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

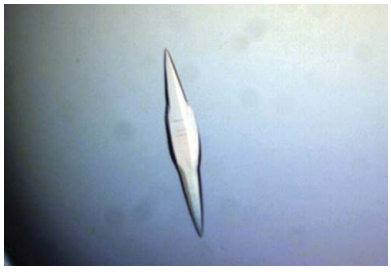
Purification and crystallization of Phd, the antitoxin of the *phd*/*doc* operon

Abel Garcia-Pino,[a,b] Yann Sterckx,[a,b] Guy Vandenbussche[c] and Remy Loris[a,b]*

[a]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium, [b]Department of Molecular and Cellular Interactions, VIB, Pleinlaan 2, B-1050 Brussel, Belgium, and [c]Structural Biology and Bioinformatics Center, Université Libre de Bruxelles, COP 206/2, Boulevard Du Triomphe, B-1050 Brussels, Belgium

Correspondence e-mail: reloris@vub.ac.be

Received 12 October 2009
Accepted 30 November 2009



The antitoxin Phd from the *phd*/*doc* module of bacteriophage P1 was crystallized in two distinct crystal forms. Crystals of His-tagged Phd contain a C-terminally truncated version of the protein and diffract to 2.20 Å resolution. Crystals of untagged Phd purified from the Phd–Doc complex diffract to 2.25 Å resolution. These crystals are partially merohedrally twinned and contain the full-length version of the protein.

1. Introduction

Prokaryotic toxin–antitoxin (TA) modules constitute small regulatory networks that play a role in the response to nutritional stress (Buts *et al.*, 2005; Engelberg-Kulka *et al.*, 2006; Gerdes *et al.*, 2005). TA operons encode two proteins termed the toxin and the antitoxin. The toxin component of the TA module interferes with a fundamental physiological process such as translation, transcription or replication. For example, the toxins CcdB and ParE act on gyrase (Bernard & Couturier, 1992; Jiang *et al.*, 2002), while RelE, YoeB, HigB and MazF cut mRNA in a codon-specific manner (Christensen & Gerdes, 2003; Christensen *et al.*, 2003; Christensen-Dalsgaard & Gerdes, 2006; Kamada & Hanaoka, 2005; Pedersen *et al.*, 2003; Zhang *et al.*, 2003).

The antitoxin proteins typically contain an N-terminal DNA-binding/dimerization domain followed by an intrinsically disordered toxin-binding domain. The antitoxin neutralizes, or even reverses, the action of the toxin by forming a noncovalent complex with this virulence factor. This binding also increases the affinity of the antitoxin for its operator DNA, resulting in tight regulation of expression of TA modules (for a review, see Buts *et al.*, 2005).

The *phd*/*doc* family of TA modules was first discovered as an addiction system in bacteriophage P1; it induces post-segregational killing of *Escherichia coli* if the P1 plasmid is lost during cell division (Lehnherr *et al.*, 1993). Members of this family of TA modules have since been identified in the genomes of both Gram-positive and Gram-negative bacteria. The 126-amino-acid toxin Doc is similar to Fic (Garcia-Pino, Christensen-Dalsgaard *et al.*, 2008), a protein that has been implicated in regulation of cell division in *E. coli* (Utsumi *et al.*, 1982). Fic domains are found in both prokaryotes and eukaryotes and have recently been shown to catalyze ATP-mediated AMPylation of a conserved tyrosine residue in the switch I region of Rho GTPases (Worby *et al.*, 2009). Doc itself contains a modified version of the Fic active-site motif and has been shown to inhibit translation through binding to ribosomes (Liu *et al.*, 2008).

The 73-amino-acid antitoxin Phd shows weak sequence similarity to the 92-amino-acid antitoxin YefM from the *yoeB/yefM* module (14% sequence identity; Kamada & Hanaoka, 2005). A peptide corresponding to its 22 C-terminal amino acids ($Phd^{52–73}$) was intrinsically disordered in solution but folded into a kinked α-helix when bound to Doc (Garcia-Pino, Christensen-Dalsgaard *et al.*, 2008). In this $Phd^{52–73}$–Doc complex structure, the peptide is bound to the Doc protein in a position that corresponds to the C-terminal α-helix of Fic (which Doc lacks). Therefore, it was suggested that Doc

might have originated from a Fic-like ancestor that transferred one of its α-helices to the C-terminus of a Phd-like DNA-binding domain.

