

Ana M. Garrote,<sup>a</sup> Pilar Redondo,<sup>a</sup> Guillermo Montoya<sup>a,b\*</sup> and Inés G. Muñoz<sup>a\*</sup>

<sup>a</sup>Macromolecular Crystallography Group, Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), Melchor Fernández Almagro 3, 28029 Madrid, Spain, and <sup>b</sup>Structural Biology Group, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark

Correspondence e-mail: gmontoya@cnio.es, imunoz@cnio.es

Received 20 December 2013

Accepted 4 February 2014

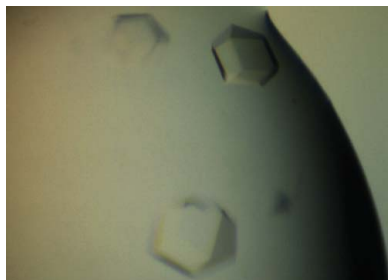
## Purification, crystallization and preliminary X-ray diffraction analysis of the kinase domain of human tousled-like kinase 2

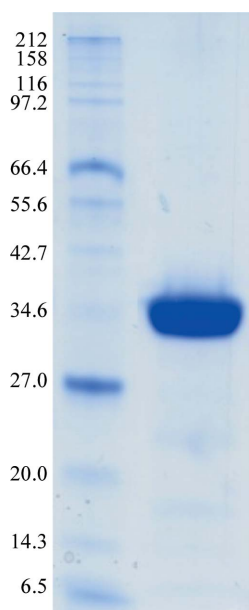
Tousled-like kinases (TLKs) are an evolutionarily conserved family of serine/threonine protein kinases involved in chromatin dynamics, including DNA replication and repair, transcription and chromosome segregation. The two members of the family reported in humans, namely TLK1 and TLK2, localize to the cell nucleus and are capable of forming homo- or hetero-oligomers by themselves. To characterize the role of TLK2, its C-terminal kinase domain was cloned and overexpressed in *Escherichia coli* followed by purification to homogeneity. Crystallization experiments in the presence of ATP- $\gamma$ -S yielded crystals suitable for X-ray diffraction analysis belonging to two different space groups: tetragonal  $I4_122$  and cubic  $P2_13$ . The latter produced the best diffracting crystal (3.4 Å resolution using synchrotron radiation), with unit-cell parameters  $a = b = c = 126.05$  Å,  $\alpha = \beta = \gamma = 90^\circ$ . The asymmetric unit contained one protein molecule, with a Matthews coefficient of  $4.59$  Å<sup>3</sup> Da<sup>-1</sup> and a solvent content of 73.23%.

### 1. Introduction

The tousled-like kinases (TLKs) are a family of conserved serine/threonine protein kinases that are found in both plants and animals. They were named after the first member of the family to be described, tousled (TSL) protein, which is encoded by the *Arabidopsis thaliana* TSL gene (Roe, Nemhauser *et al.*, 1997; Roe *et al.*, 1993). Mutations in this gene result in abnormal flower development, affecting leaf morphology, flowering time and organogenesis (Roe, Nemhauser *et al.*, 1997; Roe *et al.*, 1993; Moshkin *et al.*, 2009). The human TLKs, namely TLK1 and TLK2, localize to the cell nucleus and interact with themselves, forming homo- or hetero-oligomers, as proven by yeast two-hybrid experiments (Roe, Durfee *et al.*, 1997; Silljé *et al.*, 1999). They share 84% similarity with each other and almost 50% similarity with *Arabidopsis* TSL (Takahata *et al.*, 2009). Both TLKs have their catalytic domains placed at the C-terminus, where the ATP-binding domain displays a GXGXXS motif instead of the canonical consensus sequence GXGXXG (Hanks *et al.*, 1988). In the N-terminal domain, conserved features include three potential nuclear localization sequences and three putative coiled-coil regions (Silljé *et al.*, 1999).

