Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of wild-type Plk-1 kinase domain in complex with a selective DARPin

Tiago M. Bandeiras,[a]‡ Roman Christian Hillig,[b]*‡ Pedro M. Matias,[c] Uwe Eberspaecher,[b] Jörg Fanghänel,[b] Mónica Thomaz,[a] Sandra Miranda,[a] Kerstin Crusius,[b] Vera Pütter,[b] Patrick Amstutz,[d] Maya Gulotti-Georgieva,[d] H. Kaspar Binz,[d] Caterina Holz,[e]§ Arndt A. P. Schmitz,[b] Christine Lang,[e] Peter Donner,[b] Ursula Egner,[b] Maria A. Carrondo[c]* and Beate Müller-Tiemann[b]

[a]Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780 Oeiras, Portugal, [b]Bayer Schering Pharma AG, Global Drug Discovery, 13342 Berlin, Germany, [c]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal, [d]Molecular Partners AG, Grabenstrasse 11a, 8952 Zürich-Schlieren, Switzerland, and [e]OrganoBalance GmbH, Gustav-Meyer-Allee 25, 13355 Berlin, Germany

‡ These authors contributed equally to this work.
§ Present address: PSF Biotech AG, Robert-Rössle-Strasse 10, 13125 Berlin, Germany.

Correspondence e-mail: roman.hillig@bayerhealthcare.com, carrondo@itqb.unl.pt

Received 16 October 2007
Accepted 22 December 2007

PDB Reference: kinase domain of Plk-1, 2v5q, r2v5qsf.

As a key regulator of mitosis, the Ser/Thr protein polo-like kinase-1 (Plk-1) is a well validated drug target in cancer therapy. In order to enable structure-guided drug design, determination of the crystal structure of the kinase domain of Plk-1 was attempted. Using a multi-parallel cloning and expression approach, a set of length variants were identified which could be expressed in large amounts from insect cells and which could be purified to high purity. However, all attempts to crystallize these constructs failed. Crystals were ultimately obtained by generating designed ankyrin-repeat proteins (DARPins) selective for Plk-1 and using them for cocrystallization. Here, the first crystal structure of the kinase domain of wild-type apo Plk-1, in complex with DARPin 3H10, is presented, underlining the power of selective DARPins as crystallization tools. The structure was refined to 2.3 Å resolution and shows the active conformation of Plk-1. It broadens the basis for modelling and cocrystallization studies for drug design. The binding epitope of 3H10 is rich in arginine, glutamine and lysine residues, suggesting that the DARPin enabled crystallization by masking a surface patch which is unfavourable for crystal contact formation. Based on the packing observed in the crystal, a truncated DARPin variant was designed which showed improved binding characteristics.

1. Introduction

Human polo-like kinase 1 (Plk-1), a serine/threonine protein kinase, is a key regulator of progression through mitosis (van de Weerdt & Medema, 2006). It is the best characterized member of the Plk family, which in man also comprises Plk-2, Plk-3 and Plk-4. However, Plk-1 appears to fulfil most of the known functions of the single Plk that is present in invertebrates (Barr *et al.*, 2004).

The expression, activity and localization of Plk-1 are dynamically regulated during the cell cycle. Plk-1 expression increases from the late S phase to mitosis, the phase when Plk-1 is most active (Golsteyn *et al.*, 1995; Hamanaka *et al.*, 1995; Lee *et al.*, 1995). Plk-1 has been implicated in centrosome maturation, spindle formation, sister chromatid separation, promotion of anaphase onset and, finally, mitotic exit and cytokinesis (Barr *et al.*, 2004). All four members of the Plk family comprise an N-terminal kinase domain and two conserved polo-box sequence motifs (only one in Plk-4). The two polo-boxes of Plk-1 fold into one structural unit, the polo-box domain (PBD; Elia, Rellos *et al.*, 2003; Cheng *et al.*, 2003), which serves as a phosphopeptide-recognition domain recruiting the respective Plk to its cellular targets and/or substrates (Elia, Cantley *et al.*, 2003). The single polo-box motif of Plk-4 forms an intermolecular homodimer (Leung *et*

al., 2002) which, however, is not expected to recognize phosphopeptides (Elia, Rellos *et al.*, 2003).

Plk-1 is overexpressed in highly proliferating cells and in many human tumours, where this is often associated with a poor prognosis (Eckerdt *et al.*, 2005; Takai *et al.*, 2005). It has therefore been considered to be a very attractive target for cancer drug development (Strebhardt & Ullrich, 2006). Several small-molecule inhibitors have been described (Schwede *et al.*, 2003; Bearss *et al.*, 2006; Davis-Ward *et al.*, 2004; Steegmaier *et al.*, 2007; Andrews *et al.*, 2004) and reviewed (Strebhardt & Ullrich, 2006; McInnes *et al.*, 2005). Despite its thorough validation as a cancer drug target, successful production of soluble Plk-1 kinase domain and crystal structure determination have only been reported very recently and only for a mutated form of Plk-1 (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007). We found that production of the Plk-1 kinase domain based on bona fide bioinformatically predicted domain boundaries resulted in insoluble protein and poor expression levels. Following a multi-parallel cloning, purification and crystallization approach for screening a large variety of length variants (Malawski *et al.*, 2006), we have now identified alternative Plk-1 kinase-domain constructs that result in high-level expression of recombinant soluble Plk-1 kinase domain and can be purified to homogeneity. However, we only obtained crystals when designed ankyrin-repeat proteins (DARPins) selected for one of these Plk-1 constructs were generated and used as cocrystallization agents. Selective DARPins have recently been generated for a range of target proteins (Binz *et al.*, 2004; Stumpp & Amstutz, 2007) and have successfully been used in cocrystallization experiments (Binz *et al.*, 2004; Amstutz *et al.*, 2005; Schweizer *et al.*, 2007; Sennhauser *et al.*, 2006).

Here, we present the cloning, expression and purification of Plk-1 constructs suitable for crystallization, the generation of a selective DARPin which binds to Plk-1 with high affinity and enabled crystallization and the crystal structure determination of wild-type Plk-1.

