that activates more than 30 genes of the *pho* regulon. Crystals of the receiver domain of PhoB were obtained by applying the hanging-drop vapour-diffusion method. X-ray diffraction data have been collected using synchrotron radiation to 1.88 Å resolution. The crystals belong to the orthorhombic space group $P2_12_12_1$ with unit-cell constants a = 34.11, b = 60.42, c = 119.97 Å. The Matthews parameter suggests that PhoB crystallizes with two molecules per asymmetric unit, suggesting that activating dimerization occurs in the crystal.

1. Introduction

Under inorganic phosphate starvation conditions a two-component signal transduction system is activated in *E. coli* and other bacteria (Tae-Yoon *et al.*, 1989). This system is controlled by the transmembrane sensor kinase PhoR and the response regulator PhoB, the latter being activated by the transfer of a phosphate group from the former in the presence of magnesium (Makino *et al.*, 1989). PhoB is a transcription factor that, once activated, induces the gene expression of the phosphate (*pho*) regulon (Shinagawa *et al.*, 1983), which includes genes whose products are involved in phosphorous uptake and use (Wanner, 1996). Phosporylation of PhoB takes place at its N-terminal 127-amino-acid receiver domain (Hiratsu *et al.*, 1995), the sequence and structure of which are homologous to those of other response-regulator superfamily components (Volz, 1993). Interaction of PhoB with *pho* promoters occurs *via* the approximately 100-residue-long C-terminal DNA-binding or output domain (Makino *et al.*, 1996), which belongs to the σ70 RNA-polymerase subunit-dependent class (Pao & Saier, 1995) and presumably contains a winged-helix motif, as does the related OmpR protein (Martinez-Hackert & Stock, 1997). PhoB binds to DNA at a specific consensus sequence called the *pho* box, which consists of a seven base-pair direct repeat (Makino *et al.*, 1986). Because of the repeat in the *pho*-box sequence, it is believed that DNA binding of PhoB takes place in dimers (Makino *et al.*, 1989). In fact, PhoB has an intrinsic dimerization tendency, enhanced upon phosphorylation (McCleary, 1996), which can be emulated by the excision of the output DNA-binding domain (Fiedler & Weiss, 1995).

In this paper we describe the cloning, overexpression, purification and crystallization of the PhoB receiver domain and the subsequent X-ray data collection using synchrotron radiation. Calculation of the Matthews parameter (Matthews, 1968) suggests that PhoB crystallizes with two molecules per asymmetric unit.

2. Materials and methods

2.1. *Cloning of PhoB receiver domain*

The DNA for the PhoB receiver domain was obtained by the polymerase chain reaction (PCR) amplification technique using the following designed dideoxyoligonucleotides: primer 1, GCGCCCATGGCGAGACGTATTCTGGTC, and primer 2, GTCCGAAGCTTTTACTACGCCATTGGCGAAATACGG, with melting points of 331 and 333 K, respectively; the oligonucleotides introduced NcoI and HindIII restriction sites and initiation and termination codons at the 5′ and 3′ ends, respectively. The PCR solution contained, as the DNA template, a bacterial sample punctuated from a single colony of *E. coli* XL1-blue strain. It also contained 1 unit of polVent$_R$ DNA polymerase (Biolabs) in 100 µl of 10 m*M* KCl, 10 m*M* $(NH_4)_2SO_4$, 20 m*M* Tris–HCl (pH 8.8 at 298 K), 2 m*M* $MgSO_4$, 200 µ*M* of each dNTP, 0.1% Triton X-100 and 1 µ*M* of each primer. Three PCRs were performed at temperatures of 328, 333 and 335 K.

