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

The crystallization of apo-form UMP kinase from *Xanthomonas campestris* is significantly improved in a strong magnetic field

Jhe-Le Tu,[a] Ko-Hsin Chin,[a,b] Andrew H.-J. Wang[c,d] and Shan-Ho Chou[a,b]*

[a]Institute of Biochemistry, National Chung-Hsing University, Taichung 40227, Taiwan, [b]National Chung-Hsing University Biotechnology Center, National Chung-Hsing University, Taichung 40227, Taiwan, [c]Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei, Taiwan, and [d]Core Facility for Protein Crystallography, Academia Sinica, Nankang, Taipei, Taiwan

Correspondence e-mail: shchou@nchu.edu.tw

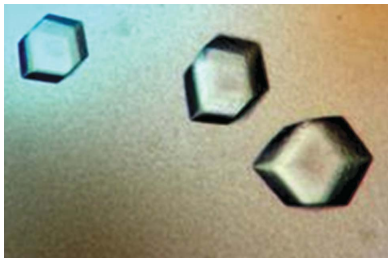
Received 16 March 2007
Accepted 16 April 2007

Bacterial UMP kinases (UMPKs) are crucial enzymes that are responsible for microbial UTP biosynthesis. Interestingly, eukaryotic and prokaryotic cells use different enzymes for UMP-phosphorylation reactions. Prokaryotic UMPKs are thus believed to be potential targets for antimicrobial drug development. Here, the cloning, expression and crystallization of SeMet-substituted XC1936, a bacterial UMPK from *Xanthomonas campestris* pathovar *campestris*, are reported. The crystallization of the apo-form UMPK was found to be significantly improved in a strong magnetic field; the crystals diffracted to a resolution of 2.35 Å, a dramatic improvement over the original value of 3.6 Å. Preliminary structural analyses of apo-form XC1936 using crystals grown in a strong magnetic field clearly reveal well defined loop regions involved in substrate-analogue binding that were previously not visible. Crystallization in a strong magnetic field thus was found to be indispensable in determining the flexible region of the XC1936 UMPK structure.

1. Introduction

XC1936 (gi|21112431) from the plant pathogen *Xanthomonas campestris* pathovar *campestris* strain 17 is classified as belonging to the amino-acid kinase superfamily in the Pfam database using a bioinformatics approach (Bateman *et al.*, 2000). This family includes kinases that phosphorylate a variety of amino-acid substrates (aspartate, acetylglutamate), as well as uridylate kinase and carbamate kinase. A *BLAST* search of the PDB, however, returned several UMPKs (PDB codes 2brx, 2ij9, 1z9d, 1ybd, 2a1f and 2bnd) with high identities (ranging from 26.7% to 54.3%), indicating that XC1936 is very likely to be a UMPK. Nevertheless, in the determined *Escherichia coli* UMPK structure (2bnd) the loop-region structure for binding nucleotide phosphates could only be determined when it was bound to UMP, UDP or UTP (Briozzo *et al.*, 2005), while in the *Pyrococcus furiosus* UMPK structure (2brx) the corresponding loop region was invisible in the electron-density map (Marco-Marin *et al.*, 2005). This phenomenon indicates that the loop regions for binding nucleotide phosphates are very flexible and their structures are difficult to determine without extra effort. Since bacterial UMPKs are specific for the phosphorylation of UMP only (Briozzo *et al.*, 2005; Marco-Marin *et al.*, 2005) and differ from eukaryotic dual-specificity UMP/CMP kinases (Yan & Tsai, 1999), they are considered to be potential targets for antibacterial drug development. We have therefore endeavoured to solve the *X. campestris* UMPK structure both in its apo form and substrate/substrate-analogue-bound form in order to better examine the conformational changes associated with ligand binding.

