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

Crystallization and preliminary X-ray diffraction analysis of inositol 1,3,4,5,6-pentakisphosphate kinase from *Arabidopsis thaliana*

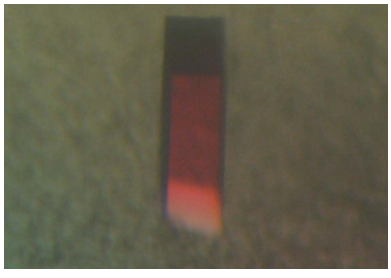
Jose Ignacio Baños-Sanz,[a]‡ Maider Villate,[a]‡ Julia Sanz-Aparicio,[a] Charles Alistair Brearley[b] and Beatriz González[a]*

[a]Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física 'Rocasolano', CSIC, Serrano 119, 28006 Madrid, Spain, and [b]School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, England

‡ These authors contributed equally to this work as first authors.

Correspondence e-mail: xbeatriz@iqfr.csic.es

Received 2 October 2009
Accepted 26 November 2009



Inositol 1,3,4,5,6-pentakisphosphate kinase (IP_5 2-K) is an enzyme involved in inositol metabolism that synthesizes IP_6 (inositol 1,2,3,4,5,6-hexakisphosphate) from inositol 1,3,4,5,6-pentakisphosphate (IP_5) and ATP. IP_6 is the major phosphorus reserve in plants, while in mammals it is involved in multiple cellular events such as DNA editing and chromatin remodelling. In addition, IP_6 is the precursor of other highly phosphorylated inositols which also play highly relevant roles. IP_5 2-K is the only enzyme that phosphorylates the 2-OH axial position of the inositide and understanding its molecular mechanism of substrate specificity is of great interest in cell biology. IP_5 2-K from *Arabidopsis thaliana* has been expressed in *Escherichia coli* as two different fusion proteins and purified. Both protein preparations yielded crystals of different quality, always in the presence of IP_6. The best crystals obtained for X-ray crystallographic analysis belonged to space group $P2_12_12_1$, with unit-cell parameters a = 58.124, b = 113.591, c = 142.478 Å. Several diffraction data sets were collected for the native enzyme and two heavy-atom derivatives using a synchrotron source.

1. Introduction

Inositol phosphates (InsPs) are molecules that play important biological roles in various key cell metabolic and signalling processes. Their levels are strictly regulated by various inositol kinases (InsP Ks) and phosphatases. In the latter steps of InsP metabolism, highly phosphorylated InsPs such as inositol 1,2,3,4,5,6-hexakisphosphate (IP_6; also known as phytic acid) are formed. IP_6 constitutes a major phosphorus reserve in plants and plays an essential role in processes as diverse as mRNA export (York *et al.*, 1999), DNA repair (Hanakahi & West, 2002), maintenance of basal resistance to plant pathogens (Murphy *et al.*, 2008), apoptosis (Agarwal *et al.*, 2009) and regulation of chromatin structure (Shen *et al.*, 2003; Steger *et al.*, 2003). In addition, IP_6 is the precursor of pyrophosphate inositols IP_7 and IP_8 (Mulugu *et al.*, 2007). Finally, several experiments have shown that IP_6 is a potential antitumour agent (Diallo *et al.*, 2008; Bozsik *et al.*, 2007).

The family of enzymes that catalyze the phosphorylation of inositol 1,3,4,5,6-pentakisphosphate (IP_5) to form IP_6 are known as IP_5 2-kinases (IP_5 2-Ks). IP_5 2-Ks have been cloned and characterized from several sources (Verbsky *et al.*, 2002; Phillippy *et al.*, 1994; Ives *et al.*, 2000; Sweetman *et al.*, 2006; Sun *et al.*, 2007; Suzuki *et al.*, 2007). Disruption of the IP_5 2-K gene (*ipk1*) yields nonviable murine embryos (Verbsky *et al.*, 2005). IP_5 2-K is also involved in the establishment of left–right organ asymmetry in mammals (Sarmah *et al.*, 2005). These enzymes present low sequence identity from yeast to mammals (about 11%) and do not present sequence homology to any other family of enzymes, including other families of inositol kinases. IP_5 2-K is the only InsP K that phosphorylates the 2-axial position of the *myo*-inositide ring, whereas the other enzymes act on equatorial positions. For this reason, IP_5 2-K represents an intriguing key point in cell biology, the understanding of which is crucial to obtaining insights into the regulation of inositol metabolism.

