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

Expression, purification and preliminary X-ray diffraction studies of the transcriptional factor PyrR from *Bacillus halodurans*

Rodrigo Arreola,[a] Anita Vega-Miranda,[b] Armando Gómez-Puyou,[a] Ruy Pérez-Montfort,[a] Enrique Merino-Pérez[c] and Alfredo Torres-Larios[a]*

[a]Departamento de Bioquímica, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Apartado Postal 70-243, Mexico City 04510, Mexico, [b]Instituto Nacional de Enfermedades Respiratorias, Calzada de Tlalpan 4502, México DF 14080, Mexico, and [c]Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apartado Postal 510-3, Cuernavaca, Morelos, Mexico

Correspondence e-mail: torres@ifc.unam.mx

Received 7 April 2008
Accepted 17 June 2008

The PyrR transcriptional regulator is widely distributed in bacteria. This RNA-binding protein is involved in the control of genes involved in pyrimidine biosynthesis, in which uridyl and guanyl nucleotides function as effectors. Here, the crystallization and preliminary X-ray diffraction analysis of two crystal forms of *Bacillus halodurans* PyrR are reported. One of the forms belongs to the monoclinic space group $P2_1$ with unit-cell parameters a = 59.7, b = 87.4, c = 72.1 Å, β = 104.4°, while the other form belongs to the orthorhombic space group $P22_12_1$ with unit-cell parameters a = 72.7, b = 95.9, c = 177.1 Å. Preliminary X-ray diffraction data analysis and molecular-replacement solution revealed the presence of four and six monomers per asymmetric unit; a crystallographic tetramer is formed in both forms.

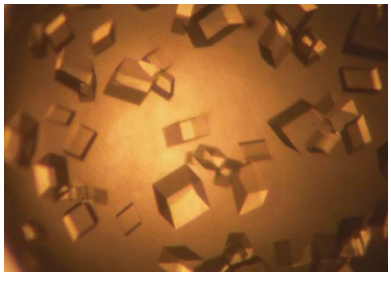
1. Introduction

Essential enzymes for *de novo* pyrimidine biosynthesis are encoded in many bacteria by an operon called *pyr*. Pyrimidine-regulatory protein (PyrR) regulates the expression of this operon at the transcriptional (reviewed in Switzer *et al.*, 1999) or translational level (Fields & Switzer, 2007) by interacting with a conserved sequence and structural motif on the mRNA located upstream of the operator (Bonner *et al.*, 2001). A repression mechanism results from the formation of a terminator hairpin or alternatively from the sequestration of the ribosome binding site. It has been proposed that binding of the protein to mRNA is regulated by the ratio of intracellular uridine to guanosine (Jorgensen *et al.*, 2008).

The PyrR-binding site has been found upstream of the *pyr* operon in 41 bacterial species and in six phyla, including pathogenic organisms such as *Mycobacterium tuberculosis* and *Staphylococcus aureus* (Griffiths-Jones *et al.*, 2005; Abreu-Goodger & Merino, 2005). This RNA site can be considered as a riboswitch-like element as recognition depends on the presence of structured RNA (Abreu-Goodger *et al.*, 2004).

The crystal structures of four PyrR proteins from different bacteria are known (Tomchick *et al.*, 1998; Chander *et al.*, 2005; Kantardjieff *et al.*, 2005; PDB code 1ufr deposited by RIKEN Structural Genomics/Proteomics Initiative). PyrR belongs to the PRT protein family, the members of which are involved in nucleotide synthesis and salvage (Sinha & Smith, 2001); in addition, PyrR proteins possess a highly basic cleft on one side which probably forms part of the mRNA-binding site (Tomchick *et al.*, 1998; Savacool & Switzer, 2002). The oligomeric organization of the protein has been found to involve crystallographic dimeric, tetrameric and hexameric oligomers. Although the biologically relevant oligomer has been suggested to be a dimer (Kantardjieff *et al.*, 2005), an equimolar binding ratio of protein:RNA also has been proposed (Bonner *et al.*, 2001).