Although the cloning, purification and expression of XC1936 were straightforward and large amounts of SeMet-substituted XC1936 protein could be obtained, its crystallization in the absence of ligand was difficult: we were unable to obtain apo-XC1936 crystals that diffracted to a resolution better than 3.6 Å. In order to overcome this problem, we adopted two additional steps: firstly, we screened a modified buffer set in order to obtain the optimum solubility buffer that increased the solubility and monodispersity of XC1936 (Collins *et al.*, 2005; Izaac *et al.*, 2006; Jancarik *et al.*, 2004); secondly, we



crystallized the XC1936 protein in the optimum buffer in a strong magnetic field in order to help orient the XC1936 molecules (Lin *et al.*, 2000; Saijo *et al.*, 2005). Using these approaches, we were finally able to obtain XC1936 crystals that diffracted to a resolution of at least 2.35 Å. Importantly, the flexible-loop regions that were previously invisible in the structures of the UMPKs from *E. coli* and *P. furiosus* are now very well defined, as revealed by the clear electron-density map from the preliminary crystal structure analysis.

2. Materials and methods

2.1. Cloning, expression and purification

The XC1936 gene fragment was PCR-amplified directly from a local *X. campestris* genome (*X. campestris* pathovar *campestris* strain 17) with a forward 5′-*TACTTCCAATCCAATGCT*ATGTCTGAACTTTCCTATCGCCGCATC primer and a reverse 5′-*TTATCCACTTCCAATG*TCAGCTGCGGCCCTGTACCAAC primer (the linker sequences are italicised). A ligation-independent cloning (LIC) approach (Aslanidis & de Jong, 1990) was carried out in order to obtain the desired construct according to a previously published protocol (Wu *et al.*, 2005). The final construct codes for an N-terminal His_6 tag, a 17-amino-acid linker and the XC1936 target protein (240 amino acids) under the control of a T7 promoter. The vector was transformed into *E. coli* BL21 (DE3)-pLysS host cells, which were grown in LB medium at 310 K until an OD_{600} of 0.8 was attained. Overexpression of the His_6-tagged target protein was induced by the addition of 0.5 m*M* IPTG at 293 K for 21 h. The cells were harvested, resuspended in equilibration buffer (20 m*M* Tris–HCl, 80 m*M* NaCl pH 8.0) and lysed using a microfluidizer (Microfluidics). Most of the tagged target protein was present in the soluble fraction (Fig. 1). After centrifugation, the target protein was purified by immobilized metal-affinity chromatography (IMAC) on a nickel column (Sigma) and eluted with 20 m*M* Tris pH 8.0, 80 m*M* NaCl in a gradient of 50–300 m*M* imidazole. The fractions containing XC1936 were monitored by SDS–PAGE, recombined and dialyzed repeatedly against 20 m*M* Tris–HCl, 80 m*M* NaCl pH 8.0. After buffer exchange, the His_6 tag and linker were cleaved from XC1936 by TEV (tobacco etch virus) protease at 295 K for 16 h and removed by immobilized metal-affinity chromatography (IMAC) on a nickel column (Sigma). For crystallization, XC1936 was further purified by FPLC (AKTA, Pharmacia Inc.) on an anion-exchange column. The fractions eluted with 20 m*M* Tris pH 8.0, 1 *M* NaCl were combined and dialyzed against 20 m*M* Tris pH 8.0 and 80 m*M* NaCl. The final target protein (240 amino acids) has greater than 99% purity (Fig. 1) and contains only an extra tripeptide (SNA) at the N-terminal end. The SeMet-substituted XC1936 was expressed in a similar way, except that the host *E. coli* cells were induced in M9 minimum medium containing SeMet and when an OD_{600} of 0.8 was reached. The overexpression and purification of SeMet-substituted XC1936 was monitored by SDS–PAGE as shown in Fig. 1.