To date, the structures of several inositol kinases involved in soluble inositol regulation have been solved (González *et al.*, 2004;

Miller & Hurley, 2004; Holmes & Jogl, 2006; Miller *et al.*, 2005). However, the structure of the IP_5 2-K family remains unknown. Structural analysis of IP_5 2-K will yield highly valuable information to fully understand the molecular mechanism of the substrate specificity of this peculiar family of InsP Ks that are essential for the functioning of mammals. In this study, we describe the expression, purification, crystallization and preliminary X-ray crystallographic analysis of IP_5 2-K from *Arabidopsis thaliana*.

2. Experimental

2.1. Expression in *Escherichia coli*

The *ipk1* gene encoding the IP_5 2-K protein from *A. thaliana* was amplified from At5G 42810 cDNA (Sweetman *et al.*, 2006) by polymerase chain reaction. The PCR products were cloned into two different vectors encoding different fusion proteins. The forward primer 5′-C**GAGCTC**ATGGAGATGATTTTGGAGG-3′ and the reverse primer 5′-CCG**CTCGAG**TTAGCTGTGGGAAGGTTTTG-AGTTGC-3′ were used to clone the PCR product into pKLSL$_t$ vector (Mancheño Gómez & Angulo Herrera, 2009). The forward primer 5′-GGAGATATA**CATATG**GAGATGATTTTGGAGGAG-3′ and the reverse primer 5′-GG**GGTACC**TTAGCTGTGGGAAGGTTTTGA-GTTGC-3′ were used to clone the PCR product into pOPTG vector. The primers include restriction sites for *Sac*I, *Xho*I, *Nde*I and *Kpn*I (shown in bold), respectively, for cloning the PCR product into the above vectors. The resulting proteins are fused with glutatione *S*-transferase (GST) and *Laetiporus sulphureus* lectin (LSL$_t$) tags, respectively, and were expressed in *E. coli* Rosetta (DE3) pLysS strain. The bacteria were grown at 310 K until they reached an OD_{600} of 1 and expression was induced by adding 0.3 m*M* IPTG and shaking for 15 h at 289 K.

2.2. Purification of the fusion proteins

For the GST-fusion protein, the cells were resuspended in buffer *A* (20 m*M* Tris–HCl pH 8, 50 m*M* NaCl, 2 m*M* DTT) plus protease inhibitors (complete EDTA-free tablets from Roche) and disrupted with a French press. After addition of 0.1% Triton X-100, the lysate was applied onto a heparin column, washed with buffer *A* and eluted with a 0–1 *M* NaCl gradient. The protein was loaded onto a GST-trap HP column (GE Healthcare) and washed with buffer *B* (20 m*M* Tris–HCl pH 8, 200 m*M* NaCl, 2 m*M* DTT). The GST tag was removed by on-column incubation with TEV protease at 277 K overnight (protease:protein mass ratio 1:80). Finally, the IP_5 2-K cleaved from GST (sample I) was eluted, concentrated and applied onto a gel-filtration column (16/60 Superdex 200) equilibrated in buffer *B*.

For the LSL$_t$-fusion protein, the bacteria were lysed in buffer *C* (20 m*M* Tris–HCl pH 8, 150 m*M* NaCl, 2 m*M* DTT). The lysate was applied onto a Sepharose CL-4B column washed with buffer *C* and the protein was eluted with buffer *C* plus 200 m*M* lactose. The sample was diluted threefold to reduce the salt concentration, loaded onto a heparin column washed with buffer *A* and eluted with a 0–1 *M* NaCl gradient. After elution from the heparin column, the fused protein was cleaved by TEV protease (protease:protein mass ratio 1:80), shaking the sample gently at 277 K overnight, and IP_5 2-K (sample II) was further purified and separated from LSL$_t$ by gel filtration using similar conditions as for sample I. The purified fraction was finally passed through a small Sepharose CL-4B column equilibrated in buffer *B* to remove the residual fused protein.

Both of the recombinant enzymes, which will be referred to as sample I and sample II, respectively, produced a well isolated peak in the gel-filtration chromatogram that could correspond to a monomer. The final yields of the purifications were 1 and 10 mg per litre of culture, respectively. The purity of the samples was confirmed by SDS–PAGE (Fig. 1). Both samples presented IP_5 2-K activity, which was measured as described previously (Sweetman *et al.*, 2006). The final pure proteins were concentrated to around 10 mg ml^{-1} and stored at 193 K.