2.2. *Protein expression and purification*

The amplified PhoB receiver-domain DNA fragment, obtained from the PCR at the highest annealing restrictive temperature, was ligated into the T7*lac* expression vector pBAT-4 (Peränen *et al.*, 1996) and cloned into the BL21 strain of *E. coli* by electroporation. A single colony was punctuated from transformed bacteria and grown in 4 ml of LB liquid culture (Sambrook *et al.*, 1989) at 310 K until an OD of ~0.4 at 600 nm was reached. IPTG (Promega) was then added to a final concentration of 0.167 m*M* in order to induce overexpression. As these cells showed good overexpression of the receiver domain, several large liquid cultures of 1 l were grown and overexpression induced as before. Cells were extracted from these cultures by centrifugation at 4000*g* and the pellet was resuspended in 20 m*M* Tris–HCl pH 8.0 (4 ml per litre of culture). Afterwards, cells were lysed by sonification, and their insoluble and soluble fractions were separated by centrifugation at 15000*g* for 30 min. The supernatant of this last centrifugation was treated with DNAse (Boehringer) and concentrated on Centriprep 30 concentrators (Amicon) until precipitation began to appear. A 4000*g* centrifugation of this last solution yielded a pellet containing the precipitated protein. The pellet was solubilized by denaturation with 6 *M* urea, and refolding of the protein was achieved with a dilution of 10 volumes of 20 m*M* Tris–HCl pH 8.0; concentration of the last solution followed. Both the initial and the refolding solutions were filtered by means of a 0.22 µm filter and subse-

quently loaded on a gel filtration HiLoad 26/60 Superdex 75 column (Pharmacia) previously equilibrated with 20 m*M* Tris–HCl pH 8.0 (Fluka). Several cycles of gel filtration alternating with concentration of the purest eluted fractions followed, until the PhoB receiver-domain purity reached more than 95% (*C* in Fig. 1). Finally, purified samples were flash frozen in liquid nitrogen and stored at 203 K.

2.3. *Protein concentration determination and mass spectroscopic analysis*

The molar extinction coefficient for the PhoB receiver domain was determined by following the method of Gill & von Hippel (1989). The formula

$$\varepsilon_{M,\mathrm{nat}} = (\mathrm{Abs}_{\mathrm{nat}})(\varepsilon_{M,\mathrm{Gdn.HCl}})/(\mathrm{Abs}_{\mathrm{Gdn.HCl}})$$

was used for a wavelength of 280 nm yielding a value of 12940 M^{-1} cm^{-1}. The purified PhoB receiver domain was analysed by mass spectroscopy at the EMBL Peptide and Protein Service, Heidelberg (Germany). The analysis yielded a mass of 14367.3 $\pm$ 2.6 Da.

2.4. *Crystallization*

Crystallization trials were performed using the hanging-drop vapour-diffusion method at 395 K. 3 µl of purified protein solution at 6.2 mg ml^{-1} concentration were mixed with 6 µl of crystallization solution equilibrated against 700 µl of mother-liquor reservoir. The crystallization device consisted of tissue-culture plates (Linbro) with mother-liquor-filled wells covered with 22 $\times$ 22 mm coverslips which had been siliconized with dichlorodimethylsilane (Merck). Each well and the corresponding coverslip were sealed together with high-vacuum grease (Dow Corning). Initial screening for crystallization conditions was performed using the Hampton Research crystallization kits Crystal Screen I and II.

2.5. *Native data collection, analysis and molecular-replacement calculations*

For conservation and manipulation, crystals were harvested with a modified crystallization solution containing 30% polyethylene glycol. A native data set was collected using synchrotron radiation (λ = 0.885 Å) at DESY (Hamburg) on the EMBL-BW7B beamline with a MAR Research 300 imaging-plate detector at 120 K using an Oxford cryosystem device. Protection from cryocooling was achieved by passing a harvested crystal mounted on a loop through an ethylene glycol drop just prior to mounting in the nitrogen stream. High- (up to 1.88 Å; Fig. 2) and low-resolution data were collected separately in 0.5 and 1° oscillation ranges, respectively. Each data set was indexed and processed with *DENZO* (Otwinowski & Minor, 1997) and scaled and merged with *SCALEPACK* (Otwinowski & Minor, 1997). Molecular-replacement calculations were carried out with *AMoRe* (Navaza, 1994) and *XPLOR* (Brünger, 1992) using starting models based on CheY (PDB code 1CHN; Volz & Matsumura, 1991), NarL (PDB code 1RNL; Baikalov *et al.*, 1996), Spo0F (PDB code 1SRR; Madhusudan *et al.*, 1996) and CheY–T87I mutant (PDB code 1VLZ; Ganguli *et al.*, 1995) modified according to the PhoB sequence. A heavy-ion-containing derivative search was performed by soaking native crystals in appropriate solutions.