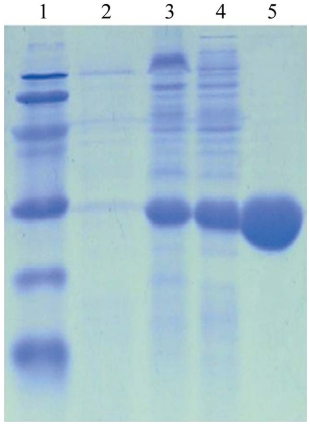
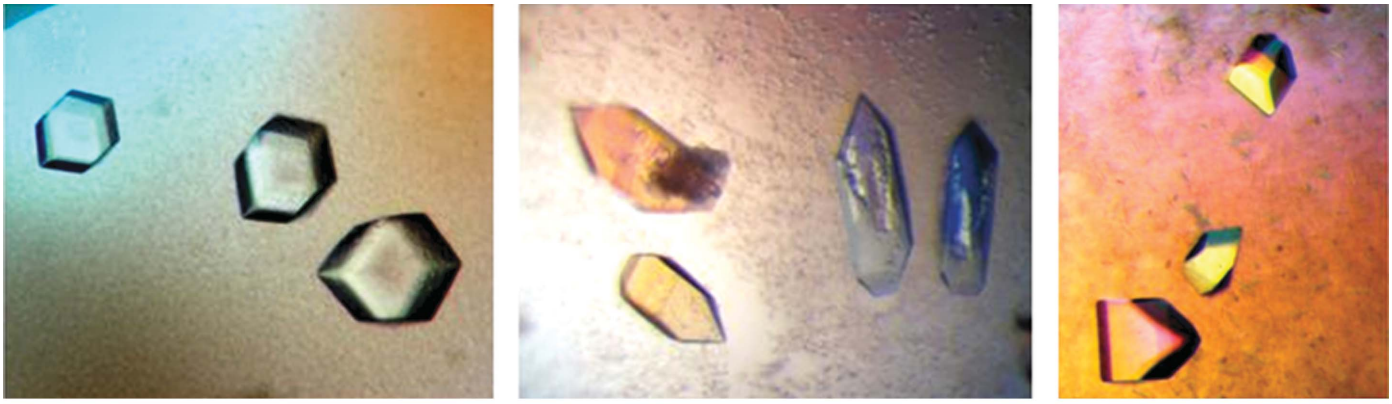


Figure 1
SDS–PAGE monitoring of the overexpression and purification of XC1936. Lane 1, protein molecular-weight markers; lane 2, whole cell lysate before IPTG induction; lane 3, whole cell lysate after IPTG induction; lane 4, soluble-form XC1936; lane 5, purified XC1936 after TEV cleavage.

2.2. Optimum solubility screen

Initial crystallization trials of XC1936 dissolved in the default purification buffer (20 m*M* Tris–HCl, 80 m*M* NaCl pH 8.0) using Hampton sparse-matrix Crystal Screens 1 and 2, a systematic PEG–pH screen and the PEG/Ion Screen only produced crystals that diffracted to a resolution of approximately 4 Å (see Fig. 2*a*). Since the non-optimum solubility of XC1936 in the original purification buffer may cause its aggregation and hence inhibit its crystal nucleation and growth (Ferré-D'Amaré & Burley, 1997; Habel *et al.*, 2001), we



(*a*) (*b*) (*c*)

Figure 2
Crystals of SeMet-substituted XC1936 from *X. campestris* grown by the hanging-drop vapour-diffusion method under various conditions. (*a*) The protein was dissolved in the original purification buffer and cube-shaped crystals grew from the crystallization drop using 1.26 *M* $(NH_4)_2SO_4$. The average dimensions of these crystals were approximately 0.15 × 0.15 × 0.1 mm. (*b*) The protein was dissolved in the optimum buffer and crystals of XC1936 with a rugged surface grew from the drop using 0.1 *M* sodium citrate pH 5.6, 20%(*w*/*v*) PEG 3350, 0.05 *M* $(NH_4)_2SO_4$. The average dimensions of these crystals were approximately 0.1 × 0.05 × 0.07 mm. (*c*) The protein was dissolved in the same optimum buffer and grown using the same crystallization buffer, but in plates that were placed alongside an 18.7 T superconducting magnet during crystallization. The average dimensions of these crystals were approximately 0.13 × 0.1 × 0.1 mm after 7 d.