The crystallization of protein–RNA complexes is still a challenging task. One strategy to overcome the crystallization problem is to try



different protein and RNA targets. As a step toward elucidating the detailed biochemical mechanisms of PyrR, including RNA-binding properties and complex formation with *pyr* mRNA, we cloned its gene from *Bacillus halodurans* and purified the protein. Crystals of PyrR protein were obtained by the sitting-drop vapour-diffusion method. The crystals diffracted well and data were collected to resolutions of 1.8 and 3.6 Å from crystals belonging to two different space groups.

2. Materials and methods

2.1. Cloning

The gene encoding *B. halodurans* PyrR (BhPyrR; gi:15615104) was amplified from genomic DNA of *B. halodurans* strain A-59, obtained from the American Type Culture Collection (ATCC), using the primers 5′-TACTTCCAATCCAATGCTTCGCGGTTAATTTTAG-ATGAGCAAC-3′ and 5′-TTATCCACTTCCAATGTTATGAGGT-TTTTTGTTCAATCGTTAC-3′. The insert was introduced into the pMCSG7 vector, a pET-based expression vector that has a His_6 tag at the N-terminus and a TEV protease site, by ligation-independent cloning (LIC) using published protocols (Stols *et al.*, 2002). Briefly, the vector was treated with *Ssp*I followed by T4 DNA polymerase in the presence of dGTP and the PCR product was treated with polymerase in the presence of dCTP. After annealing, freshly prepared competent *Escherichia coli* BL21 (DE3) pLysS cells were transformed and expression assays were performed with four colonies. Plasmids were isolated from cultures that showed a highly enriched band at the expected molecular weight of PyrR (20.2 kDa) and the insert was sequenced to verify the absence of mutations.

2.2. Expression and purification

E. coli BL21 (DE3) pLysS cells were grown at 310 K in LB medium containing ampicillin and chloramphenicol until an OD of 0.6 at 600 nm was reached. At this point, they were induced for 3 h with isopropyl β-D-1-thiogalactopyranoside. The cell pellet from 1 l culture was suspended in 15 ml buffer *A* (50 m*M* sodium phosphate buffer pH 8.0, 300 m*M* NaCl) supplemented with 10 m*M* imidazole. Cells were lysed by sonication and centrifuged at 20 000*g* for 30 min. The supernatant was loaded onto a column containing 10 ml Ni–NTA agarose resin (Qiagen). The protein was eluted with a linear gradient of buffer *A* supplemented with 500 m*M* imidazole and dialyzed for 2 h against a buffer containing 50 m*M* Tris pH 8.0, 0.5 m*M* EDTA and 1 m*M* dithiothreitol. The protein was then cleaved using purified recombinant His-tagged TEV protease expressed from the vector pRK508 (Kapust & Waugh, 1999) and purified to homogeneity. The protease was added at a ratio of 1 µg protease per 50 µg of BhPyrR and incubated at 303 K for 18 h. The mixture was then concentrated using Amicon Ultra filters (10 000 Da molecular-weight cutoff, Millipore) and dialyzed against buffer *A* supplemented with 10 m*M* imidazole. The His-tagged TEV protease was subsequently removed by batch treatment with 4 ml Ni–NTA agarose. The protein was precipitated with ammonium sulfate at 70% saturation. A hydrophobic interaction chromatography step was then carried out (Butyl-Toyopearl resin) in the presence of ammonium sulfate. After these purification steps, protein samples were >99% pure as judged by denaturing gel electrophoresis. Approximately 20 mg pure protein was obtained per litre of culture. The protein was stored for up to two weeks at 277 K in a buffer containing 20 m*M* HEPES pH 7.4.