2. Materials and methods

2.1. Cloning and expression of Plk-1 kinase domain constructs

Recombinant baculovirus expressing full-length His-tagged human Plk-1 (NM_05030, residues 1–603) was produced using the BaculoGold kit (BD Pharmingen). Based on the results of limited proteolysis, different-length constructs of the kinase domain of human Plk-1, both with and without the activation-segment mutation T210D (Fig. 1), were cloned using the Gateway System (Invitrogen). The mutations at nucleotides 628/629, ACC/GAC (T210D) were introduced using the QuikChange kit (Stratagene). The forward primers used fuse a TOPO cloning signal and a thrombin cleavage site (5′-CACCCTGGTTCCGCGTGGATCC-3′) to the 5′ end of the Plk-1 N-terminal sequence starting at the residues depicted in Fig. 1. The reverse primers introduced a stop codon (5′-TTA-3′) after the C-terminal Plk-1 residues shown in Fig. 1. One forward primer was paired with one reverse primer in the combinations shown in Fig. 1 and employed in a PCR reaction with a Plk-1 WT or mutant template and the Advantage-HF 2 PCR kit (Clontech). The resulting separate amplicons were cloned into pENTR/D-TOPO (Invitrogen) and their sequences were verified.

For the expression of GST-tagged or His-tagged Plk-1 kinase in *Escherichia coli*, the Gateway destination vectors pD-ECO1 or pD-ECO3 were used as previously described (Malawski *et al.*, 2006) after recombination-cloning the set of Plk-1 pENTRY plasmids using Gateway LR Clonase Enzyme Mix (Invitrogen). For expression of the same GST-tagged Plk-1 kinase constructs in insect cells, the Plk-1 plasmids were recombination-cloned into pD-INX1 and subsequently into pD-INS1. In order to construct vector pD-INX1, the GST sequence of pGEX-KT (GE Healthcare) was PCR-amplified while introducing *Eco*47III sites at both ends as well as an optimized Kozak sequence (5′-CCCAGCGCTAAAAT-GAGCCCTATACTAGGTTATTGG-3′; 5′-ATCCAGCGC-TATCCGATTTTGGAGGATGG-3′). The amplicon was cloned into the *Eco*47III site of pXINSECT-DEST38 (Invitrogen). For the construction of pD-INS1, the GST sequence of pGEX-KT (GE Healthcare) was PCR-amplified while introducing *Bam*HI and *Sma*I sites at the ends as well as an optimized Kozak sequence (5′-CGCGGATCCAAAAT-GAGCCCTATACTAGGTTATTGG-3′; 5′-GTACCCGGGG-TCCGATTTTGGAGGATGGTCG-3′). The terminal restriction sites were used to ligate the amplicon into pVL1393 (Becton Dickinson). The Reading Frame Cassette A of the Gateway Vector Conversion System (Invitrogen) was then introduced into the *Sma*I site of this vector.

E. coli expression was performed in BL21-CodonPlus (DE3)-RIL cells (Stratagene). Cultures were grown in LB_{Amp} medium to an OD_{600} of 0.7–0.8. Protein expression was induced by the addition of 0.5 m*M* IPTG. Cells were harvested

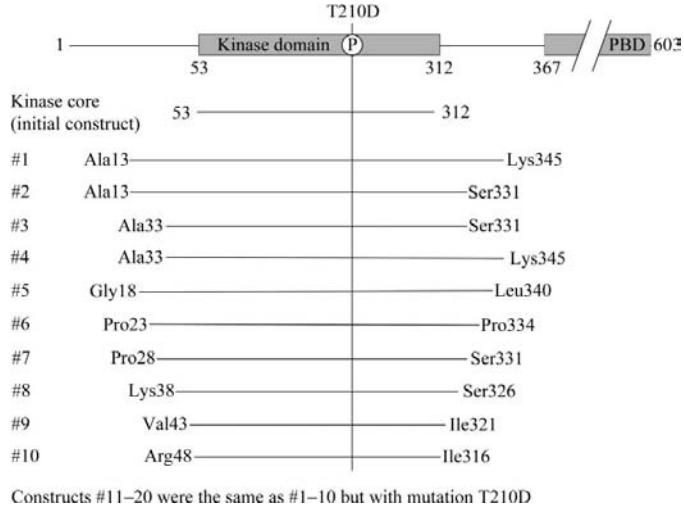

Figure 1
Plk-1 construct design and examples of multi-parallel expression tests. The domain structure of human Plk-1 is shown, with P indicating the phosphorylation site Thr210. Below, an initial construct and a set of 20 length variants (#1–#20) is shown. Residue numbers for the N- and C-terminal amino acids are indicated.

by centrifugation after 20 h at 288 K and lysed as described by Malawski *et al.* (2006).

Transient insect-cell expression was performed by transfection of High Five cells (Invitrogen) with the respective pD-INX1 expression constructs. High Five cells in the logarithmic growth phase cultivated in Insect Xpress (BioWhittaker) cell-culture medium were seeded at a cell density of 1.5×10^6 cells in six-well dishes and transfected with 10 µl Fugene6 and 2 µg DNA per well. Cells were harvested by centrifugation after 72 h and extracted in lysis buffer consisting of 50 m*M* Tris–HCl pH 8.0, 120 m*M* NaCl, 10%(*v*/*v*) glycerol, 0.5%(*v*/*v*) NP40, 5 m*M* DTT with Complete protease-inhibitor cocktail 1× (Roche).

Expression levels were determined by enriching the Plk-1 kinase-domain variants by affinity chromatography using either MicroSpin GST Purification Modules (GE Healthcare) following the manufacturer's instructions or on chelating columns. The expression levels of the His-tagged proteins were analyzed after separation of the soluble and insoluble fractions by centrifugation. Expression of the six shortest well expressed length variants (constructs #3–#8) and of mutant #14 was then up-scaled in insect cells by generating recombinant baculoviruses using the BaculoGold system (Becton Dickinson). Transfer vectors were generated by recombination-cloning of the respective entry clones into pD-INS1. For protein expression, High Five cells at 2×10^6 cells ml^{-1} were infected with recombinant viruses at an MOI of 3 pfu per cell and the cells were harvested by centrifugation after 48 h.