Table 1
The modified optimum-solubility screen buffers.

No.	Buffer (100 m*M*)	pH
1	Glycine	3
2	Citric acid	3.2
3	Citric acid	4
4	Sodium acetate	4.5
5	Sodium acetate	5
6	Sodium phosphate	5.5
7	Sodium phosphate	6.5
8	Sodium phosphate	7.5
9	Potassium phosphate	5
10	Potassium phosphate	6
11	Potassium phosphate	7
12	Potassium phosphate	7.4
13	Sodium citrate	4.7
14	Sodium citrate	5.5
15	Bis-Tris	6
16	Bis-Tris	6.5
17	Bis-Tris	7
18	ADA	6.2
19	ADA	6.8
20	Bis-Tris propane	6.5
21	Bis-Tris propane	7
22	Bis-Tris propane	8
23	Bis-Tris propane	8.5
24	Ammonium acetate	7
25	Ammonium acetate	7.5
26	MOPS	7
27	HEPES	7
28	HEPES	7.5
29	HEPES	8
30	Tris	7.5
31	Tris	8
32	Tris	8.5
33	MES	5.8
34	MES	6.2
35	MES	6.5
36	Bicine	8.5
37	Bicine	9
38	(H)EPPS	8
39	Imidazole	8
40	CHES	9
41	CHES	9.5
42	CAPS	10
43	PIPES	6.5
44	TAPS	8.5
45	Cacodylate	6.5
46	PIPPS (polyisoprenyl phosphate)	3.7
47	Sodium formate	
48	Sodium sulfate	
49	Lithium chloride	
50	Magnesium chloride	
51	Calcium chloride	
52	Potassium chloride	
53	Ammonium chloride	

performed a buffer-screening procedure before crystallization by modifying the recently developed optimum-solubility (OS) screening method (Jancarik *et al.*, 2004) in order to improve the solubility of XC1936 and to obtain the most homogeneous and monodisperse protein condition. To find a better buffer condition, we expanded the original set of 24 buffers to 53 buffers, each at a concentration of 100 m*M*, with a pH ranging from 3 to 10, as shown in Table 1. Before screening, the protein sample was dialyzed against distilled H_2O and concentrated to about 30 mg ml^{-1} using an Amicon Ultra-10 (Millipore). The suitable protein concentration for screening was first examined using formulation No. 6 from Hampton Crystal Screen 1 [0.1 *M* Tris–HCl pH 8.5, 0.2 *M* magnesium chloride and 30%(*w*/*v*) PEG 4K]. If the protein drop remained clear after 20 min, the protein concentration was considered possibly to be too low and to need to be concentrated further. When a suitable protein concentration had been determined, a 1 µl aliquot was pipetted into 1 µl each of the buffers listed in Table 2 in a transparent tube. The contents were

Table 2
Data-collection statistics for SeMet-substituted XC1936.

Values in parentheses are for the highest resolution shell.

	Without buffer screening	With buffer screening	With buffer screening and in high magnetic field
Space group	*I*23	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	*a* = *b* = *c* = 156.3	*a* = 111.35, *b* = 120.17, *c* = 125.82	*a* = 111.45, *b* = 120.07, *c* = 125.79
Temperature (K)	100	100	100
Wavelength (Å)	0.963981	0.97963	0.96411
Resolution range (Å)	50.0–3.6 (3.73–3.6)	30–3.6 (3.71–3.6)	30–2.35 (2.45–2.35)
Mosaicity (°)	0.6	0.42	0.45
Unique reflections	35852	186235	277840
Redundancy	5	3	4
Completeness (%)	99.6 (100)	98.6 (97.3)	99.4 (99.9)
$R_{\rm merge}$ (%)	9.6	12	5.7
Mean $I/\sigma(I)$	10.5 (8.2)	9.2 (3.1)	15.8 (6.2)
Solvent content (%)	58	50	50

mixed well and kept in a freezer at 193 K for half an hour before being transferred to a refrigerator at 277 K until the frozen protein samples had thawed. The degrees of precipitation of the protein samples depend on the solubility of protein in the given buffer and can be monitored under a light microscope.