The TLKs are regulated by cell-cycle-dependent phosphorylation and their activity is strongly linked to DNA replication, with maximal activity during the S phase. Also, they exhibit sensitivity to DNA-damaging agents and inhibitors of DNA replication (Takahata *et al.*, 2009). TLKs are involved in chromatin assembly through the binding and phosphorylation of the human chromatin assembly factors Asf1a and Asf1b (Silljé & Nigg, 2001), and have been associated with numerous replicative and transcriptional processes, including chromosome condensation and segregation (Sunavala-Dossabhoy *et al.*, 2003; Hashimoto *et al.*, 2008; Yap *et al.*, 2011), gene silencing (Wang *et al.*, 2006) and DNA repair (Li *et al.*, 2001; Canfield *et al.*, 2009); the latter is further supported by their role as targets of ATM and Chk1, two kinases that are involved in the DNA-damage checkpoint upon irradiation (Krause *et al.*, 2003). Recent studies have demonstrated their direct implication in the development of different types of cancer such as neuroblastoma and pancreatic and prostate cancers (Heidenblad *et al.*, 2005; Wang *et al.*, 2006; Ronald *et al.*, 2011), being key regulators of Kaposi's sarcoma-associated herpesvirus and Epstein-Barr virus reactivation, since their expression is required for





**Figure 1**  
12% SDS-PAGE of purified TLK2\_KD used for crystallization trials. Molecular-weight markers are labelled in kDa.

the maintenance of viral latency (Dillon *et al.*, 2013). This suggests the potential use of TLKs as biomarkers for diagnosis and prognosis, without neglecting the possibility of their use as drug targets affecting not only pathogenesis but also the response of the patient to chemotherapy and/or radiotherapy. In this study, we report the purification, crystallization, diffraction data collection and preliminary crystallographic analysis of the C-terminal kinase domain of TLK2.

## 2. Materials and methods

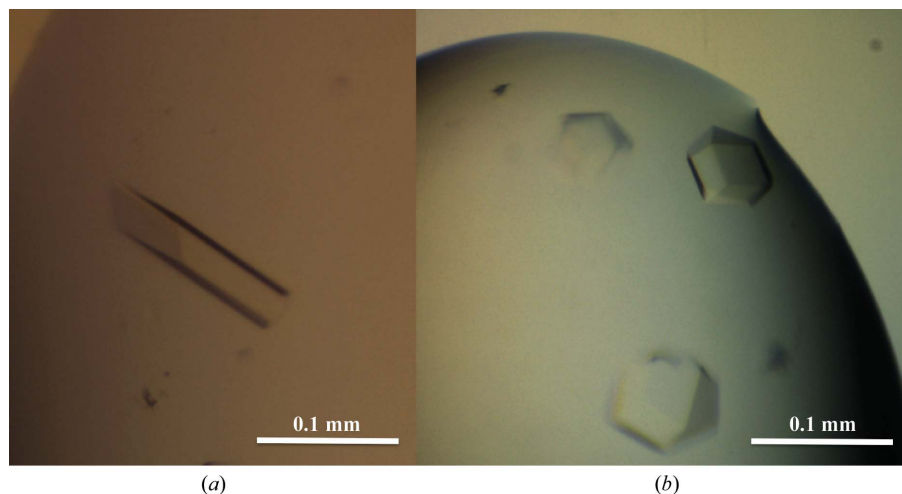
### 2.1. Cloning and expression

TLK2 is composed of 772 amino acids. The kinase domain (TLK2\_KD; amino acids 451–752) was amplified by PCR using the cDNA of TLK2 as a template (GenBank accession ID gi:6063019). The PCR product was inserted into multiple cloning site 1 (MCS1) of