2.2. Purification and crystallization screening of Plk-1 constructs

Cell pellets were lyzed by thawing and the Plk-1 variants were purified *via* GSH affinity chromatography, on-column thrombin-tag cleavage, ion-exchange chromatography, gel filtration and hydrophobic interaction chromatography following standard procedures. Sample purity was checked by SDS–PAGE and mass spectrometry (MS) and samples were concentrated to 8–18 mg ml^{-1} and used directly for crystallization trials. Despite intensive screening, no protein crystals could be obtained.

2.3. Generation and characterization of Plk-1-specific DARPins

DARPins specific for Plk-1#4 were generated as described by Binz *et al.* (2004). In brief, GST-Plk-1#4 and GST alone (as a control) were immobilized on GSH plates (Pierce). Immobilization was confirmed by ELISA using a mouse anti-GST antibody (Sigma) or a mouse anti-Plk-1 antibody as a primary detection reagent and a goat anti-mouse-IgG-alkaline phosphatase conjugate (Qiagen) as a secondary detection reagent (data not shown). Ribosome-display selections on immobilized GST-Plk-1#4 without substrate were performed for four consecutive rounds including prepanning steps on immobilized GST, as has been described previously (Binz *et al.*, 2004). The fourth selection round comprised either a competitive elution step (using 2.5 µ*M* GST-free Plk-1 for 90 min) for complex elution or extensive washing (seven washes over 2 h) prior to standard ribosomal complex disruption by EDTA (Binz *et al.*, 2004). Individual binders of the third and the fourth selection round were then screened by crude *E. coli* extract ELISA for GST-Plk-1 binding as described previously (Binz *et al.*, 2004). Several binders were identified, of which 12 were characterized in more detail. These proteins were expressed in *E. coli* XL-1 Blue and purified on Ni–NTA columns (Qiagen) as described by Binz *et al.* (2004). The affinities and specificities of these proteins were analyzed by ELISA. Specificity was tested comparing the binding signal of 50 n*M* of the respective DARPin on immobilized GST-Plk-1#4 with that of immobilized GST only. The affinity was analyzed by competition ELISA by measuring the binding signal of 50 n*M* of the respective DARPin on immobilized GST-Plk-1#4 in the presence or absence of varying concentrations of free Plk-1#4 (0, 75 or 200 n*M* Plk-1#4; data not shown). Five of these binders showed nanomolar affinity and binding could be competed, indicating high target specificity (data not shown).

2.4. Generation and purification of truncated DARPin 3H10Δ

The DNA sequence encoding DARPin 3H10Δ (residues 1–143) was PCR-amplified from pQE 30 3H10 [using primers which incorporated an *Nde*I site at the 5′ end of the gene (5′-GAACATATGGACCTGGGTAAGAAACTGCTG-3′) and a *Xho*I site at the 3′ end (5′-GAGCTCGAGTCAGTCC TGA-GCGTTAACGTCAGC-3′)], ligated into vector pET-15b and transformed into *E. coli* BL21(DE3) cells. Expression was carried out in CircleGrow medium (200 mg l^{-1} ampicillin) at 310 K. Expression was induced at an OD_{600} of 1.0 (100 µ*M* isopropyl β-D-thiopyranogalactoside; IPTG) and the cells were harvested after 5 h and stored at 253 K.

Cells were resuspended in buffer *A* (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 40 m*M* imidazole, 10% glycerol, 1 m*M* DTT with one Complete protease-inhibitor cocktail tablet), disrupted using a Microfluidizer and centrifuged. The supernatant was loaded onto a 5 ml His-Trap column (GE Healthcare) which was washed with ten column volumes (CV) of buffer *A* and ten CV of cleavage buffer (50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 0.5 m*M* EDTA, 10% glycerol). After on-column cleavage (1 µg thrombin per milligram of 3H10Δ at 279 K overnight), the column was washed with five CV of buffer *A* to elute the cleaved protein. Peak fractions were pooled and purified by gel filtration [16/60 Superdex 75 (GE Healthcare); 50 m*M* Tris–HCl pH 7.5, 400 m*M* NaCl, 20 m*M* $MgCl_2$, 1 m*M* DTT]. The expression yield of 3H10Δ was 30 mg per litre of culture.

2.5. *In vitro* kinase assay

To determine the influence of DARPins on Plk-1 activity, scintillation proximity assays (SPA) were performed for 90 min at 295 K in the presence of a tenfold molar excess of DARPins in a total volume of 31 µl using casein from bovine milk (Sigma) as substrate [0.66 µg ml^{-1} full-length His-tagged Plk-1, 0.7 µ*M* biotinylated casein, 50 m*M* HEPES pH 7.5,

10 m*M* $MgCl_2$, 3 m*M* $MnCl_2$, 1 m*M* DTT, 0.01%(*v*/*v*) Nonidet P40, 3% DMSO, 0.5 µ*M* ATP, 50 nCi γ-^{33}P-ATP]. Reactions were terminated by the addition of 50 µl SPA suspension [100 µ*M* ATP, 10 m*M* EDTA, 0.2%(*v*/*v*) Triton X-100, 2.5 mg ml^{-1} streptavidin-coated SPA beads (Amersham Biosciences)] in phosphate-buffered saline (0.138 *M* NaCl, 0.0027 *M* KCl pH 7.4; Sigma P3813-10PAK). SPA beads were allowed to sediment overnight at 295 K. The incorporated ^{33}P was determined using a TopCount scintillation counter (Perkin–Elmer). The experiment was carried out three times for each DARPin.