Crystallization of full-length TA antitoxins in the absence of their cognate toxin is not simple. Most structures of antitoxins obtained to date have been determined by NMR spectroscopy in their free states (Madl *et al.*, 2006; Oberer *et al.*, 2007) or by crystallography as complexes with their toxin partners (Mattison *et al.*, 2006; Meinhart *et al.*, 2003; Miallau *et al.*, 2009; Schumacher *et al.*, 2009; Takagi *et al.*, 2005). The crystal structure of MazE, the first antitoxin to be crystallized in the absence of its cognate toxin, was determined employing a llama antibody VHH domain as a crystallization aid (Loris *et al.*, 2003). The only full-length antitoxin for which a crystal of its free state is available is a YefM-like protein from *Mycobacterium tuberculosis*. In this structure, the C-terminal region is involved in extensive crystal-packing interactions and folds into an α-helical conundrum that lacks biological relevance (Kumar *et al.*, 2008). Here, we report the crystallization of the Phd antitoxin (73 amino acids; 8133.10 Da) in the absence of it endogenous toxin partner Doc. The resulting crystal structures of Phd and of its N-terminal domain are likely to provide new information on the mechanism of action of Phd, in particular on the conformational states that this protein can adopt and on the conformational changes that occur between the Doc-bound and unbound conformations of the protein.

2. Materials and methods

2.1. Expression and purification of Phd in the absence of Doc

The *phd* gene (SwissProt entry Q06253 PHD_BPP1) was cloned in pET15b vector using the *Nde*I and *Bam*HI restriction sites and the plasmid was transformed into BL21 (DE3) cells. This strategy places an N-terminal His tag on the protein (tag sequence MGSSHHHHH-HSSGLVPRGSH, followed by the wild-type Phd sequence including its N-terminal methionine). 5 l LB medium (supplemented with 1% glucose and ampicillin) was inoculated with an overnight preculture. Cells were grown to an OD_{600} of 0.8–0.9 at 310 K and induced with 0.5 m*M* IPTG. For overexpression, the temperature was switched to 301 K. 3 h after induction, the culture was harvested by centrifugation (9000*g* for 20 min) and the cells were lysed with a cell cracker. The supernatant was loaded onto an Ni-Sepharose column, which was subsequently washed extensively with 50 m*M* Tris pH 7.5 to remove nonspecifically bound proteins. Phd was eluted with a gradient of imidazole (from 0 to 1.0 *M*). Fractions containing Phd were pooled together, concentrated and loaded onto a gel-filtration column equilibrated in 50 m*M* HEPES pH 7.2, 150 m*M* NaCl. The sample quality was verified by SDS–PAGE and CD spectroscopy.

2.2. Expression and purification of Phd from the Phd–Doc complex

The coding region of the *phd*/*doc* operon (SwissProt entries Q06253 PHD_BPP1 and Q06259 DOC_BPP1) was cloned in pET21b vector and the plasmid was transformed into BL21 (DE3) cells as described by Garcia-Pino, Dao-Thi *et al.* (2008). The cloning places a C-terminal His tag on Doc, while Phd retains its wild-type sequence. 10 l LB was inoculated with an overnight preculture selected with ampicillin. The culture was grown at 310 K to an OD_{600} of 0.8–0.9, induced with IPTG for 2 h and harvested by centrifugation (9000*g* for 20 min). The cells were lysed with a cell cracker and the supernatant was loaded onto an Ni-Sepharose column. After extensive washing of the column with 50 m*M* Tris pH 7.5, 150 m*M* NaCl, we separated Phd from Doc using a step gradient (0.0, 1.5 and 3.0 *M*) of guanidinium hydrochloride (GdHCl), with Phd eluting at both 1.5 and 3.0 *M* GdHCl. The Phd-containing fractions were pooled, diluted ten times in 50 m*M* Tris pH 7.0 and dialysed overnight against the same buffer to completely remove GdHCl. Sample quality was verified by SDS–PAGE, CD spectroscopy and mass spectrometry.