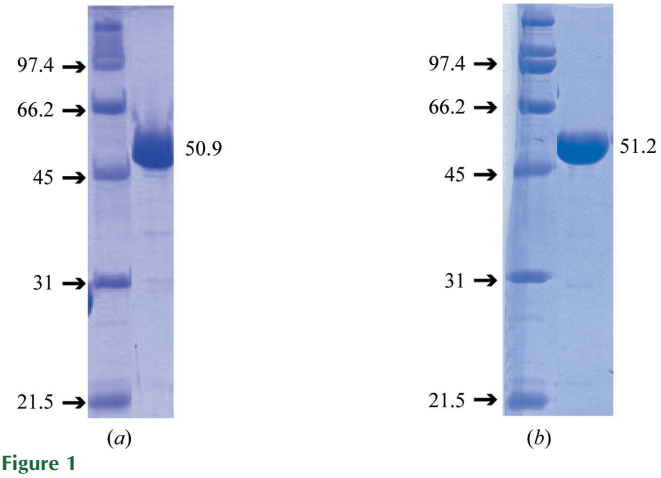


Figure 1
SDS–PAGE analysis of IP_5 2-K from *A. thaliana* for (*a*) sample I obtained from the GST-fusion protein and (*b*) sample II obtained from the LSL$_t$-fusion protein. Both gels show the positions of the molecular-mass markers used; their molecular masses are indicated in kDa. Sample I and sample II have the same sequence except for the first N-terminal residues (GSH in sample I and GEFEL in sample II).

2.3. Selenomethionine-substituted protein expression and purification

LSL$_t$-IP_5 2-K was expressed in minimal medium using the method reported by van Duyne *et al.* (1993). The selenomethionine-derivative protein was purified in a similar way to the native protein. The yield of the purified protein decreased dramatically in this protein form to a value of 0.6 mg per litre of culture.

2.4. Crystallization and cryoprotection

Initial crystallization conditions for protein samples I and II were investigated by high-throughput techniques with a NanoDrop robot (Innovadyne Technologies Inc.) using the commercial screen solutions Index and SaltRx from Hampton Research and PACT Suite and JCSG+ Suite from Qiagen. Duplicate crystallization assays were carried out for each protein sample with and without IP_6 using the sitting-drop vapour-diffusion method at 291 K in 96-well plates (Innovaplate SD-2 microplates, Innovadyne Technologies Inc.). In this method, drops consisting of 250 nl protein and 250 nl precipitant solution were mixed and equilibrated against 60 µl well solution. Many solutions containing PEG 3350 and PEG 6000 yielded crystals, but only in the assays using protein previously mixed with 2 m*M* IP_6. Several strategies were used to optimize the crystallization conditions, which included adjusting the precipitant concentration and pH values, screening different additives (Additive Screen, Hampton Research) and streak-seeding. The final conditions were scaled up in 24-well plates (Cryschem plates, Hampton Research) using sitting-drop experiments by mixing 1 µl protein solution with 1 µl precipitant solution and equilibrating against 500 µl well solution. The optimization was approached independently for each protein form, including the selenomethionine-substituted protein. The best crystals were obtained from sample II mixed with 2 m*M* IP_6 and 2 m*M* ADP at 22% PEG 3350 and 0.1 *M* bis-tris pH 5.9. Several cryoprotectants

were tested, including glycerol; the most successful cryoprotection resulted from gradually increasing the PEG content to 35% of the crystal mother-liquor solution before flash-cooling to 100 K.

2.5. Formation of heavy-atom derivatives

The best crystals (obtained in the presence of ADP and IP_6) were subjected to soaks in various heavy-atom solutions. The heavy-atom compounds used were lead acetate and trimethyllead acetate (TMLA). For the soaks, these compounds were included in the cryoprotectant solution together with 2 m*M* IP_6 and 2 m*M* ADP and the crystals were left from 1 to 30 min in the heavy-atom high-PEG solution. The final soaking times and heavy-atom concentrations of the crystals used for X-ray diffraction analysis were 25 min and 100 m*M* in both cases.