2.6. Isothermal titration calorimetry (ITC)

ITC experiments were performed with a MicroCal VP-ITC at 298 K. All sample buffers were exchanged using PD-10 columns (GE Healthcare) in order to ensure identical buffer conditions for both interaction partners. Protein concentrations were determined by UV–Vis spectrometry (Nanodrop) using calculated molar extinction coefficients (ε_{280} = 18 490 M^{-1} cm^{-1} for Plk-1#4 and ε_{280} = 3840 M^{-1} cm^{-1} for both DARPin 3H10 and 3H10Δ). The staurosporine sample was prepared by diluting a 10 m*M* DMSO stock solution into the protein buffer. All titration experiments were carried out in 50 m*M* Tris pH 8.0, 300 m*M* NaCl, 10%(*v*/*v*) glycerin and 0.1 m*M* EDTA. The titration protocol consisted of 18 injections of 15 µl Plk-1 (108 µ*M*) titrated into a sample cell filled with 10 µ*M* DARPin 3H10, 3H10Δ or staurosporine (stirring speed 310 rev min^{-1}, feedback-gain mode set to high). Control experiments were performed to measure the heat of dilution of Plk-1; to determine the thermodynamic values, the heat of dilution was subtracted from the heat evolved during the respective titration experiment. The data were analyzed using the *ORIGIN* software provided by the manufacturer of the instrument (MicroCal). Binding isotherms were fitted according to a single-site binding model.

2.7. Plk-1#4 expression and purification for DARPin cocrystallization

Protein production was performed in a 10 l stirred bioreactor (B. Braun, Germany) inoculated with High Five cells grown in Insect X-Press media (Cambrex, US) in a 2 l bioreactor at a cell concentration of 3–4 × 10^5 cells ml^{-1} with Pluronic F-68 added to the medium to a final concentration of 2 g l^{-1}. The temperature was constant at 301 K and $p(O_2)$ was maintained above 30% by varying the agitation rate between 60 and 185 rev min^{-1}. The aeration was adjusted to supply a mixture of air, oxygen and nitrogen at a flow rate of 100 ml min^{-1}. Infection was performed at a cell concentration of 2–3 × 10^6 cells ml^{-1} using an MOI of 2–5 pfu per cell. Cells were harvested 48 h after infection by centrifugation at 120*g*. 30 g of the cell pellet was resuspended in 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 10%(*v*/*v*) glycerol, 0.5 m*M* EDTA, 5 m*M* DTT and disrupted twice in a continuous high-pressure homogenizer (50 MPa). The supernatant from ultracentrifugation was loaded onto 20 ml Gluthathione Sepharose 4B (GE Healthcare) and washed with the same buffer followed by 50 m*M* Tris–HCl pH 8.0, 10%(*v*/*v*) glycerol, 300 m*M* NaCl, 1 m*M* DTT. After opening and addition of 1500 units of thrombin (GE Healthcare), the column was left rotating overnight at 277 K. The washing step was restarted and the flowthrough containing the cleaved Plk-1#4 was collected and passed onto a Benzamidine Sepharose 4 FF (GE Healthcare) column to remove thrombin. Plk-1#4 was further purified using a Hiload Superdex 75 column (GE Healthcare) and concentrated to 3.3 mg ml^{-1} in an Amicon Ultra4 (10 kDa cutoff; Bradford).

2.8. Expression of DARPin 3H10 and purification for crystallization

Fermentation using the corresponding *E. coli* strain expressing DARPin 3H10 was carried out in LB medium in the presence of ampicillin (200 mg ml^{-1}). Expression was induced with 500 m*M* IPTG and the fermentation (10 l) was continued at 310 K for 4 h. Cells were harvested by centrifugation at 3000*g*. 30 g of the cell pellet was disrupted twice in a continuous high-pressure homogeniser (90 MPa) in a buffer containing 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl. The supernatant from the ultracentrifugation was loaded onto 80 ml Chelating Sepharose FF (GE Healthcare) packed in a XK26/20 column, washed with 50 m*M* Tris–HCl pH 8.0, 10%(*v*/*v*) glycerol, 150 m*M* NaCl. The protein was eluted with a linear two CV gradient to 500 m*M* imidazole, desalted using a Sephadex G25 column (GE Healthcare) to 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 10%(*v*/*v*) glycerol, 0.05%(*v*/*v*) *n*-octyl-β-glucoside (βOG), 5 m*M* DTT and concentrated to 7.7 mg ml^{-1} (Bradford).

2.9. Formation and crystallization of the complex of Plk-1#4 and DARPin 3H10

Freshly purified Plk-1#4 and 3H10 were mixed in a 1:1 molar ratio and incubated overnight at 277 K with stirring. The solution was filtered (0.2 mm), purified by gel filtration (Hiload Superdex 75) and concentrated to 12 mg ml^{-1} in an Amicon Ultra4 (10 kDa cutoff; Bradford). The concentrated Plk-1–3H10 sample in 50 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 10%(*v*/*v*) glycerol, 0.05%(*v*/*v*) βOG and 5 m*M* DTT was used for crystallization trials with a Cartesian robot using the vapour-diffusion method at 277 and 293 K in 96-well plates (Molecular Dimensions). Drops were equilibrated against 576 conditions from the crystallization screens Classics Lite, pHClear and Cryos (Nextal), NR-LBD+Extension and MemStart/Sys (Molecular Dimensions) and Index (Hampton Research). Small thin crystal plates appeared after 4 d in condition No. 79 of Index Screen [0.1 *M* bis-tris pH 6.5, 0.2 *M* ammonium acetate, 25%(*w*/*v*) PEG 3350]. Optimization was carried out in drops made up from 1.5 µl protein solution and 1.5 µl reservoir solution in 24-well plates (Hampton Research) and included changes in PEG, buffers and pH, additive screening and different temperatures, resulting in higher quality crystals that grew from 0.1 *M* Tris–HCl pH 8.0, 8%(*w*/*v*) PEG 5000 MME, 0.01 *M* EDTA sodium salt at 303 K. Obtaining well diffracting crystals from this condition

required multiple seeding steps. Initial thin crystal plates were used for streak-seeding into 2 d-old drops containing no crystals. This produced well shaped crystals along with thin plates. The best crystals were used to streak-seed into 4 h-old drops, yielding only well shaped crystals which appeared within 1 d and continued to grow for two weeks to a maximum dimension of 500 µm.