2.3. Size-exclusion chromatography

Size-exclusion chromatography was performed using a Superdex 75 10/300GL analytical column (GE Healthcare) on an ÄKTA FPLC System (GE Healthcare). 300 µl protein solution at a concentration of 1 mg ml^{-1} (50 µ*M* monomer) was loaded onto the column, which was pre-equilibrated with 50 m*M* Tris pH 8.0 and 150 m*M* NaCl. The column was run at a flow rate of 0.5 ml min^{-1} and the absorbance at 280 nm was measured. The size-exclusion column was calibrated with a gel-filtration standard that contained the following globular protein markers (molecular weights and retention volumes are reported): γ-globulin (bovine; 158 kDa, 8.5 ml), ovalbumin (chicken; 43 kDa, 10.5 ml), myoglobin (horse; 17 kDa, 12.7 ml) and vitamin B_{12} (1.35 kDa, 18.2 ml).

2.4. Crystallization and data collection

Crystallization assays were carried out using a refrigerated incubator at 282 K. The protein was crystallized by the sitting-drop vapour-diffusion method using 96-well Corning crystallization plates with three drop positions. 1 µl protein solution containing 12 or

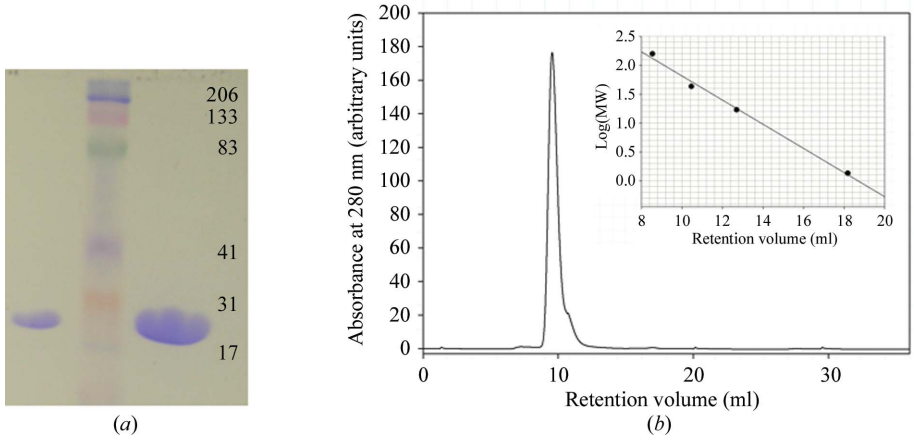
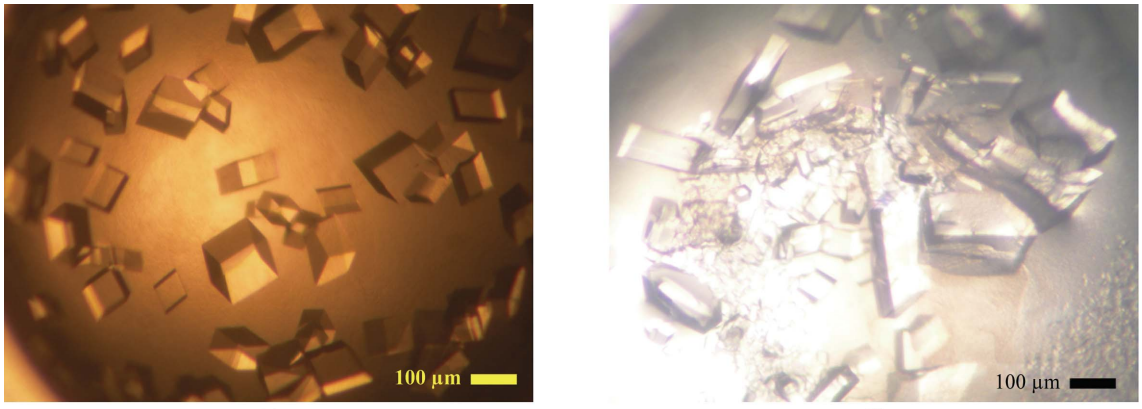