2.6. Data collection and processing

Crystals were tested using in-house and synchrotron-radiation facilities. Native and MAD data sets were collected on various ESRF beamlines (Grenoble, France). The data sets were processed using the program *MOSFLM* (Leslie, 1990) and the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Heavy-atom positions were located using *SHELX* (Sheldrick, 2008) and were refined using *autoSHARP* (Vonrhein *et al.*, 2007). Crystal parameters and diffraction statistics are shown in Table 1.

3. Results and discussion

IP_5 2-K from *A. thaliana* has been expressed as two different fusion proteins that allowed protein purification to apparent homogeneity. Interestingly, a good method for producing a considerable amount of IP_5 2-K has been obtained (the method used to produce sample II).

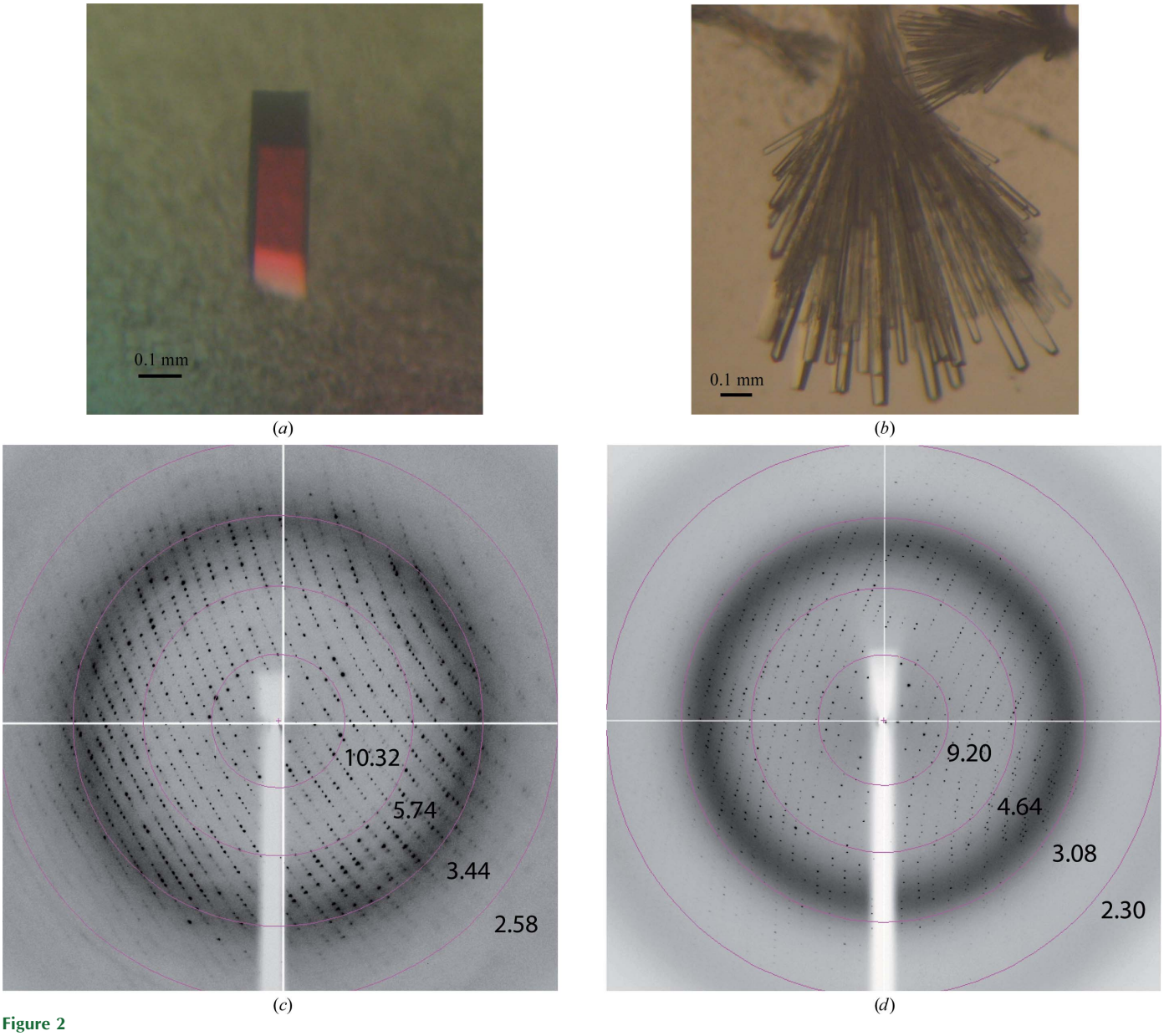


Figure 2
Rod-shaped crystals of recombinant IP_5 2-K from (*a*) sample I (0.25 × 0.11 × 0.11 mm) and (*b*) sample II (0.4 × 0.05 × 0.03 mm). (*c*, *d*) X-ray diffraction patterns of the crystals in (*a*) and (*b*) obtained using a synchrotron source. Numbers show the resolutions (Å) of the rings.

Table 1
Data-collection statistics for the IP_5 2-K crystals.

Values in parentheses are for the highest resolution shell.

	Native				Lead (peak)	
Data set	Sample I + IP_6	Sample II + IP_6	Sample II + IP_6	Sample II + IP_6, ADP	Sample II + IP_6, ADP	Sample II + IP_6, ADP
Cryoprotectant†	Glycerol	Glycerol	PEG 3350	PEG 3350	PEG 3350 + TMLA	PEG 3350 + lead acetate
Wavelength (Å)	0.93300	0.93300	0.93430	0.93340	0.94845	0.94780
Source	ESRF	ESRF	ESRF	ESRF	ESRF	ESRF
Beamline	ID14-2	ID14-2	ID14-1	ID14-1	ID23-1	ID23-1
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)						
a	60.172	59.538	57.669	58.124	57.046	58.063
b	115.067	117.523	112.501	113.591	112.583	110.970
c	236.254	150.446	139.720	142.478	141.081	138.694
Resolution (Å)	64.55–2.90 (3.06–2.90)	74.53–3.20 (3.37–3.20)	112.42–3.10 (3.27–3.10)	71.25–2.30 (2.42–2.30)	112.51–3.20 (3.37–3.20)	138.27–3.20 (3.37–3.20)
Unique reflections	37268 (5319)	18140 (2597)	17181 (2436)	42769 (6186)	15659 (2227)	15449 (2201)
R_{merge}‡	14.2 (42.5)	15.6 (47.1)	12.5 (44.5)	12.9 (45.6)	17.1 (40.9)	16.9 (44.2)
Completeness (%)	99.8 (99.8)	100.0 (100.0)	100.0 (100.0)	99.9 (100.0)	100.0 (100.0)	100.0 (100.0)
Mean multiplicity	5.3 (4.9)	5.7 (5.9)	7.6 (7.9)	7.3 (7.3)	7.7 (7.8)	6.7 (6.7)
Mean $I/\sigma(I)$	11.1 (2.1)	12.7 (2.9)	12.0 (4.4)	11.3 (4.2)	11.2 (5.9)	12.1 (4.0)