2.10. Diffraction data collection

Crystals of the Plk-1#4–DARPin 3H10 complex were incubated using a stepwise approach in a mother-liquor solution containing increasing concentrations of glycerol [5%, 15%, 25%(*v*/*v*), 3 min per step] and flash-frozen under a stream of nitrogen gas at 100 K. An initial diffraction data set was collected at ESRF beamline ID23-2 using a MAR Research 225 CCD detector. The images were integrated to 3.25 Å resolution with *MOSFLM* (Leslie, 1992). A second diffraction data set was subsequently collected at the ESRF beamline ID29 using an ADSC Quantum 315 detector and the images were processed to 2.3 Å resolution with *MOSFLM*. For both data sets, data scaling, merging and intensity conversion to structure-factor amplitudes were carried out with *SCALA* and *TRUNCATE* from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). The crystal data and a summary of the data-collection statistics are listed in Table 1.

2.11. Structure determination

The structure was solved by molecular replacement (MR) using the program *Phaser* (Storoni *et al.*, 2004). For the DARPin search model, coordinates from PDB entries 1mj0 (Kohl *et al.*, 2003) and 1svx (Binz *et al.*, 2004) were used. As no crystal structure of Plk-1 was available at the time, a series of homology models were generated. Sequence comparisons revealed Aurora-A, PKB and PDK1 to be the closest homologues of known structure with respect to the overall kinase domain (32–30% sequence identity) and B-Raf to be the closest homologue with respect to the ATP site (47% identity). Homology models were calculated with *MODELLER* (Marti-Renom *et al.*, 2000) using default settings as implemented in *Discovery Studio* (Accelrys Inc., San Diego, USA). The structures of one PDK1–inhibitor complex and of two Aurora-A–inhibitor complexes were used as templates, since these two kinases feature a leucine gatekeeper as observed in Plk-1 [PDK1, PDB entry 1uu9 (Komander *et al.*, 2004); Aurora-A, active and inactive conformations, PDB entries 1ol5 (Bayliss *et al.*, 2003) and 2bmc (Fancelli *et al.*, 2005), respectively]. In addition, homology models were built based on B-Raf (PDB code 1uwh; Wan *et al.*, 2004) because it shares with Plk-1 the presence of a phenylalanine residue underneath the adenine base. Initial MR calculations were carried out using the standard *Phaser* protocol (*CCP*4*i*; Potterton *et al.*, 2003) and the 3.25 Å data set. The solutions indicated that the asymmetric unit consisted of a heterotetramer formed by two Plk-1#4–DARPin 3H10 heterodimers related by a noncrystallographic twofold rotation axis. Inspection of the MR solution revealed that the DARPin molecules were packed tail-to-tail, with extensive clashing of the C^{α} atoms in the C-terminal capping ankyrin repeats of the full DARPin 3H10 search models, indicating that these C-terminal regions were oriented differently or perhaps even disordered. This prompted modified search models for DARPin 3H10 in which the C-terminal capping repeat was truncated after residue 141 to avoid C^{α} clashes.

The best MR solutions were obtained using the homology model from 1ol5 for Plk-1 molecule *A*, the homology model from 1uu9 for Plk-1 molecule *B*, the truncated DARPin from 1mj0 for DARPin 3H10 molecule *C* and the truncated DARPin from 1svx for DARPin 3H10 molecule *D*. The electron density was further improved using *ARP*/*wARP* in atom-update and refinement mode (Perrakis *et al.*, 2001). However, the homology models for Plk-1 lacked significant portions of both the N- and C-terminal regions and although some electron density for the missing residues could be seen in the electron-density maps, it was deemed doubtful that the 3.25 Å

Table 1
Data-collection and refinement statistics for Plk-1–DARPin 3H10.

Values in parentheses are for the highest resolution shell.

Crystal data		
Beamline	ESRF ID23-2	ESRF ID29
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)		
a	61.53	62.33
b	136.28	135.22
c	140.13	136.82
Data collection		
Wavelength (Å)	0.8726	1.037
Resolution (Å)	70.01–3.25 (3.43–3.25)	67.60–2.30 (2.42–2.30)
$R_{\rm merge}$	0.081 (0.333)	0.062 (0.401)
$I/\sigma(I)$	6.3 (2.1)	7.4 (1.7)
Observations	141780 (20739)	180727 (26660)
Unique reflections	19276 (2749)	51654 (7427)
Completeness (%)	100.0 (100.0)	99.3 (99.3)
Multiplicity	7.4 (7.5)	3.5 (3.6)
Estimated $B_{\rm overall}$ (Å^2)	104.2	45.7
Refinement		
Resolution (Å)		61.1–2.30
No. of reflections		49051
$R_{\rm work}/R_{\rm free}$†		0.181/0.225
No. of atoms		
Protein		6559
Water		406
R.m.s. deviations		
Bond lengths (Å)		0.011
Bond angles (°)		1.26
Mean *B* factors‡ (Å^2)		
Protein main chain/side chain		
Plk-1 *A*		41.4/44.9
Plk-1 *B*		46.2/50.9
DARPin *C*		35.7/39.9
DARPin *D*		50.0/53.8
Solvent		40.7
PROCHECK G factors		
Plk-1 *A*		0.13
Plk-1 *B*		0.11
DARPin *C*		0.24
DARPin *D*		0.25

† $R_{\rm free}$ was calculated from a random sample containing 5% of the total number of independent reflections measured (Brünger, 1992). ‡ Mean *B* factors were calculated from equivalent isotropic *B* values, including the TLS contribution for the protein atoms.

data could be used to successfully build a more complete Plk-1 structural model.