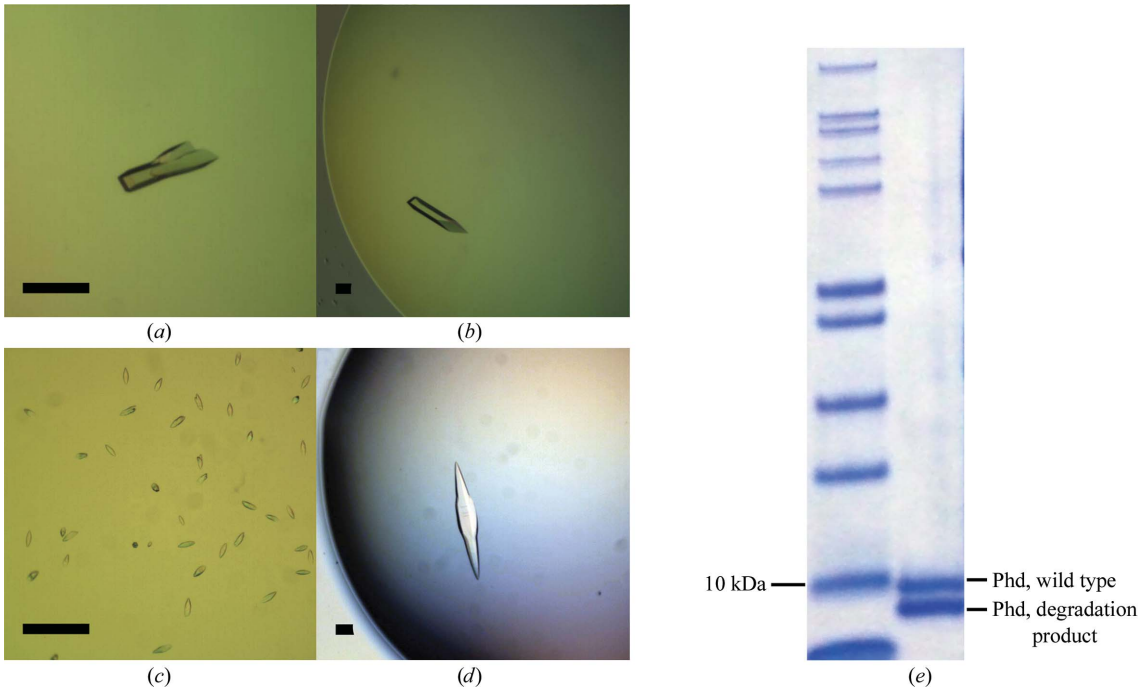


Figure 1
Different crystal forms of the free Phd. (*a*, *b*) Typical crystals of the N-terminal fragment of Phd (residues 1–58). (*c*) Initial crystals of full-length free Phd. (*d*) Large trigonal crystals of Phd obtained after optimization of the conditions used for the crystals in (*c*). The scale bars correspond to 0.1 mm. (*e*) SDS–PAGE of purified Phd after several days of incubation at 277 K, showing the appearance of a second band arising from proteolytic degradation of the wild-type protein.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Form I	Form II
Space group	$C222_1$	$P3_121$ or $P3_221$
Unit-cell parameters		
a (Å)	107.0	71.9
b (Å)	122.5	71.9
c (Å)	61.2	68.0
α (°)	90.0	90.0
β (°)	90.0	90.0
γ (°)	90.0	120.0
Resolution (Å)	34.0–2.19 (2.29–2.19)	12.0–2.23 (2.32–2.23)
Completeness (%)	99.1 (99.9)	98.3 (93.1)
No. of measured reflections	164406 (15896)	145806 (25106)
No. of unique reflections	20014 (1935)	9720 (1685)
Redundancy	8.2	15.0
$\langle I/\sigma(I)\rangle$	9.0 (6.9)	14.9 (6.7)
$R_{\rm merge}$†	0.110 (0.320)	0.074 (0.248)