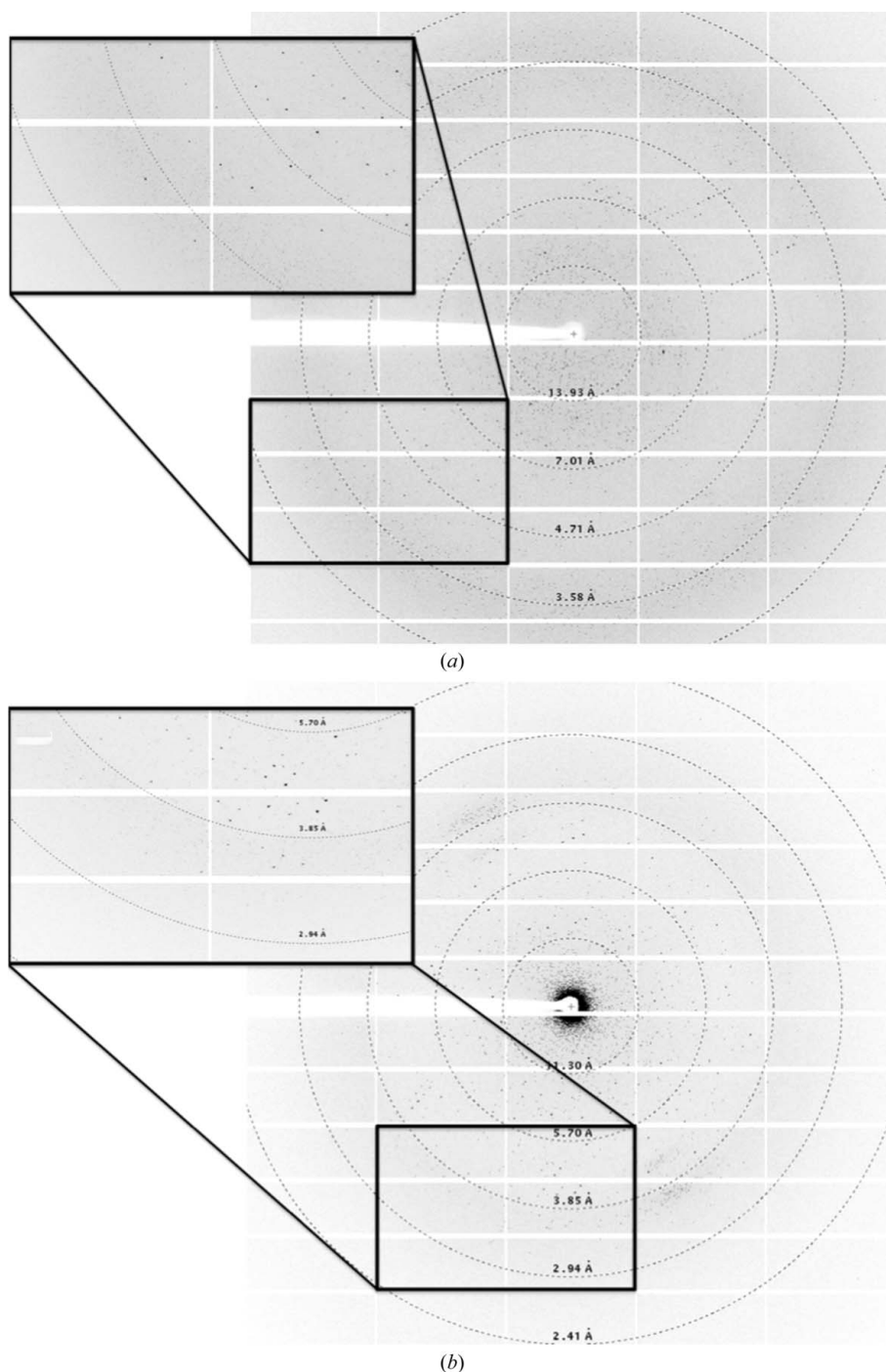
pETDuet-1 vector (Novagen) in frame with an N-terminal hexahistidine, thus adding an MGSSHHHHHH purification tag. This vector includes the lambda protein phosphatase ORF at the MCS2, expression of which permitted the dephosphorylation of TLK2\_KD to occur during its expression. The resulting construct was used to transform chemically competent *Escherichia coli* Rosetta pLysS cells grown in Luria–Bertani (LB) broth supplemented with 35  $\mu\text{g ml}^{-1}$  chloramphenicol and 100  $\mu\text{g ml}^{-1}$  ampicillin. Cells were cultured at 310 K and 200  $\text{rev min}^{-1}$  for 3 h after induction with 0.3 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) when the  $\text{OD}_{600}$  was between 0.6 and 0.8. Cells were then harvested by centrifugation at 4000  $\text{rev min}^{-1}$  for 20 min at 277 K and cooled by immersion in liquid nitrogen until further use.