Careful fine-tuning of the crystallization conditions and the use of streak-seeding then produced larger crystals. A second data set was measured from one of these to 2.3 Å resolution. The MR procedure was repeated, using the homology model from 1ol5 for Plk-1 molecules *A* and *B* and the DARPin from 1mj0 truncated after residue 141 for DARPin 3H10 molecules *C* and *D*. The electron density was further improved by using *ARP*/*wARP* in atom-update and refinement mode (Perrakis *et al.*, 2001) and a partial model was automatically built using *ARP*/*wARP* in automated model-building mode starting from the MR phases (Perrakis *et al.*, 1999). This improved the initial *Phaser* figure of merit of 0.435 to 2.3 Å resolution to 0.673 (*ARP*/*wARP* in atom-update and refinement mode) and 0.695 (*ARP*/*wARP* in automated model-building mode), respectively. The resulting maps were used to correct the model prior to refinement.

2.12. Structure refinement

Refinement was carried out with the program *REFMAC* (Murshudov *et al.*, 1997) from the *CCP*4 suite. At a later stage, a translation/libration/screw (TLS; Schomaker & Trueblood, 1968) rigid-body motion refinement was performed prior to restrained refinement of atomic positions and thermal motion parameters. One rigid body was defined for each domain in each Plk-1 molecule (domain 1 comprised residues 39–131 and domain 2 included residues 132–323) and for each DARPin 3H10 molecule. In the final refinement stages, a total of 411 solvent molecules were located with *ARP*/*wARP* (Lamzin & Wilson, 1993) and individual restrained *B* factors were refined for all non-H atoms. The final values of R and $R_{\rm free}$ (Brünger, 1992) were 0.182 and 0.224, respectively. Throughout the refinement, the model was periodically checked and corrected with *TURBO* (Roussel *et al.*, 1990) and *Coot* (Emsley & Cowtan, 2004) against $2|F_{\rm o}| - |F_{\rm c}|$ and $|F_{\rm o}| - |F_{\rm c}|$ electron-density maps. The maximum-likelihood estimate of the overall coordinate error obtained with *REFMAC* was 0.15 Å. For both independent Plk-1 monomers only residues 39–323 were visible in the electron density, while residues 12–141 could be seen for DARPin molecule *C* and residues 13–141 for DARPin molecule *D*. In addition, a total of 18 side-chain atoms could not be seen in the final electron-density maps and were assigned zero occupancy. The main refinement statistics are presented in Table 1.

The structure was validated using *PROCHECK* (Laskowski *et al.*, 1993). The Ramachandran (ψ, φ) plot (Ramachandran & Sasisekharan, 1968) showed only three residues (of 250) outside the most favoured regions for Plk-1 molecule *A* (Lys91, Arg136 and Lys146) and only two (of 250) for Plk-1 molecule *B* (Arg136 and Lys146). The unusual (ψ, φ) conformation of Arg136 and Lys146 in both Plk-1 monomers is supported by their well defined electron density. However, Lys91 in Plk-1 molecule *A* is located in a loop region with poor electron density. All five residues are located in the so-called 'generously allowed regions' and are only marginally outside the most favoured regions.

3. Results

3.1. Construct design

An initial Plk-1 construct comprised residues 53–312 (Fig. 1) and was designed as the minimal kinase domain based on alignments with other kinases of known structure. This construct was cloned and expressed in *E. coli* and insect cells using either an N-terminal GST or His tag (data not shown). However, none of these approaches yielded sufficient amounts of soluble protein.

In order to obtain more suitable constructs, limited proteolysis experiments were carried out using recombinant full-length Plk-1 and a series of proteases. Trypsin and elastase both produced stable fragments covering the kinase domain. N-terminal sequencing and MS revealed that trypsin cleaved N-terminally after Arg12 and C-terminally after Lys345, whereas elastase cleaved after Ala32 and Ser331. Starting from these borders, further length variants were designed by truncation mutagenesis, in which the termini were gradually deleted in steps of five amino acids (constructs #1–#10; Fig. 1). In an attempt to stabilize the activated conformation, phosphorylation at Thr210 in the activation segment was mimicked by introducing the mutation T210D in all ten length variants, resulting in construct set #11–#20.

3.2. Expression, purification and analysis of recombinant Plk-1 and DARPins

Multi-parallel expression tests with all 20 Plk-1 length variants showed that the GST-tagged proteins were generally better expressed than the His-tagged versions, which were therefore abandoned. After expression in *E. coli* followed by GSH affinity chromatography, all expressed constructs were contaminated with GroEL chaperone protein as confirmed by MS (Supplementary Fig. 1*a*[1]). GroEL could not be removed in subsequent purification steps even under harsh conditions such as the addition of 2 *M* urea to the purification buffers. This pointed towards non-optimal folding during bacterial expression and was overcome by expression in insect cells (Supplementary Fig. 1*c*). Expression tests in *Saccharomyces cerevisiae* and *Pichia pastoris* produced chaperone-free and soluble Plk-1 for *S. cerevisiae*, but significantly degraded proteins in *P. pastoris*, which was therefore abandoned (data not shown).

Insect cells and *S. cerevisiae* yielded consistent expression data in that only the longer constructs #1–#8 (and #11–#18) were produced in significant amounts as soluble proteins, indicating that constructs #9 and #10 (and #19 and #20) had been designed too short. *S. cerevisiae* was not further pursued as this expression system provided lower yields compared with insect cells. Based on the behaviour in test purifications and

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: HV5097). Services for accessing this material are described at the back of the journal.

Table 2
ITC characterization of the interaction of Plk-1#4 with staurosporine and two DARPins.

Ligand	ΔH (kJ mol^{-1})	$T\Delta S$ (kJ mol^{-1})	K_d (nM)	N
Staurosporine	−18.0	23.9	49	0.96
DARPin 3H10	∼4.6†	∼45.2†	71†	0.74
DARPin 3H10Δ	−28.9	∼26.8†	∼0.9†	0.87

† In these cases the thermodynamic parameters could not be determined with high accuracy owing to the small binding enthalpy and/or the high affinity of the interaction.

starting from the shortest well expressed and soluble constructs, constructs #3–#8 and subsequently the T210D mutant #14 were purified on a larger scale from insect cells. Supplementary Fig. 1(*b*) shows an SDS–PAGE gel with purified Plk-1#4 as a representative example.