The drops in which no precipitation was observed were chosen as starting buffers and these were further screened using dynamic light-scattering (DLS) analysis. The clear XC1936 protein solution was first diluted into the same reservoir solution in a ratio of 1:15 to give a final concentration of 2 mg ml^{-1}. The DLS analyses were then performed using a PD2000 laser light-scattering detector (Precision Detectors Inc.). Because no ideal monodisperse protein sample was observed, additive screening using the four buffers that gave the best DLS readings [potassium phosphate pH 7, sodium citrate pH 5.5, *N*-(2-acetamido)-2-iminodiacetic acid (ADA) pH 6.8 and HEPES pH 7.5] was performed. 15 additives were selected based on their benefit to protein solubility. Two further additives, 0.1% of the nonionic surfactants Tween-20 and Triton X-100, were included in the original screen recipe to prevent agitation-induced aggregation of the recombinant fusion protein (Chou *et al.*, 2005). 5 m*M* β-mercaptoethanol, along with 100 m*M* ADA pH 6.8, were found to serve as the optimum buffer for solubilizing XC1936 protein sample. The polydispersity and radius values for XC1936 were improved from 75% and 54.5 nm, respectively, in the original purification buffer to 26% and 9.63 nm in the final optimum buffer.

2.3. Crystallization

For crystallization, the XC1936 protein sample was dialyzed against the optimum buffer (20 m*M* ADA pH 6.8, 5 m*M* β-mercaptoethanol and 0.02% NaN_3) and concentrated to 8.5 mg ml^{-1} using an Amicon Ultra-10 (Millipore). Screening for crystallization conditions was performed using sitting-drop vapour diffusion in 96-well plates (Hampton Research) at 293 K by mixing 0.5 µl protein solution with 0.5 µl reagent solution. Initial screens included the Hampton sparse-matrix Crystal Screens 1 and 2, a systematic PEG–pH screen and the PEG/Ion Screen and were performed using a Gilson C240 crystallization workstation. We observed crystal formation using a reservoir solution comprising 0.1 *M* sodium citrate pH 5.6, 20%(*w*/*v*) PEG 3350, 0.05 *M* $(NH_4)_2SO_4$; the crystals reached maximum dimensions of 0.1 × 0.05 × 0.07 mm in 9 d (Fig. 2*b*). The crystals grown using the optimum buffer (Fig. 2*b*) have sharper edges than those grown with the original buffer (20 m*M* Tris–HCl pH 8.0 and 80 m*M* NaCl; Fig. 2*a*),

although some cavities were still observed in the surface of the crystals.

Although the crystal quality had been improved somewhat at this stage, the diffraction data were still not good enough for structural determination; the resolution only reached 3.6 Å. We were unable to determine the required protein phases using the MAD approach. Interestingly, it has been reported that the quality of crystals can be significantly improved by applying a strong magnetic field during crystallization (Lin *et al.*, 2000; Saijo *et al.*, 2005). The improved hydrodynamic properties are attributed to the reduced convection and reduced gravity induced by the magnetic force. We therefore placed the crystallization plates alongside a strong magnetic field (18.7 T) to help orient the protein molecules. Good-quality crystals of XC1936 appeared in 7 d from a reservoir solution containing the optimum buffer at 298 K and reached maximum dimensions of 0.13 × 0.1 × 0.1 mm. The crystals grown under these conditions not only have sharper edges (Fig. 2*c*) than those obtained in the absence of a strong magnetic field, but also exhibit greatly improved resolution (2.35 Å).