### 2.2. Purification

To purify TLK2\_KD, the cell pellet was resuspended in 1:10(*w:v*) buffer *A* (50 mM Tris pH 8.8, 500 mM NaCl, 0.3 mM TCEP, 5% glycerol), including protease inhibitors (complete EDTA-free tablets, Roche). The resuspended cells were lysed by sonication (Sonic Dismembrator 500, Fisher Scientific) while cooling on ice for 15 min (2 s on/1 s off). The insoluble cell lysate was removed by centrifugation at 34 000*g* for 50 min at 277 K. We have estimated that approximately 80% of the protein is present in the soluble fraction. The supernatant was loaded onto a pre-equilibrated 5 ml  $\text{Ni}^{2+}$ -loaded HisTrap Chelating HP column (GE Healthcare). The bound protein was eluted over 30 column volumes using a linear imidazole gradient to a final concentration of 100% buffer *B* (20 mM Tris pH 8.8, 25 mM NaCl, 0.3 mM TCEP, 5% glycerol, 500 mM imidazole). The elution peak of the kinase domain occurs at 12.6% buffer *B*, when it reaches a salt content of 440 mM. After pooling the protein fractions, the protein was loaded onto a HiTrap Q HP column (GE Healthcare) pre-equilibrated in buffer *C* (20 mM Tris pH 8.8, 25 mM NaCl, 0.2 mM TCEP). The high salt concentration in the loading sample avoids the binding of TLK2\_KD to the column. The loading process was repeated several times to maximize the amount of protein contaminants bound to the column. In our experience, when the protein sample was dialyzed before loading against a buffer with lower salt concentration (50 mM NaCl), approximately 30% of the TLK2\_KD was bound to the ion-exchange column along with other contaminants, thus considerably decreasing the final yield. The TLK2\_KD fractions were concentrated to a final volume of 5 ml by



**Figure 2**  
Crystals of TLK2\_KD corresponding to space groups  $I4_122$  (a) and  $P2_13$  (b).



**Figure 3**  
X-ray diffraction pattern images corresponding to the space groups  $I4,22$  (a) and  $P2,3$  (b).

centrifugation ( $3000 \text{ rev min}^{-1}$  at 277 K) using an Amicon Ultra concentrator (Millipore) with a 10 kDa cutoff filter. The protein was then further purified and buffer-exchanged into buffer *D* (20 mM HEPES pH 7.5, 50 mM NaCl, 0.2 mM TCEP) by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare). Fractions corresponding to the purified protein were collected and concentrated to  $12.3 \text{ mg ml}^{-1}$  prior to crystallization. The homogeneity of the protein was determined by Coomassie Blue staining of a 12% SDS-PAGE gel (Fig. 1). The identity of the protein was confirmed by in-gel trypsin digestion (Sigma) and mass-spectrometric analysis (Proteomic Unit, CNIO, Madrid). The

molecular mass of the protein was determined to be 36 695.8 Da by electrospray ionization mass spectrometry (ESI-MS), which matches the sequence-predicted mass of 36 696 Da and does not include the initial methionine (Frottin *et al.*, 2006).

### 2.3. Crystallization

The crystallization trials were set up in 96-well sitting-drop plates (Innovadine) using a Cartesian MicroSys robot. The nanodrops were 0.2  $\mu\text{l}$  in size, consisting of 0.1  $\mu\text{l}$  reservoir solution plus 0.1  $\mu\text{l}$  protein solution at  $10 \text{ mg ml}^{-1}$ , and were equilibrated over 60  $\mu\text{l}$  reservoir

**Table 1**

Data-collection statistics for the two TLK2\_KD crystal forms.

Values in parentheses are for the outermost shell.

	Crystal form A	Crystal form B
Space group	<i>I</i> 4 <sub>1</sub> 22	<i>P</i> 2 <sub>1</sub> 3
Unit-cell parameters (Å, °)	<i>a</i> = <i>b</i> = 88.36, <i>c</i> = 342.87, α = β = γ = 90	<i>a</i> = <i>b</i> = <i>c</i> = 126.05, α = β = γ = 90
Data collection		
Temperature (K)	100	100
Wavelength (Å)	0.99987	1.00002
Beamline	XS06A, SLS	XS06A, SLS
Resolution (Å)	85.71–3.70 (3.90–3.70)	89.13–3.40 (3.58–3.40)
Total reflections	82242 (12186)	125158 (17818)
Unique reflections	7737 (1088)	9456 (1354)
Multiplicity	10.6 (11.2)	13.2 (13.2)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Mean <i>I</i> /σ( <i>I</i> )	19.1 (3.5)	13.5 (3.6)
<i>R</i> <sub>merge</sub> †	0.065 (0.79)	0.133 (0.71)