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

2.3. Crystallization

Crystallization conditions were screened by vapour diffusion using the hanging-drop method. Hampton Research Crystal Screens I and II were used for initial screening. Hanging drops consisting of 2 µl protein solution and 2 µl precipitant solution were equilibrated against 400 µl precipitant solution. All crystallization trials were conducted at 293 K. The protein solution consisted of 10 mg ml^{-1} untagged Phd in 50 m*M* Tris pH 7.0 or 10 mg ml^{-1} His-tagged Phd in 50 m*M* HEPES pH 7.2, 150 m*M* NaCl. Optimization was achieved by subsequently varying the protein concentration as well as the precipitant concentration.

2.4. Data collection and analysis

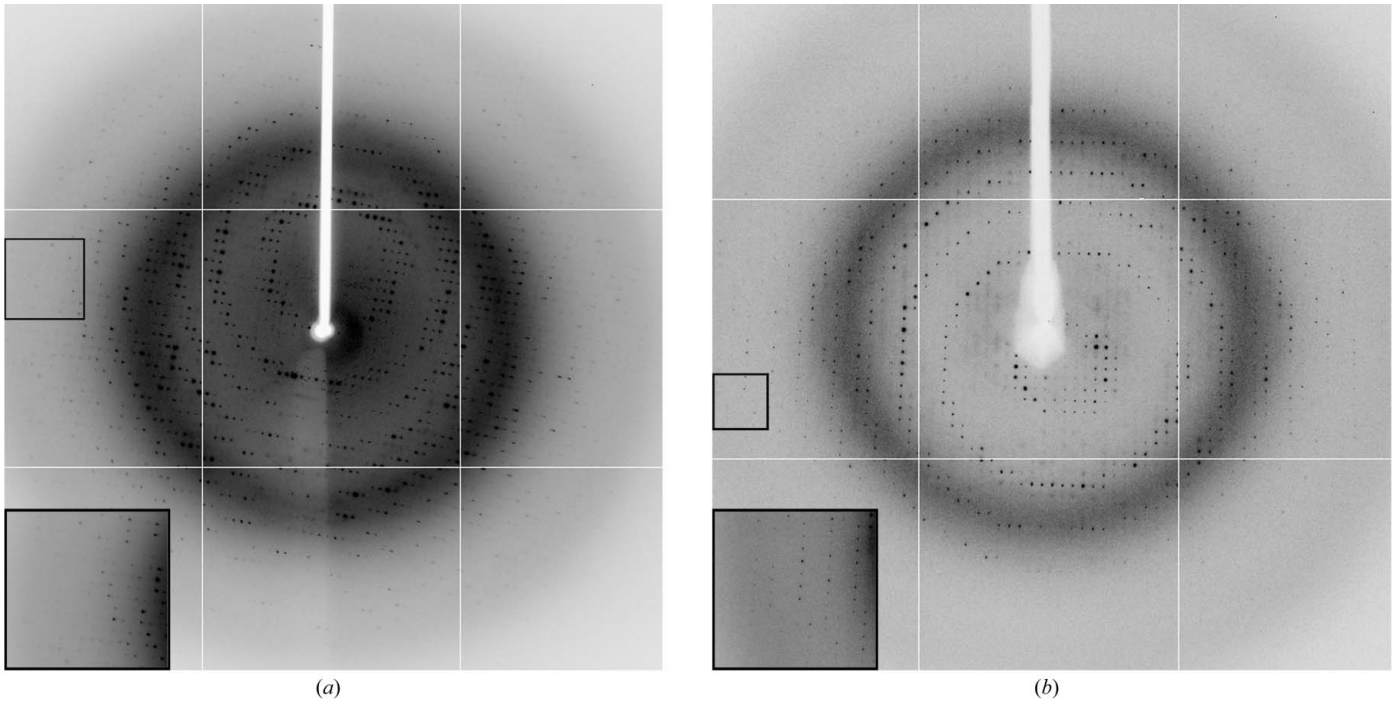
A crystal of form I was cryoprotected by transfer to a solution consisting of 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* magnesium acetate, 20%(*w*/*v*) PEG 8000 and 15%(*v*/*v*) glycerol and subsequently vitrified in liquid nitrogen. X-ray data were collected on beamline Proxima-1 of the SOLEIL synchrotron (Gif-Sur-Yvette, France) using an ADSC Quantum Q315 CCD detector and a wavelength of 0.98 Å.