To analyze the ability of the purified Plk-1 constructs to bind small molecules into their ATP-binding site, we employed isothermal titration calorimetry to thermodynamically follow the interaction with the unspecific kinase inhibitor staurosporine. Representative results are shown for Plk-1#4 (Fig. 2*a*, Table 2). The interaction was of exothermic nature and the high affinity turned out to result from comparable enthalpic and entropic contributions, indicative of specific polar interactions (such as hydrogen bonds) as well as hydrophobic interactions. The stoichiometry factor N of 0.96 showed that within the range of the accuracy of this experiment all Plk-1#4 molecules were capable of ligand binding, confirming a homogenous protein population. Five DARPins identified in the ribosome-display selections (termed 3H10, 2F9, 1C9, 1G7 and 1E1) were purified in high amounts as soluble and active proteins as indicated by gel-filtration and ELISA experiments (data not shown).

3.3. Characterization of the Plk-1–DARPin interaction

Complex formation between Plk-1#4 and the various DARPins was confirmed by gel-filtration experiments, which demonstrated a clear molecular-weight shift. *In vitro* kinase-activity assays showed that the Plk-1 activity was partially inhibited by binding of the DARPins (Fig. 3), suggesting that either the binding epitopes on Plk-1#4 partially overlapped with the active site causing competitive inhibition or that the DARPins acted as allosteric inhibitors. Additional incubation of all five DARPins with an ATP-site inhibitor of the thiazolidinone series (Schwede *et al.*, 2003; Santamaria *et al.*, 2007) revealed that with all DARPins inhibition could be further increased in the presence of the small-molecule inhibitor (data not shown), demonstrating that DARPin binding did not block ligand access to the ATP site.

Determination of the thermodynamic parameters of the Plk-1#4–3H10 interaction by isothermal titration calorimetry (ITC) showed that the reaction was slightly endothermic (Fig. 2*b* and Table 2). The observed small heat change during the reaction allowed only a rough estimation of the K_d value, which was determined to be 71 nM. Competition ELISA experiments estimated the K_d value of this interaction as approximately 75 nM and thereby confirmed the order of magnitude of the value obtained by ITC. Based on this estimation of K_d the binding reaction is endothermic (ΔH =

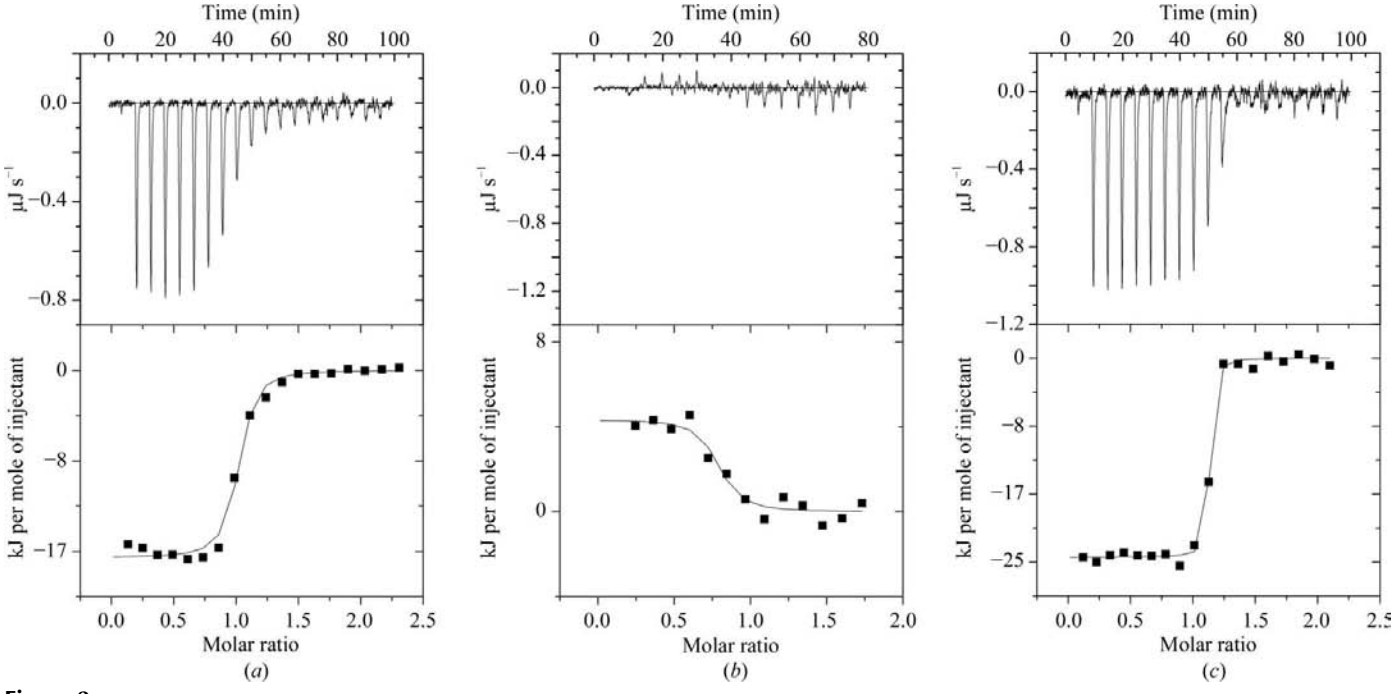

Figure 2
ITC characterization of Plk-1#4. Microcalorimetry titration curves for (*a*) staurosporine, (*b*) DARPin 3H10 and (*c*) DARPin 3H10Δ titrated with Plk-1#4. The derived thermodynamic parameters are listed in Table 2. The sample cell was filled with 10 µM of the respective Plk-1 interaction partner. Each peak results from the injection of 15 µl of a 108 µM solution of Plk-1#4 into the ligand solution (upper panels). Integration of titration signals resulted in ΔH_{ITC} values which were corrected for the heat of dilution and then plotted against the molar ratio between Plk-1#4 and the ligand (lower panels). The resulting titration curves were fitted to a single-site binding model by nonlinear least-squares analysis.