† *R*<sub>merge</sub> is defined according to *XDS* (Kabsch, 2010).

solution. The initial screening conditions were from the commercially available kits Crystal Screen, Crystal Screen 2 and Natrix (Hampton Research), Wizard I and II (Emerald BioSystems), JBScreen Kinase 1–4 and JBScreen Classic 1–10 (Jena), and The JCSG+ and Protein Complex Suites (Qiagen). For each screen the protein was tested either alone or in the presence of 2 mM ATP-γ-S. Duplicated crystal trays were stored at 277 and 291 K. These trials yielded crystals at 277 K under two conditions, Natrix condition No. 1 (crystal type A; 20 mM HEPES pH 7, 2 M Li<sub>2</sub>SO<sub>4</sub>, 10 mM MgCl<sub>2</sub>) and Wizard II condition No. 37 (crystal type B; 1 M sodium/potassium tartrate, 0.1 M Tris pH 7, 200 mM Li<sub>2</sub>SO<sub>4</sub>), and only when the protein was mixed with ATP-γ-S. Initial crystals were saved for seeding. Optimization of the crystallization conditions was performed by scaling up the initial hits using 24-well hanging-drop Linbro plates (Hampton Research). The crystals were reproducible using larger drops by mixing 1 μl reservoir solution and 1 μl protein solution at 5 mg ml<sup>-1</sup>, and in both cases they reached larger sizes when the reservoir solution was taken directly from the original Natrix and Wizard II kits (Fig. 2).

#### 2.4. Data collection

To prepare the crystals for data collection, they were cryoprotected by a 5 s soak in solutions consisting of 20 mM HEPES pH 7, 1 M Li<sub>2</sub>SO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 20% glycerol or 1 M sodium/potassium tartrate, 0.1 M Tris pH 7, 200 mM Li<sub>2</sub>SO<sub>4</sub>, 20% glycerol. They were then immediately vitrified in liquid nitrogen. Diffraction data were collected at 100 K on beamline XS06A at the Swiss Light Source (Villigen, Switzerland). 300 images of 0.5° oscillation per frame were collected from crystal type A, with a 620 mm crystal-to-detector distance and an exposure time of 0.5 s per image. 600 images of 0.2° oscillation per frame were collected from crystal type B, with a 500 mm crystal-to-detector distance and an exposure time of 0.2 s per image. Data were recorded for both crystals on a Pilatus detector; they were processed with *XDS* (Kabsch, 2010) and integrated and scaled using the *CCP4* program suite (Winn *et al.*, 2011). Data-collection statistics are summarized in Table 1.

### 3. Results

The kinase domain of TLK2 (TLK2\_KD) has been expressed and purified to homogeneity (Fig. 2). The yield of the pure protein was 1.7 mg per litre of culture. Purified TLK2\_KD concentrated to 5 mg ml<sup>-1</sup> produced crystals at 277 K under two different conditions,

resulting in crystals belonging to two different space groups. The crystals grown using Natrix condition No. 1 diffracted to 3.7 Å resolution and belonged to space group *I*4<sub>1</sub>22, with unit-cell parameters *a* = *b* = 88.36, *c* = 342.87 Å, α = β = γ = 90°, and the crystals grown under Wizard II condition No. 37 diffracted to 3.4 Å resolution and belonged to space group *P*2<sub>1</sub>3, with unit-cell parameters *a* = *b* = *c* = 126.05 Å, α = β = γ = 90° (Fig. 3). Cell-content calculations (Matthews, 1968) predict that each asymmetric unit contains one molecule of TLK2\_KD (*V*<sub>M</sub> = 4.60 Å<sup>3</sup> Da<sup>-1</sup> for crystal type A and *V*<sub>M</sub> = 4.59 Å<sup>3</sup> Da<sup>-1</sup> for crystal type B), with approximately 73% solvent in both cases, values that are consistent with the fragility of the crystals. Unfortunately, dehydration trials did not help to improve either their robustness or their resolution. Structure solution was obtained by molecular replacement using *Phaser* (McCoy *et al.*, 2007) with a search model generated by *I-TASSER* (Roy *et al.*, 2010) that was mainly based on the structure of Aurora A kinase (PDB entry 3h10; Aliagas-Martin *et al.*, 2009), to which TLK2\_KD is predicted to have a similar secondary structure (46% sequence homology). Refinement and model building are ongoing.