3. Results and discussion

3.1. *Genetic construct, overexpression and purification*

The oligonucleotides for the PCR were designed based on the DNA sequence obtained from the EMBL database (code X04026). The open reading frame of the DNA sequence was deduced by comparing the three possible amino-acid sequences (obtained starting from three consecutive bases) to the true sequence of the protein. The last residue for the N-terminal domain was chosen following the superposition of the CheY superfamily proteins reviewed by Volz (1993) and considering that the last residue should not possess hydro-

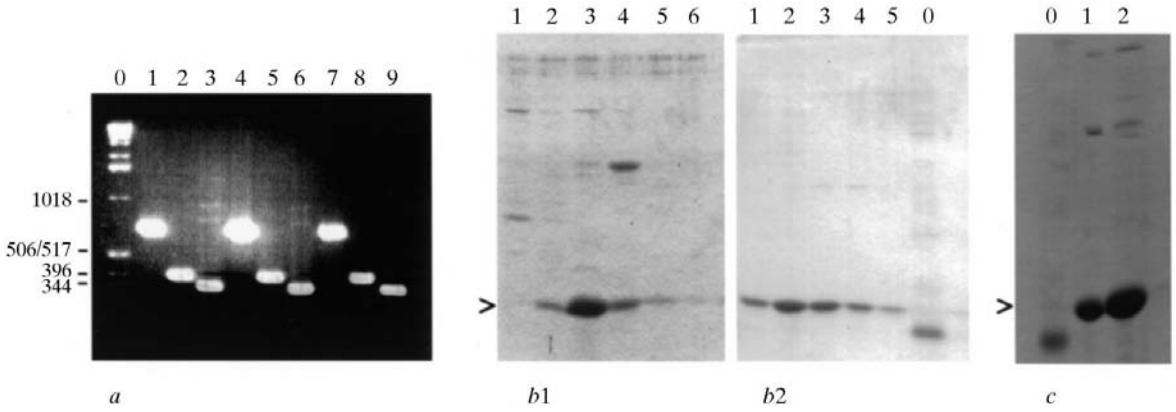

Fig. 1. The PCR yielded a high amount of amplified DNA fragments. In *a*, lanes 1, 4 and 7 contain the whole PhoB DNA, lanes 2, 5 and 8 contain amplified DNA of the PhoB receiver domain and lanes 3, 6 and 9 contain the DNA of the PhoB DNA-binding domain. In *b*1 and *b*2, eluted fractions from gel filtration were collected together and concentrated for subsequent crystallization. PhoB receiver domain is indicated with an arrowhead. In *b*1, the purest fractions (lanes 2 and 3) eluted from the soluble cell fraction aliquot were pooled with similar fractions from other gel-filtration loadings. In *b*2, all eluted fractions from the refolding solution (lanes 1 to 5) were pooled, together with the fractions in *b*1. Highly pure PhoB receiver domain (*b*), suitable for subsequent crystallization trials, was achieved after two cycles of gel filtration alternating with pooling of the purest aliquotes and concentration. Lanes 0 are standards; for SDS–acrylamide gels, lactalbumin ($M_w \simeq$ 14200 Da, Sigma) was used.

Table 1. *Data-collection and processing statistics*

	High resolution	Low resolution	Combined data
Resolution limits	20–1.88 Å	40–3.11 Å	40–1.88 Å
Number of observations	55230	16935	—
Number of unique reflections	17118	4473	17784
Completeness (whole range)	79.2	87.5	84.8
Completeness (last shell)	73.4 (1.95–1.88 Å)	91.9 (3.23–3.11 Å)	73.4 (1.95–1.88 Å)
$R_{\rm merge}$† (%)	3.9	3.2	2.4
$R_{\rm merge}$† (last shell) (%)	11.2	5.1	—
$\langle I/\sigma(I)\rangle$	28.2	41.8	—
$\langle I/\sigma(I)\rangle$ (last shell)	7.5	20.1	—
Data with $I > \sigma(I)$ (%)	77.4	86.2	—
Data with $I > \sigma(I)$ (last shell) (%)	69.2	89.3	—
Average multiplicity	3.22	3.78	—