A crystal of form II was directly vitrified in liquid nitrogen from the crystallization solution [0.1 *M* HEPES pH 7.5, 10%(*w*/*v*) PEG 8000, 25%(*v*/*v*) ethylene glycol] without the need for any additional cryoprotectant. X-ray data for form II were collected on beamline ID-29 of the ESRF synchrotron (Grenoble, France) using an ADSC Quantum 4 CCD detector and a wavelength of 1.07 Å.

All data were indexed and integrated using *DENZO* and subsequently scaled and merged using *SCALEPACK* (Otwinowski & Minor, 1997). Intensities were converted to structure-factor amplitudes using the *CCP*4 program *TRUNCATE* (Collaborative Computational Project, Number 4, 1994). Matthews coefficients were calculated with the program *MATTHEWS_COEF* for cell-content analysis (Collaborative Computational Project, Number 4, 1994). As considerable twinning was identified in the crystals of full-length Phd (crystal form II), the X-ray data of this crystal form were examined further using *phenix.xtriage* (Zwart *et al.*, 2005) from the *PHENIX* suite (Afonine *et al.*, 2005) to determine possible twin laws and twinning fractions.

2.5. Mass-spectrometric analysis

For the acquisition of mass-spectrometric data for Phd, samples were loaded into a nanoflow capillary electrospray ionization source (Proxeon, Odense, Denmark). ESI mass spectra were acquired on a quadrupole time-of-flight instrument (Q-Tof Ultima, Waters/Micromass) equipped with a Z-spray nanoelectrospray source and operating in the positive-ion mode. The spectra were recorded in the V mode and represent the combination of 1 s scans. The molecular mass of the protein was determined after processing the spectra with the software *MaxEnt*1 (Ferrige *et al.*, 1992).



(*a*) (*b*)

Figure 2
Typical diffraction patterns of (*a*) crystals of the N-terminal region of Phd (residues 1–58) and (*b*) full-length Phd.

3. Results and discussion

Initially, crystals were obtained of the His-tagged version of Phd that was overproduced in the absence of Doc (Figs. 1*a* and 1*b*). These crystals were grown from 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* magnesium acetate, 20%(*w*/*v*) PEG 8000 and did not require any further optimization. Mass-spectrometric analysis of the crystals showed that the protein was truncated at Ser58. This truncation is likely to be the result of uncontrolled proteolysis and has also been observed for other antitoxin proteins (REF). Fig. 1(*c*) shows the degradation pattern of His-tagged Phd that appeared after several days of storage of the purified protein at 277 K. Although this crystal form did not require further optimization, it was only obtained after a crystallization process lasting several months. This substantial length was likely to be related to degradation of the His-tagged Phd being necessary for crystallization. The crystals belong to space group $C222_1$, with unit-cell parameters a = 107.0, b = 122.5, c = 61.2 Å, and diffracted to 2.2 Å resolution (Fig. 2*a*, Table 1). Analysis of the unit-cell contents with the *CCP*4 program *MATTHEWS_COEF* suggested that the asymmetric unit contains between two (Matthews coefficient of 3.12 Å^3 Da^{-1}; 61% solvent) and four dimers (Matthews coefficient of 1.56 Å^3 Da^{-1}; 21% solvent). Analysis of the self-rotation function (Fig. 3*a*) showed a series of prominent nontrivial peaks on the κ = 180° section that agreed with the presence of two tetrameric entities harbouring internal 222 symmetry (likely to be a dimers of dimers, as the dimer is expected to be the biologically functional molecule). The κ = 120° and κ = 90° sections, on the other hand, did not show any significant peaks.