Figure 1
(*a*) SDS–PAGE analysis of purified BhPyrR. The first lane was loaded with 2 µg purified BhPyrR. 8 µg of the same protein was loaded in the third lane. Kaleidoscope prestained standards (BioRad) are shown in the middle lane. The corresponding molecular weight in kDa for each protein in the marker is shown on the right side of the figure. (*b*) Analytical gel filtration on a Superdex 75 10/300GL column of purified BhPyrR. 300 µg protein was loaded onto the column and eluted as described in §2. The protein eluted at 9.5 ml, which corresponds to an apparent molecular weight of ∼80 kDa (the molecular weight of the monomer is 20.2 kDa). The inset shows the relative retention volumes of the protein molecular-weight standards (MW in kDa; see §2).

25 mg ml^{-1} protein, 3 m*M* UTP and 5 m*M* $MgCl_2$ was mixed with 1 µl reservoir solution. Cube-shaped crystals were obtained using the Crystal Screen HT kit from Hampton Research. Further optimization yielded better shaped crystals after one or two weeks of incubation. The best single crystals were grown at 282 K; one crystal form (*A*) was obtained using 10 mg ml^{-1} protein solution and a reservoir solution consisting of 100 m*M* HEPES pH 7.5, 2% PEG 400 and 2 *M* ammonium sulfate and a second crystal form (*B*) was obtained with 100 m*M* MES pH 6.5, 200 m*M* ammonium sulfate and 30% PEG MME 5000. The crystals of form *A*, which had approximate dimensions of 0.15 × 0.15 × 0.15 mm, were cryoprotected by soaking them for 2 min in a solution containing 100 m*M* HEPES pH 7.5, 2% PEG 400 and 2 *M* lithium sulfate and frozen in liquid nitrogen. The crystals of form *B* were frozen directly in liquid nitrogen. Diffraction data were collected on the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory, Illinois, USA using a MAR Mosaic MX-300 detector. The data were processed with *MOSFLM* (Leslie, 1992) and reduced with *SCALA* (Collaborative Computational Project, Number 4, 1994). Molecular-replacement solutions were searched for using a monomer of *B. caldolyticus* PyrR protein (Chander *et al.*, 2005; PDB code 1non) as a model with the program *Phaser* (McCoy *et al.*, 2007).