Wilson *B* factor (Å^2)	80.7	72.0	67.1	36.7	43.2	47.7

† The cryoprotectant includes the same products as were added to the sample. ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observed amplitude of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean amplitude for measurements of reflection *hkl*.

Preliminary crystallization trials for both recombinant samples led to crystal clusters of thin plates or needles using PEG 3350 or PEG 6000 as precipitant. After refinement of several parameters, including a very fine tuning of the pH value, we obtained clusters of thicker plates or large rods; only a few assays with sample I led to isolated thick rods (Fig. 2). The best plate-shaped crystals from both samples appeared using 20% PEG 3350 and 0.1 *M* bis-tris propane pH 6.5. However, rods appeared at 17.5% PEG 3350 and 0.1 *M* sodium citrate pH 5.2 for sample I and 22% PEG 3350 and 0.1 *M* bis-tris pH 5.9 for sample II. The different protein-purification protocols and/or the remaining N-terminal residues after tag removal could explain the different crystal properties of the two samples. We focused on the crystals that appeared as clusters of rods, since the X-ray diffraction data from the plate-shaped crystals could not be indexed. We initially used glycerol as a cryoprotectant and found that the sample I rods diffracted to 2.9 Å resolution (space group $P2_12_12_1$, unit-cell parameters $a = 60.175$, $b = 115.067$, $c = 236.254$ Å; Table 1), while the thin sample II rods diffracted to 3.2 Å resolution (space group $P2_12_12_1$, unit-cell parameters $a = 59.538$, $b = 117.523$, $c = 150.448$ Å; Table 1). Only the crystals from sample II were further optimized, as an analysis suggested that the sample I crystals were probably twinned. Cryoprotection was optimized by increasing the PEG concentration to 35% in the mother-liquor solution, which produced visibly lower damage to crystals than glycerol and a substantial change in the crystal unit cell (space group $P2_12_12_1$, unit-cell parameters $a = 57.669$, $b = 112.501$, $c = 139.720$ Å; Table 1). In addition, we observed that protein incubation with both products of the reaction, 2 m*M* ADP and 2 m*M* IP_6, yielded crystals of higher quality, a few of which diffracted to 2.3 Å resolution (Table 1). In general, these crystals appeared in 4–5 days and continued growing over the following two weeks to maximum dimensions of 0.3 × 0.05 × 0.03 mm.