2.4. Data collection

Crystals were soaked in the mother liquor and then flash-cooled at 100 K under a stream of cold nitrogen. X-ray diffraction data were collected using Cu $K\alpha$ radiation from a Rigaku MicroMax007 rotating-anode generator equipped with Osmic mirror optics and a R-AXIS IV^{++} image plate. An SeMet-substituted data set was obtained to a maximum resolution of 2.35 Å from crystals grown using optimum buffer in a magnetic field. The data were indexed and integrated using the *HKL* processing software (Otwinowski & Minor, 1997), giving a data set that was 99.4% complete with an overall $R_{\rm merge}$ of 5.7% on intensities. The crystals belong to the monoclinic space group $P2_12_12_1$. The data-collection statistics for the crystals grown using the original buffer, using the optimum buffer and using further magnetic field treatment are summarized in Table 2; an X-ray diffraction image of the latter collected in-house is shown in Fig. 3.

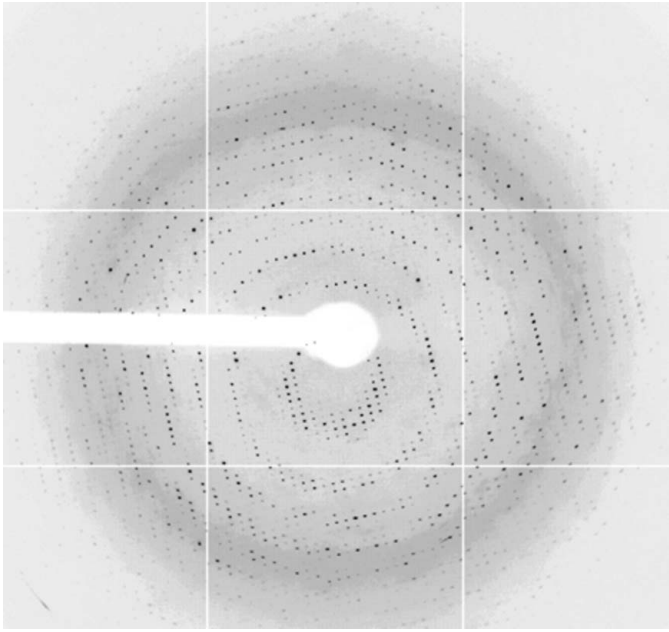


Figure 3
A diffraction pattern of an SeMet XC1936 crystal grown under optimum buffer screening and magnetic field treatment and flash-frozen in reservoir cryoprotectant. The exposure time was 10 s, with an oscillation range of 1° and a crystal-to-film distance of 300 mm. The crystal diffracted to a resolution of 2.35 Å.

3. Results and discussion

Although we were able to obtain cube-shaped crystals of XC1936 (Fig. 2*a*) using the original purification buffer, we found that their diffraction quality was not good enough to solve the required protein phase angles. We used two additional steps to obtain crystals that were suitable for the determination of the *X. campestris* UMPK structure in the apo form. The first step involved selecting a better buffer than that used during purification in order to increase the solubility and decrease the aggregation of XC1936. Initially, we used the screening buffers defined by Jancarik *et al.* (2004). However, soon after we found that buffers with different pH values may exhibit very different effects on the protein solubility. Hence, we developed an expanded set of screen buffers with finer pH steps as shown in Table 1. Several extra salts such as sodium formate, sodium sulfate, lithium chloride, magnesium chloride, calcium chloride, potassium chloride and ammonium chloride were included as it has been reported that these salts can further increase protein solubility (Izaac *et al.*, 2006). In addition, we also included two extra nonionic surfactants (0.1% Tween-20 and 0.1% Triton X-100) to expand the additive screen (Chou *et al.*, 2005). However, using the optimum buffer (20 m*M* ADA pH 6.8, 5 m*M* β-mercaptoethanol and 0.02% NaN_3) obtained from this screening step, the resulting XC1936 crystals still did not seem to be greatly improved in terms of either their appearance (the crystal surface appeared rugged; Fig. 2*b*) or their diffraction quality.