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

phobic properties. The introduction of restriction sites at both the 3′ and 5′ ends of the PCR product did not result in a change of the original PhoB sequence. Inspection with agarose gels of the PCR products at the three annealing temperatures showed a great amount of amplified DNA in all of them (*a* in Fig. 1). pBAT-4 was found to be a satisfactory vector for the PhoB receiver-domain overexpression in the BL21 *E. coli* strain. After cell lysis and centrifugation, a large amount of protein was observed to be located in the soluble cell fraction as checked by SDS–acrylamide gels. Concentration of this cell fraction resulted in opacity of the solution. The aggregated protein was extracted by centrifugation and its solubilization was achieved with a urea solution. Subsequent refolding consisted of diluting the last denaturating solution with 10 volumes of an aqueous solution. Gel filtration of both the concentrated soluble cell fraction and the refolding solution yielded fractions which displayed clear bands in SDS–acrylamide gels, where PhoB was the major component (*b* in Fig. 1).

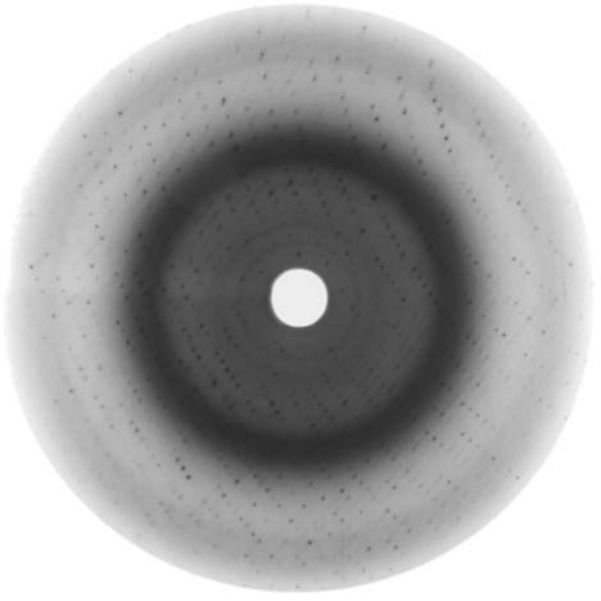

Fig. 2. X-ray diffraction image of the PhoB receiver domain corresponding to the high-resolution data set. The oscillation range was 0.5°. The highest resolution at the detector edge is 1.67 Å.

Purification yielded 177.45 mg of protein from 7.5 l of cell cultures. After confirmation that the purified protein was PhoB by mass spectroscopy, it was stored as samples of 16.5 and 6.2 mg ml^{-1} at 203 K.

3.2. *Crystallization and native data collection*

The optimal crystallization conditions were achieved by mixing 3 µl of 6.2 mg ml^{-1} protein solution with 6 µl of a crystallization solution containing 20%(*w*/*v*) polyethylene glycol 4000 (Fluka), 0.1 *M* Tris–HCl (Fluka) pH 8.0, 0.4 *M* sodium acetate (J. T. Baker) and 0.02% sodium azide (J. T. Baker). The drop was allowed to equilibrate against a reservoir consisting of 700 ml of crystallization solution. PhoB receiver-domain crystals appeared after one week and grew to a maximum size of 0.5 × 0.2 × 0.1 mm. They showed morphological and diffraction decay after about two months. Indexing clearly indicated that the crystal belonged to a primitive orthorhombic space group and systematic absences suggested it to be $P2_12_12_1$. Cell constants are a = 34.11, b = 0.42, c = 119.97 Å. The calculation of the most likely Matthews parameter (Matthews, 1968) yields a value of 2.15 Å^3 Da^{-1} for two molecules per asymmetric unit. This is consistent with the tendency of this domain to dimerize in the absence of the DNA-binding domain (DBD) (Fiedler & Weiss, 1995). Dimerization upon excision of DBD indicates that the activating dimerization of PhoB takes place through the receiver domain. It is probable that excision of the DBD leads to a conformationally activated form of the receiver domain which is close to the activated phosphorylated conformation, and this may have been reproduced in this work. Moreover, the value of the Matthews parameter for these crystals, as well as the percentage solvent content of about 33%, indicate tight packing of the molecules in the cell. From this crystal one high- and one low-resolution data set were collected and merged. Statistical parameters are given in Table 1.