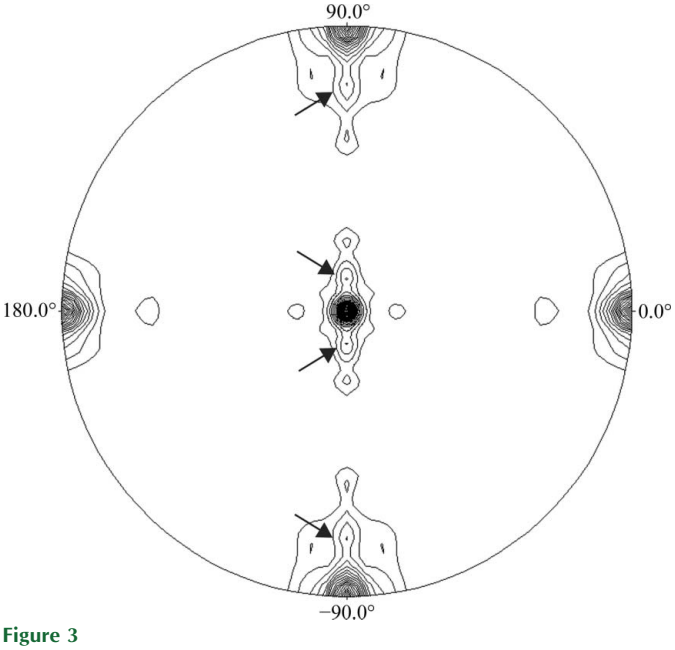


Figure 3
Plot of the self-rotation function of IP_5 2-K crystals from sample II using data between 15.0 and 2.2 Å resolution and a 25.0 Å radius of integration in the $\kappa = 180^\circ$ section. The view is down the *c* axis. $\varphi = 0^\circ$ and $\varphi = 90^\circ$ correspond to the *a* and the *b* axes, respectively. The peaks highlighted show a noncrystallographic twofold axis along the *bc* plane, which is supported by the heavy-atom positions.

Assuming that two molecules are present in the asymmetric unit, the Matthews coefficient (Matthews, 1968) of crystals from sample II was 2.33 Å^3 Da^{-1}, leading to a 47% solvent content within the unit cell. We have investigated the local symmetry relating the units in the asymmetric unit using the *CCP*4 package program *POLARRFN* (Kabsch, 1976). Several self-rotation functions were computed in the resolution range 15–3 Å, with Patterson vectors from 25 to 40 Å radius of integration. Analysis of self-rotation peaks revealed the presence of noncrystallographic twofold symmetry along the *bc* plane in the sample II crystals, although their height was rather low. The stereographic projection ($\kappa = 180^\circ$ section) of the self-rotation is shown in Fig. 3.

We have expressed and purified selenomethionine-substituted protein, which yielded crystals under very similar conditions to the native crystals but that always grew as plates. The crystals were not suitable for X-ray data analysis since they were twinned. Therefore, we set up a series of our best crystallization assays with sample II in

the presence of 2 m*M* ADP and 2 m*M* IP_6 and the crystals obtained were used for the preparation of heavy-atom derivatives. MAD experiments were carried out on the ID23-1 beamline using two crystals, one soaked in TMLA and other in lead acetate, which diffracted to >3 Å resolution (the peak data sets are shown in Table 1). It was possible to find several lead positions in the asymmetric units for both derivatives, which were refined and used for phase calculations. The phasing procedure confirmed the presence of two molecules in the asymmetric unit. Model building in the experimental electron-density maps obtained and refinement against the higher resolution native data set are currently in progress.