We thank the Swiss Light Source and ALBA beamline staff for their support. This work was supported by the Ministerio de Economía y Competitividad BFU2011-23815/BMC, the Fundación Ramón Areces and the Comunidad Autónoma de Madrid (CAM-S2010/BMD-2305 to GM). AMG is supported by a La Caixa fellowship.

#### References

- Aliagas-Martin, I. *et al.* (2009). *J. Med. Chem.* **52**, 3300–3307.
- Canfield, C., Rains, J. & De Benedetti, A. (2009). *BMC Mol. Biol.* **10**, 110.
- Dillon, P. J., Gregory, S. M., Tamburro, K., Sanders, M. K., Johnson, G. L., Raab-Traub, N., Dittmer, D. P. & Damania, B. (2013). *Cell Host Microbe*, **13**, 204–214.
- Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C. & Meinnel, T. (2006). *Mol. Cell. Proteomics*, **5**, 2336–2349.
- Hanks, S. K., Quinn, A. M. & Hunter, T. (1988). *Science*, **241**, 42–52.
- Hashimoto, M., Matsui, T., Iwabuchi, K. & Date, T. (2008). *Mutat. Res.* **657**, 63–67.
- Heidenblad, M., Lindgren, D., Veltman, J. A., Jonson, T., Mahlamäki, E. H., Gorunova, L., van Kessel, A. G., Schoenmakers, E. F. & Höglund, M. (2005). *Oncogene*, **24**, 1794–1801.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Krause, D. R., Jonnalagadda, J. C., Gatei, M. H., Sillje, H. H. W., Zhou, B.-B., Nigg, E. A. & Khanna, K. (2003). *Oncogene*, **22**, 5927–5937.
- Li, Y., DeFatta, R., Anthony, C., Sunavala, G. & De Benedetti, A. (2001). *Oncogene*, **20**, 726–738.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- 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.
- Moshkin, Y. M., Kan, T. W., Goodfellow, H., Bezstarosti, K., Maeda, R. K., Pilyugin, M., Karch, F., Bray, S. J., Demmers, J. A. A. & Verrijzer, C. P. (2009). *Mol. Cell*, **35**, 782–793.
- Roe, J. L., Durfee, T., Zupan, J. R., Repetti, P. P., McLean, B. G. & Zambryski, P. C. (1997). *J. Biol. Chem.* **272**, 5838–5845.
- Roe, J. L., Nemhauser, J. L. & Zambryski, P. C. (1997). *Plant Cell*, **9**, 335–353.
- Roe, J. L., Rivin, C. J., Sessions, R. A., Feldmann, K. A. & Zambryski, P. C. (1993). *Cell*, **75**, 939–950.
- Ronald, S., Sunavala-Dossabhoy, G., Adams, L., Williams, B. & De Benedetti, A. (2011). *Prostate*, **71**, 1367–1373.
- Roy, A., Kucukural, A. & Zhang, Y. (2010). *Nature Protoc.* **5**, 725–738.
- Sillje, H. H. & Nigg, E. A. (2001). *Curr. Biol.* **11**, 1068–1073.
- Sillje, H. H., Takahashi, K., Tanaka, K., Van Houwe, G. & Nigg, E. A. (1999). *EMBO J.* **18**, 5691–5702.
- Sunavala-Dossabhoy, G., Li, Y., Williams, B. & De Benedetti, A. (2003). *BMC Cell Biol.* **4**, 16.
- Takahata, S., Yu, Y. & Stillman, D. J. (2009). *EMBO J.* **28**, 3378–3389.
- Wang, Q. *et al.* (2006). *Cancer Res.* **66**, 6050–6062.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yap, D. B., Walker, D. C., Prentice, L. M., McKinney, S., Turashvili, G., Mooslehner-Allen, K., de Algora, T. R., Fee, J., de Tassigny, X., Colledge, W. H. & Aparicio, S. (2011). *PLoS One*, **6**, e27127.