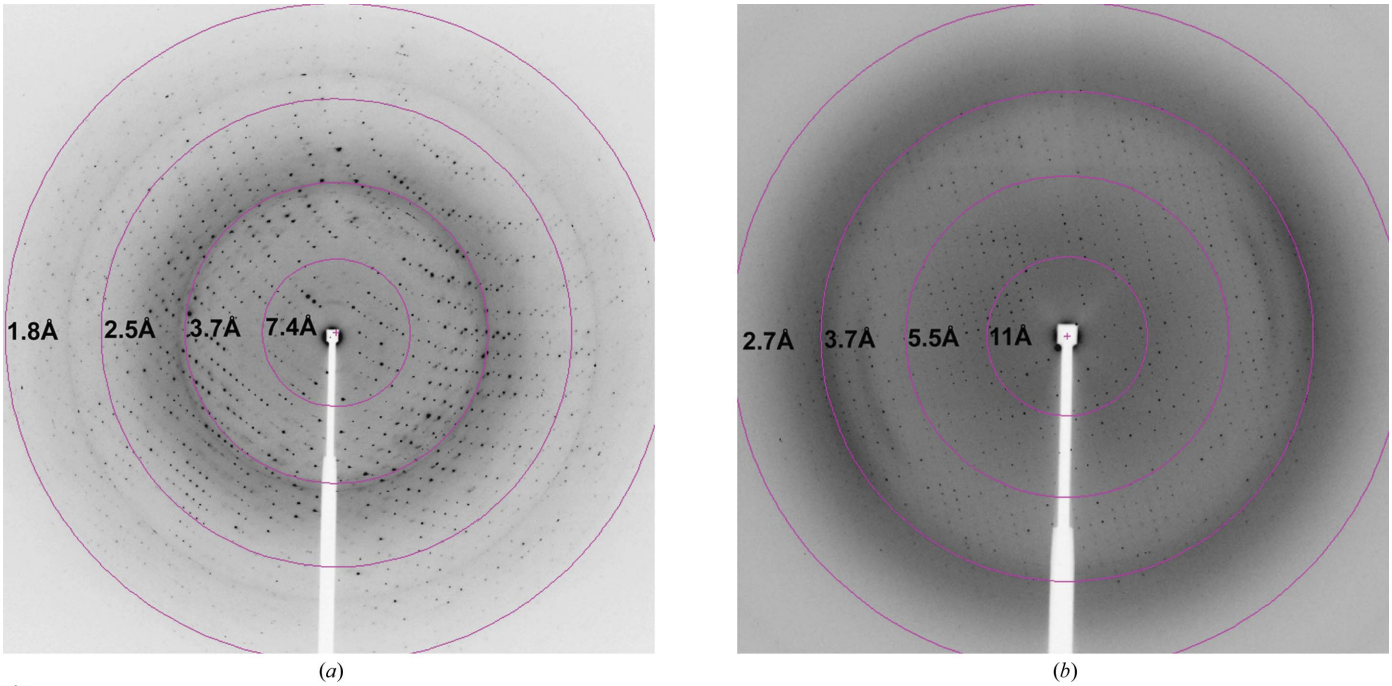
3. Results and discussion

BhPyrR protein was cloned into a LIC vector, expressed in *E. coli* BL21 (DE3) pLysS in a soluble form and purified to homogeneity (Fig. 1*a*) in three steps. The hydrophobic chromatography step was performed in order to prevent contamination with RNases and thus circumvent problems in preparing cocrystallization experiments with



(*a*) (*b*)

Figure 2
Crystals of BhPyrR protein grown using the sitting-drop method. (*a*) Crystals of form *A* were grown at 282 K with a reservoir solution consisting of 100 m*M* HEPES pH 7.5, 2% PEG 400 and 2 *M* ammonium sulfate. The dimensions of the crystal used for data collection were approximately 0.15 × 0.15 × 0.15 mm. (*b*) Crystals of form *B* were grown at 282 K with a reservoir solution consisting of 100 m*M* MES pH 6.5, 200 m*M* ammonium sulfate and 30% PEG MME 5000. The dimensions of the crystal used for data collection were approximately 0.2 × 0.05 × 0.05 mm.



(*a*) (*b*)

Figure 3
X-ray diffraction patterns recorded from the two crystal forms of BhPyrR. (*a*) Form *A*. (*b*) Form *B*.

Table 1
Data-collection statistics.

Values in parentheses are for the outer shell.