It was reasoned that the precipitation rate of the XC1936 protein may have been too fast and that the application of a strong superconducting magnetic field may partially reduce gravity and hence the sedimentation speed (Lin *et al.*, 2000). When XC1936 was dissolved in the same optimum buffer and crystallization was performed in the same manner, but with the crystallization plates placed alongside a strong 18.7 T superconducting magnet, it was observed that the crystal quality indeed improved significantly, as judged by the sharper edges of the crystals and the improved diffraction resolution (Table 2). The average dimensions of these crystals were 0.13 × 0.1 × 0.1 mm after 7 d and the crystals were highly suitable for structural determination.

Importantly, preliminary structural analysis of the apo-form XC1936 hexamer revealed a well determined electron-density map in the loop (^{56}NIFRGAGLAAS^{66}) partially responsible for substrate/substrate-analogue binding (data to be published), indicating that application of a strong magnetic field can be very beneficial in determination of the structure of flexible loops. As we have also successfully obtained cocrystals of XC1936 UMPK with several of its cofactors, such as UMP, UDP, UTP, AMPPNP and GTP (data not shown), detailed structural comparison between apo-form UMPK from *X. campestris* and its substrate/substrate analogue/cofactor complexes could reveal interesting conformational changes associated with ligand binding. We are currently collecting further experimental data and detailed structural studies are now under way.

This work was supported by an Academic Excellence Pursuit grant from the Ministry of Education and by the National Science Council, Taiwan (grant 95-2113-M-005-018-MY2) to S-HC. We thank the National Synchrotron Radiation Research Center, Taiwan and the SPring-8 Synchrotron facility in Japan for assistance during X-ray data collection. The National Synchrotron Radiation Research Center is a user facility supported by the National Science Council,

Taiwan and the Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine, Taiwan.

References

Aslanidis, C. & de Jong, P. C. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L. & Sonnhammer, E. L. L. (2000). *Nucleic Acids Res.* **28**, 263–266.
Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Barzu, O. & Gilles, A.-M. (2005). *J. Biol. Chem.* **280**, 25533–25540.
Chou, D. K., Krishnamurthy, R., Randolph, T. W., Carpenter, J. F. & Manning, M. C. (2005). *J. Pharm. Sci.* **94**, 1368–1381.
Collins, B., Stevens, R. C. & Page, R. (2005). *Acta Cryst.* F**61**, 1035–1038.
Ferré-D'Amaré, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
Habel, J. E., Ohren, J. F. & Borgstahl, G. E. O. (2001). *Acta Cryst.* D**57**, 254–259.
Izaac, A., Schall, C. A. & Mueser, T. C. (2006). *Acta Cryst.* D**62**, 833–842.
Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. (2004). *Acta Cryst.* D**60**, 1670–1673.
Lin, S.-X., Zhou, M., Azzi, A., Xu, G.-J., Wakayama, N. I. & Ataka, M. (2000). *Biochem. Biophys. Res. Commun.* **275**, 274–278.
Marco-Marin, C., Gil-Ortiz, F. & Rubio, V. (2005). *J. Mol. Biol.* **352**, 438–454.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Saijo, S., Yamada, Y., Sato, T., Tanaka, N., Matsui, T., Sazaki, G., Nakajima, K. & Matsuura, Y. (2005). *Acta Cryst.* D**61**, 207–217.
Wu, Y.-Y., Chin, K.-H., Chou, C.-C., Lee, C.-C., Shr, H.-L., Lyu, P.-C., Wang, A. H.-J. & Chou, S.-H. (2005). *Acta Cryst.* F**61**, 902–905.
Yan, H. G. & Tsai, M. D. (1999). *Adv. Enzymol. Relat. Areas Mol. Biol.* **73**, 103–134.