We also reasoned that it would be possible to obtain nontruncated and nontagged Phd by co-expressing it with Doc. The affinity between Phd and Doc is in the micromolar range (Gazit & Sauer, 1999) and it is possible to separate the proteins by treating the complex with guanidinium hydrochloride (A. Garcia-Pino & R. Loris, manuscript in preparation); the Phd–Doc complex dissociates at moderate GdHCl concentrations. Under those conditions Phd unfolds, but it can be fully refolded without any significant loss of protein. Indeed, crystals of refolded nontagged Phd were also obtained. They initially grew from 0.1 *M* HEPES pH 7.5, 10%(*w*/*v*) PEG 8000, 8%(*v*/*v*) ethylene glycol. The crystals which were obtained directly from the screen were rather small and were associated with excessive nucleation (Fig. 1*d*). Further improvement was achieved by reducing the protein concentration to 5 mg ml^{-1} and increasing the concentration of ethylene glycol to 15–20%. This resulted in reduced nucleation and larger crystals (Fig. 1*e*). Mass-spectrometric analysis of the protein present in the crystals (Fig. 4) revealed that Phd is an intact antitoxin; its experimental mass (8132.7 Da) was within 1 Da of the expected value. The crystals belong to space group $P3_121$ or $P3_221$, with unit-cell parameters a = b = 71.9, c = 68.0 Å, and

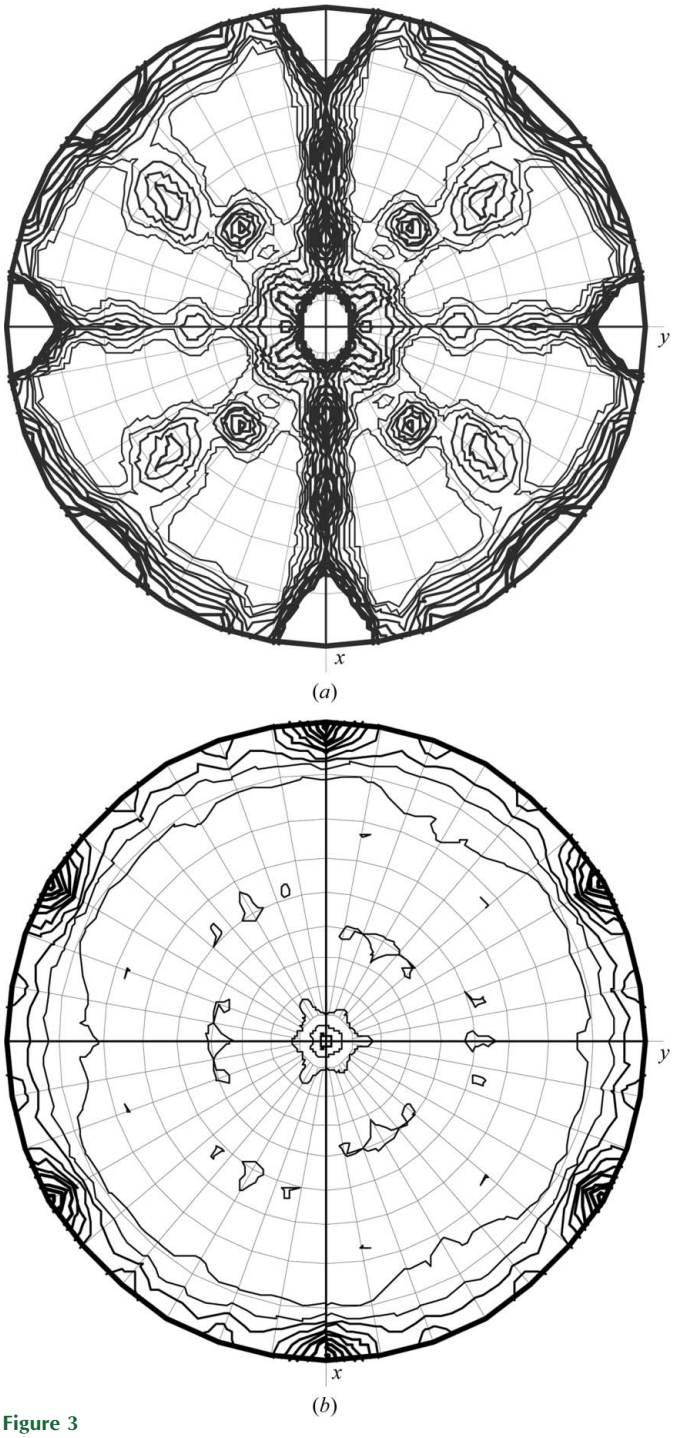


Figure 3
Self-rotation functions. (*a*) κ = 180° section of the self-rotation function calculated for crystal form I. (*b*) κ = 180° section of the self-rotation function calculated for crystal form II. Contour levels are drawn at 1σ intervals. All calculations were performed with *MOLREP* (Vagin & Teplyakov, 1997).