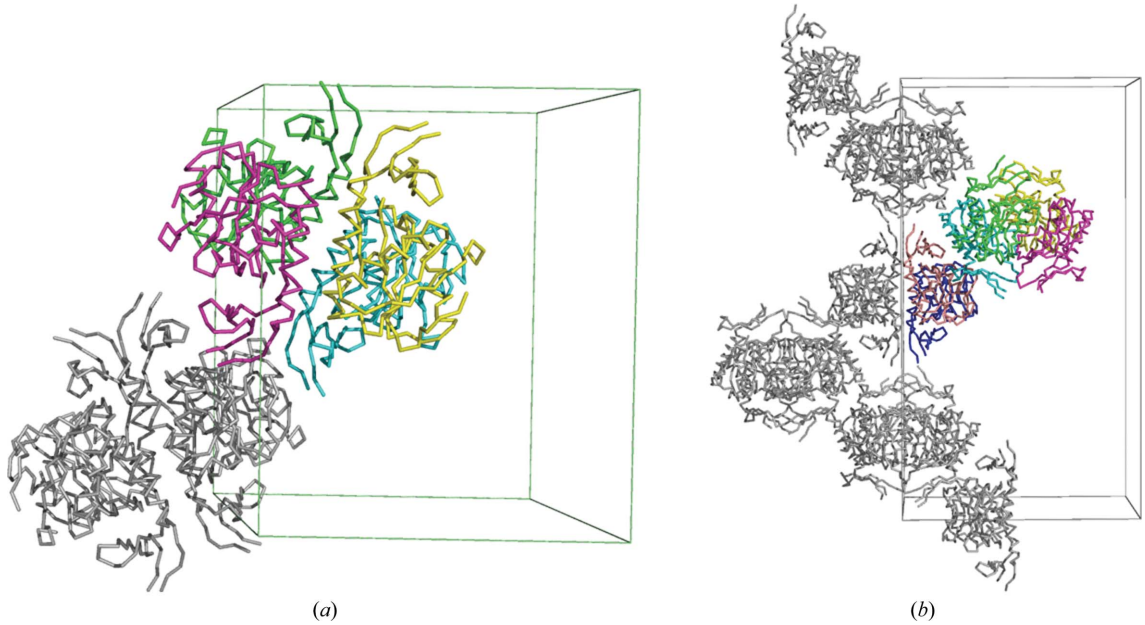
	Form *A*	Form *B*
Wavelength (Å)	1.00	1.00
Temperature (K)	100	100
Space group	$P2_1$	$P22_12_1$
Unit-cell parameters		
a (Å)	59.7	72.7
b (Å)	87.4	95.9
c (Å)	72.1	177.1
α (°)	90	90
β (°)	104.4	90
γ (°)	90	90
Matthews coefficient V_M (Å^3 Da^{-1})	2.3	2.6
Solvent content (%)	46	52
No. of molecules in ASU	4	6
Resolution range (Å)	69.8–1.8 (1.9–1.8)	69.8–3.6 (3.8–3.6)
Observed reflections	230381 (26453)	40547 (5433)
Unique reflections	64856 (8233)	13520 (1897)
Redundancy	3.5 (3.0)	3.0 (2.9)
Completeness (%)	98.8 (92.0)	92.2 (90.0)
R_{merge}† (%)	6.6 (29.9)	11.5 (34.5)
Average $I/\sigma(I)$	14.6 (3.0)	9.8 (3.1)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of a reflection and $\langle I(hkl)\rangle$ is the mean intensity of reflection *hkl*.

RNA (Torres-Larios *et al.*, 2002). Gel-filtration chromatography was performed to analyze the oligomerization state of the protein in solution. The molecular weight of PyrR was estimated to be ∼80 kDa, suggestive of a tetrameric state (Fig. 1*b*). The Hampton Research Crystal Screen HT kit was used for preliminary screening and produced crystals under several conditions. After refinement of the crystallization conditions, we obtained good-looking single crystals from two conditions after two weeks of growth at 282 K using 10 mg ml^{-1} protein; the first condition mainly used ammonium sulfate as a precipitant (crystal form *A*, obtained using 100 m*M* HEPES pH 7.5, 2% PEG 400 and 2 *M* ammonium sulfate; Fig. 2*a*); in the second condition PEG MME 5000 was employed (crystal form *B*, obtained using 100 m*M* MES pH 6.5, 200 m*M* ammonium sulfate and 30% PEG MME 5000; Fig. 2*b*). Both types of crystals were suitable for diffraction experiments. Crystal form *A* showed significant diffraction to 1.8 Å resolution, while crystal form *B* diffracted to 3.6 Å resolution (Fig. 3). The data-collection statistics are listed in Table 1.