3.3. *Molecular-replacement calculation and heavy-ion-derivative search*

Firstly, an extensive molecular-replacement search through the different possible primitive orthorhombic space groups was carried out, but no solution appeared using different models. Failure of molecular replacement could be due to structural differences between PhoB and the known receiver three-dimensional structures used as starting models and/or the

close-packing interactions between the monomers. As a consequence of the lack of progress with molecular replacement, a heavy-ion-derivative search was pursued to obtain starting phases. Knowledge of both the PhoB receiver-domain sequence and the three-dimensional structure of CheY led us initially to try compounds containing heavy ions with a low number of putative binding sites, such as uranium and lanthanides (Sm, Eu), which bind to clusters of carboxylate groups of glutamate or aspartate residues (in the active site of CheY), or compounds such as mercurials which have affinity for histidines and cysteines, which are present in a very low number in PhoB. To date, not only these specific heavy ions but also non-specific ones such as gold or platinum have been tried. 16 different data sets for soaked crystals have been collected, processed and analysed by difference Patterson techniques, but strong non-isomorphism mainly affecting the *c* cell axis (up to 9.7% variation) prevented calculation of proper phases. The search for better soaking conditions and co-crystallization with heavy ions are in progress. Multiple-wavelength anomalous diffraction (MAD) experiments are also planned, since this method does not require isomorphism of the derivative crystals.

This work was supported by the Ministerio de Educación y Ciencia (grant PB95-0224) and Generalitat de Catalunya (Centre de Referència en Biotecnologia and grant 1997SGR-275). Data collection at EMBL-DESY was supported by the EU Large Installations Project CHGE-CT93-0040.

References

Baikalov, I., Schröder, I., Kaczor-Grzeskowiak, M., Grzeskowiak, K., Gunsalus, R. P. & Dickerson, R. D. (1996). *Biochemistry*, **35**, 11053–11061.

Brünger, A. T. (1992). *XPLOR, Version* 3.1. *A System for X-ray Crystallography and NMR*. New Haven: Yale University Press.

Fiedler, U. & Weiss, V. (1995). *EMBO J.* **14**, 3696–3705.

Ganguli, S., Wang, H., Matsumura, P. & Volz, K. (1995). *J. Biol. Chem.* **270**, 17386–17393.

Gill, S. T. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.

Hiratsu, K., Nakata, A., Shinagawa, H. & Makino, K. (1995). *Gene*, **161**, 7–10.

McCleary, W. R. (1996). *Mol. Microbiol.* **20**, 1155–1163.

Madhusudan, Zapf, J., Whiteley, J. M., Hoch, J. A., Xuong, N. H. & Varughese, K. I. (1996). *Structure*, **4**, 679–690.

Makino, K., Amemura, M., Kawamoto, T., Kimura, S., Shinagawa, H., Nakata, A. & Suzuki, M. (1996). *J. Mol. Biol.* **259**, 15–26.

Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986). *J. Mol. Biol.* **190**, 37–44.

Makino, K., Shinagawa, H., Amemura, M., Yamasa, M. & Nakata, A. (1989). *J. Mol. Biol.* **210**, 551–559.

Martinez-Hackert, E. & Stock, A. M. (1997). *Structure*, **5**, 109–124.

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

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

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

Pao, G. M. & Saier, M. H. J. (1995). *J. Mol. Evol.* **40**, 136–154.

Peränen, J., Rikkonen, M. & Kääriäinen, L. (1996). *Anal. Biochem.* **236**, 371–373.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning*. New York: Cold Spring Harbor Laboratory Press.

Shinagawa, H., Makino, K. & Nakata, A. (1983). *J. Mol. Biol.* **168**, 477–488.

Tae-Yoon, L., Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1989). *J. Bacteriol.* **171**, 6593–6599.

Volz, K. (1993). *Biochemistry*, **32**, 11741–11753.

Volz, K. & Matsumura, P. (1991). *J. Biol. Chem.* **266**, 15511–15519.

Wanner, B. L. (1996). *Escherichia Coli and Salmonella Typhimurium: Cellular and Molecular Biology*, edited by F. C. Neidhardt, R. I. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. J. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger, pp. 1357–1381. Washington, DC: American Society for Microbiology.