We thank Olga Perisic and Jose Miguel Mancheño for providing us with pOPTG and $pKLSL_t$ vectors, respectively. We thank the ESRF staff for providing time for and assistance with data collection. JIB is supported by an FPU fellowship (AP2008-00916) from the Ministerio de Educación. BG is supported by a 'Ramon y Cajal' fellowship (RYC-2006-002701). This work was supported by grants from the Comunidad de Madrid-CSIC (CCG07-CSIC/GEN-2232 and CCG08-CSIC/GEN-3490), Ministerio de Ciencia e Innovación (BFU2008-02897/BMC) and Ministerio de Educación y Ciencia (RYC-2006-002701). CAB was supported by the Biotechnology and Biological Sciences Research Council of the UK (grant BB/C514090/1).

References

Agarwal, R., Mumtaz, H. & Ali, N. (2009). *Mol. Cell. Biochem.* **328**, 155–165.
Bozsik, A., Kokeny, S. & Olah, E. (2007). *Cancer Genomics Proteomics*, **4**, 43–52.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Diallo, J. S., Betton, B., Parent, N., Peant, B., Lessard, L., Le Page, C., Bertrand, R., Mes-Masson, A. M. & Saad, F. (2008). *Br. J. Cancer*, **99**, 1613–1622.
González, B., Schell, M. J., Letcher, A. J., Veprintsev, D. B., Irvine, R. F. & Williams, R. L. (2004). *Mol. Cell*, **15**, 689–701.
Hanakahi, L. A. & West, S. C. (2002). *EMBO J.* **21**, 2038–2044.
Holmes, W. & Jogl, G. (2006). *J. Biol. Chem.* **281**, 38109–38116.
Ives, E. B., Nichols, J., Wente, S. R. & York, J. D. (2000). *J. Biol. Chem.* **275**, 36575–36583.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Leslie, A. G. W. (1990). *Crystallographic Computing 5. From Chemistry to Biology*, edited by D. Moras, A. D. Podjarny & J.-C. Thierry, pp. 50–61. Oxford University Press.
Mancheño Gómez, J. M. & Angulo Herrera, I. (2009). Patent WO/2009/121994.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miller, G. J. & Hurley, J. H. (2004). *Mol. Cell*, **15**, 703–711.
Miller, G. J., Wilson, M. P., Majerus, P. W. & Hurley, J. H. (2005). *Mol. Cell*, **18**, 201–212.
Mulugu, S., Bai, W., Fridy, P. C., Bastidas, R. J., Otto, J. C., Dollins, D. E., Haystead, T. A., Ribeiro, A. A. & York, J. D. (2007). *Science*, **316**, 106–109.
Murphy, A. M., Otto, B., Brearley, C. A., Carr, J. P. & Hanke, D. E. (2008). *Plant J.* **56**, 638–652.
Phillippy, B. Q., Ullah, A. H. & Ehrlich, K. C. (1994). *J. Biol. Chem.* **269**, 28393–28399.
Sarmah, B., Latimer, A. J., Appel, B. & Wente, S. R. (2005). *Dev. Cell*, **9**, 133–145.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Shen, X., Xiao, H., Ranallo, R., Wu, W. H. & Wu, C. (2003). *Science*, **299**, 112–114.
Steger, D. J., Haswell, E. S., Miller, A. L., Wente, S. R. & O'Shea, E. K. (2003). *Science*, **299**, 114–116.
Sun, Y., Thompson, M., Lin, G., Butler, H., Gao, Z., Thornburgh, S., Yau, K., Smith, D. A. & Shukla, V. K. (2007). *Plant Physiol.* **144**, 1278–1291.
Suzuki, M., Tanaka, K., Kuwano, M. & Yoshida, K. T. (2007). *Gene*, **405**, 55–64.
Sweetman, D., Johnson, S., Caddick, S. E., Hanke, D. E. & Brearley, C. A. (2006). *Biochem. J.* **394**, 95–103.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.
Verbsky, J., Lavine, K. & Majerus, P. W. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8448–8453.
Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W. & Wente, S. R. (2002). *J. Biol. Chem.* **277**, 31857–31862.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
York, J. D., Odom, A. R., Murphy, R., Ives, E. B. & Wente, S. R. (1999). *Science*, **285**, 96–100.