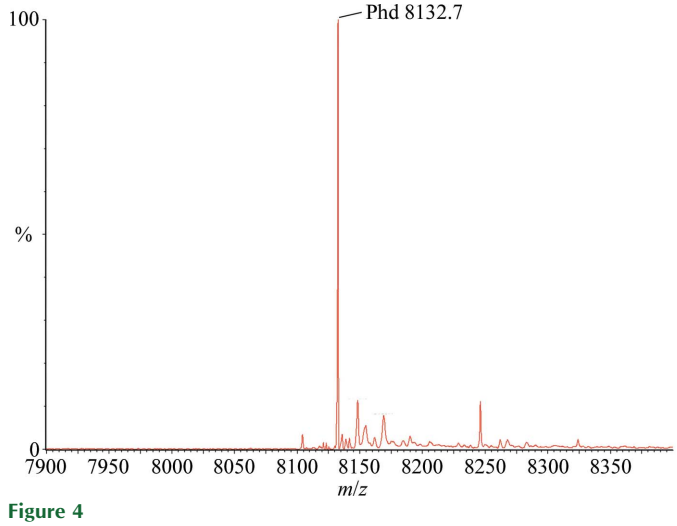


Figure 4
Mass-spectrometric analysis of Phd crystals of form II. The sample for mass spectrometry was obtained from a single crystal that was recovered from the crystallization drop, washed to remove contamination from Phd molecules still in solution and then diluted in 10 µl water.

diffraction data were collected to 2.25 Å resolution (Fig. 3*b*, Table 1). Analysis of the acentric moments of the intensity distributions indicated partial merohedral twinning. The twinning fraction was estimated to be 0.3–0.45 for the four crystals that were analyzed. The crystals are most likely to contain a single dimer of Phd in the asymmetric unit (Matthews coefficient of 3.12 $Å^3$ Da^{-1}, corresponding to a solvent fraction of 60.7%), although the presence of 1.5 dimers (one dimer formed by applying crystallographic symmetry) cannot be excluded (Matthews coefficient of 2.08 $Å^3$ Da^{-1}, corresponding to a solvent fraction of 41.1%). The self-rotation function is not informative in this respect as only trivial peaks corresponding to the crystallographic twofold axes were observed.

Several strategies are at our disposal for phasing and structure determination of Phd. The most straightforward would be molecular replacement using the N-terminal domain of Phd present in the crystals of the full-length Phd–Doc complex (Garcia-Pino, Dao-Thi *et al.*, 2008). We are currently attempting to determine the structure of this Phd–Doc complex by molecular replacement using the coordinates of the $Phd^{52–73}$–Doc complex (Garcia-Pino, Christensen-Dalsgaard *et al.*, 2008). In the case of failure, we will screen for suitable heavy-atom derivatives to obtain experimental phases. Phd does not contain any methionine residues and hence SAD or MAD phasing using selenomethionine would require the production and crystallization of a site-specific mutant.

This work was supported by research grants from OZR-VUB, VIB and FWO-Vlaanderen. YS acknowledges the receipt of an individual predoctoral fellowship from FWO-Vlaanderen. The authors acknowledge the use of the Proxima-1 beamline at SOLEIL (Gif-Sur-Yvette, France) and beamline ID-29 at the ESRF (Grenoble, France). AGP acknowledges the Belgian Federal Science Policy Office (DWTC) and the European Space Agency for financing in the context of Prodex project AO2004.