Analysis of the Matthews coefficients for the two crystal forms revealed the possible presence of four monomers in the asymmetric unit of the primitive monoclinic crystal form *A* and six monomers in the primitive orthorhombic crystal form *B*. One monomer of *B. caldolyticus* PyrR protein (Chander *et al.*, 2005; PDB code 1non) was used as a search model to solve the phase problem by molecular replacement using the program *Phaser* (McCoy *et al.*, 2007). For crystal form *A*, four copies could be located in the asymmetric unit with no clashes and high *Z* scores for the translation function (TFZ = 46) and a high value for the log-likelihood gain (LLG = 4328). The four monomers in the asymmetric unit form a tetramer (Fig. 4*a*). For crystal form *B*, four copies could initially be located (TFZ = 25, LLG = 745). Two additional monomers could be located by performing an additional search using the previously found monomers. The scores that were obtained for the final solution with six monomers were quite high (TFZ = 32, LLG = 1883). The ambiguity between the enantiomorphic orthorhombic space groups was solved by *Phaser*, favouring space group $P22_12_1$, which presented no clashes in crystal packing. The final solution contained a tetramer and a dimer in the asymmetric unit; the dimer formed a crystallographic tetramer with a twofold axis on *c* (Fig. 4*b*). The tetrameric organization of both crystal forms is very similar to that found in the crystal structures of the *B. caldolyticus* (Chander *et al.*, 2005) and *M. tuberculosis* (Kantardjieff *et al.*, 2005) PyrR proteins. The packing of the two crystal forms, together with the results of size-exclusion chromatography, suggests that the quaternary organization of *B. halodurans* PyrR corresponds to a tetramer in the absence of RNA. Crystallographic refinement is currently in progress.

Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source



(*a*) (*b*)

Figure 4
Packing of the two crystal forms of BhPyrR. The monomers that are part of the asymmetric unit are shown in colour. (*a*) Form *A*, primitive monoclinic unit cell. (*b*) Form *B*, primitive orthorhombic unit cell.

was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACyT) grant 60127-Q and Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México (PAPIIT) grant IN200507.

References

Abreu-Goodger, C. & Merino, E. (2005). *Nucleic Acids Res.* **33**, W690–W692.
Abreu-Goodger, C., Ontiveros-Palacios, N., Ciria, R. & Merino, E. (2004). *Trends Genet.* **20**, 475–479.
Bonner, E. R., D'Elia, J. N., Billips, B. K. & Switzer, R. L. (2001). *Nucleic Acids Res.* **29**, 4851–4865.
Chander, P., Halbig, K. M., Miller, J. K., Fields, C. J., Bonner, H. K., Grabner, G. K., Switzer, R. L. & Smith, J. L. (2005). *J. Bacteriol.* **187**, 1773–1782.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fields, C. J. & Switzer, R. L. (2007). *J. Bacteriol.* **189**, 6236–6245.
Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S. R. & Bateman, A. (2005). *Nucleic Acids Res.* **33**, D121–D124.
Jorgensen, C. M., Fields, C. J., Chander, P., Watt, D., Burgner, J. W. II, Smith, J. L. & Switzer, R. L. (2008). *FEBS J.* **275**, 655–670.
Kantardjieff, K. A., Vasquez, C., Castro, P., Warfel, N. M., Rho, B.-S., Lekin, T., Kim, C.-Y., Segelke, B. W., Terwilliger, T. C. & Rupp, B. (2005). *Acta Cryst.* D**61**, 355–364.
Kapust, R. B. & Waugh, D. S. (1999). *Protein Sci.* **8**, 1668–1674.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Savacool, H. K. & Switzer, R. L. (2002). *J. Bacteriol.* **184**, 2521–2528.
Sinha, S. C. & Smith, J. L. (2001). *Curr. Opin. Struct. Biol.* **11**, 733–739.
Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R. & Donnelly, M. I. (2002). *Protein Expr. Purif.* **25**, 8–15.
Switzer, R. L., Turner, R. J. & Lu, Y. (1999). *Prog. Nucleic Acid Res. Mol. Biol.* **62**, 329–367.
Tomchick, D. R., Turner, R. J., Switzer, R. L. & Smith, J. L. (1998). *Structure*, **6**, 337–350.
Torres-Larios, A., Dock-Bregeon, A. C., Romby, P., Rees, B., Sankaranarayanan, R., Caillet, J., Springer, M., Ehresmann, C., Ehresmann, B. & Moras, D. (2002). *Nature Struct. Biol.* **9**, 343–347.