References

Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl.* **42**, 1–7.
Bernard, P. & Couturier, M. (1992). *J. Mol. Biol.* **226**, 735–745.
Buts, L., Lah, J., Dao-Thi, M.-H., Wyns, L. & Loris, R. (2005). *Trends Biochem. Sci.* **30**, 672–679.
Christensen, S. K. & Gerdes, K. (2003). *Mol. Microbiol.* **48**, 1389–1400.
Christensen, S. K., Pedersen, K., Hansen, F. G. & Gerdes, K. (2003). *J. Mol. Biol.* **332**, 809–819.
Christensen-Dalsgaard, M. & Gerdes, K. (2006). *Mol. Microbiol.* **62**, 397–411.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Engelberg-Kulka, H., Amitai, S., Kolodkin-Gal, I. & Hazan, R. (2006). *PLoS Genet.* **2**, e135.
Ferrige, A. G., Seddon, M. J., Skilling, J. & Ordsmith, N. (1992). *Rapid Commun. Mass Spectrom.* **6**, 765–770.
Garcia-Pino, A., Christensen-Dalsgaard, M., Wyns, L., Yarmolinsky, M., Magnuson, R. D., Gerdes, K. & Loris, R. (2008). *J. Biol. Chem.* **283**, 30821–30827.
Garcia-Pino, A., Dao-Thi, M.-H., Gazit, E., Magnuson, R. D., Wyns, L. & Loris, R. (2008). *Acta Cryst.* F**64**, 1034–1038.
Gazit, E. & Sauer, R. T. (1999). *J. Biol. Chem.* **274**, 16813–16818.
Gerdes, K., Christensen, S. K. & Lobner-Olesen, A. (2005). *Nature Rev. Microbiol.* **3**, 371–382.
Jiang, Y., Pogliano, J., Helinski, D. R. & Konieczny, I. (2002). *Mol. Microbiol.* **44**, 971–979.
Kamada, K. & Hanaoka, F. (2005). *Mol. Cell*, **19**, 497–509.
Kumar, P., Issac, B., Dodson, E. J., Turkenburg, J. P. & Mande, S. C. (2008). *J. Mol. Biol.* **383**, 482–493.
Lehnherr, H., Maguin, E., Jafri, S. & Yarmolinsky, M. B. (1993). *J. Mol. Biol.* **233**, 414–428.
Liu, M., Zhang, Y., Inouye, M. & Woychik, N. A. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 5885–5890.
Loris, R., Marianovsky, I., Lah, J., Laeremans, T., Engelberg-Kulka, H., Glaser, G., Muyldermans, S. & Wyns, L. (2003). *J. Biol. Chem.* **278**, 28252–28257.
Madl, T., Van Melderen, L., Mine, N., Respondek, M., Oberer, M., Keller, W., Khatai, L. & Zangger, K. (2006). *J. Mol. Biol.* **364**, 170–185.
Mattison, K., Wilbur, J. S., So, M. & Brennan, R. G. (2006). *J. Biol. Chem.* **281**, 37942–37951.
Meinhart, A., Alonso, J. C., Strater, N. & Saenger, W. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 1661–1666.
Miallau, L., Faller, M., Chiang, J., Arbing, M., Guo, F., Cascio, D. & Eisenberg, D. (2009). *J. Biol. Chem.* **284**, 276–283.
Oberer, M., Zangger, K., Gruber, K. & Keller, W. (2007). *Protein Sci.* **16**, 1676–1688.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K. & Ehrenberg, M. (2003). *Cell*, **112**, 131–140.
Schumacher, M. A., Piro, K. M., Xu, W., Hansen, S., Lewis, K. & Brennan, R. G. (2009). *Science*, **323**, 396–401.
Takagi, H., Kakuta, Y., Okada, T., Yao, M., Tanaka, I. & Kimura, M. (2005). *Nature Struct. Mol. Biol.* **12**, 327–331.
Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M. & Komano, T. (1982). *J. Bacteriol.* **151**, 807–812.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C. & Dixon, J. E. (2009). *Mol. Cell*, **34**, 93–103.
Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Qing, G. & Inouye, M. (2003). *Mol. Cell*, **12**, 913–923.
Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl.* **43**, contribution 7.