+4.6 J mol^{-1}) and is entirely driven by the entropic contribution (Table 2). The observed stoichiometry factors of $N = 0.74$ and 0.87 for DARPins 3H10 and 3H10Δ, respectively, agree with 1:1 complex formation.

In addition to full-length DARPin 3H10, the binding of a C-terminally truncated 3H10Δ (see below and §4) was characterized by ITC. The resulting titration curve (Fig. 2*c*) revealed that in this case binding was accompanied by a strong exothermic heat change. Nonlinear least-squares fit analysis of the titration data according to the single-site model gave a K_d value in the low nanomolar range (Table 2). Again, the high affinity of the interaction did not allow precise determination of the association constant under the conditions used here. Nevertheless, the binding enthalpy could be accurately determined as $\Delta H = -28.9$ kJ mol^{-1}.

3.4. Structure of wild-type Plk-1

The complex of Plk-1#4 and DARPin 3H10 was crystallized and the structure was solved by molecular replacement and refined to 2.3 Å resolution (Table 1). The structure of the kinase domain is shown in Fig. 4(*a*). It features the typical fold of a protein kinase, consisting of an N-terminal and a C-terminal lobe. The N-terminal lobe mainly consists of β-strands, with an incomplete β-barrel formed by six antiparallel strands (β0–β5), with a single α-helix (αC) nearly perpendicular to the barrel axis and packing against its side. The C-terminal lobe is formed almost entirely from α-helices. Its core is a four-helical antiparallel bundle (αE, αF, αH, αI) with a small two-stranded β-sheet (β6–β7) at the domain interface and two additional helices: αD at the bottom of the bundle and almost parallel to it and αG at the top and nearly perpendicular to the bundle.

A three-dimensional structure comparison of Plk-1 molecule *A* in the crystal structure against the PDB using the protein structure-comparison service *SSM* (Krissinel & Henrick, 2004) revealed that the four closest structural homologues were the three-dimensional structures of the recently published Plk-1 T210V mutant (in complex with two different ligands; PDB codes 2ou7 and 2owb; Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007) and the catalytic domain of protein kinase Aurora-A [in two crystal forms; PDB codes 1ol5 (Bayliss *et al.*, 2003) and 1mq4 (Nowakowski *et al.*, 2002)]. A superposition of the six structures (including Plk-1 molecule *B* in the wild-type structure) is shown in Fig. 4(*b*). The individual structures, together with their r.m.s.d. and alignment quality (*Q* score) values with respect to Plk-1 molecule *A*, are presented in Fig. 4(*c*). These results confirm that the best search models for solving the three-dimensional structure of Plk-1 by MR are PDB entry 1ol5 (which had been used in this study as a starting point for homology models) and PDB entry 1mq4, which was used by Kothe and coworkers for MR (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007).

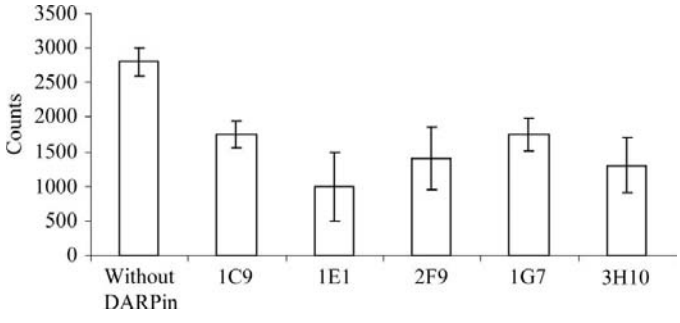

Figure 3
Characterization of the Plk-1–DARPin interaction. The Plk-1 activity assay shows that all DARPins partially inhibit the kinase activity. However, they do not fully block access to the ATP site as the degree of inhibition could be further increased by adding a small-molecule inhibitor (data not shown).

3.5. Structure of DARPin 3H10

The structure of DARPin 3H10 shows the typical ankyrin-repeat fold with an overall curved arrangement up to the fourth repeat, as shown in Fig. 5(*a*). Based on its sequence, 3H10 was expected to feature five ankyrin repeats (Fig. 5*b*). However, in the crystal structure the fifth ankyrin repeat was fully disordered. This explained why the truncated DARPin search models with four repeats led to the most convincing MR solutions (see §2). Also, the absence of the fifth ankyrin repeat allowed the formation of the close tail-to-tail contact between the two DARPins in the crystal structure (Fig. 6*a*).

3.6. Molecular recognition between Plk-1 and DARPin

The asymmetric unit in the crystal structure contains two Plk-1–DARPin heterodimers which are related by a twofold noncrystallographic symmetry axis and form a heterotetramer (Fig. 6*a*). The interaction surfaces in the crystal structure were analyzed with the *Protein Interfaces, Surfaces and Assemblies* service (*PISA*; Krissinel & Henrick, 2007). The results are summarized in Supplementary Table 1. As expected from the arrangement of the Plk-1 and DARPin molecules in the heterotetramer, the most important interaction surfaces are those within each Plk-1–DARPin heterodimer: *AD* (Plk-1 molecule *A* with 3H10 molecule *D*) and *BC* (Plk-1 molecule *B* with 3H10 molecule *C*). With interface areas of 850 and 870 Å^2, respectively, these interactions feature the largest interface areas and the most negative estimated values of total free-energy gain upon interface formation ($\Delta^t G$). They were also assigned a complexation significance score (CSS) of 1, indicating high significance for assembly formation. In the heterotetramer there are two additional interaction surfaces: *AB* (between Plk-1 molecules *A* and *B*), with the third largest area but the lowest value of Δ^tG, and *CD* (between the two DARPin molecules *C* and *D*), with the smallest area of the four but with the third largest value of $\Delta^t G$. Among all the interfaces observed in the crystal structure, these four lead to the highest estimated total free-energy gain (−159.5 kJ mol^{-1}) upon tetramer formation.

3H10 binds to a surface epitope which is located immediately in front of and below the ATP-binding pocket (Fig. 6*b*). The alignment in Fig. 5(*b*) shows that the residues employed by the DARPin (highlighted in blue) mostly consist of those

positions which were randomized for library construction as described in Binz *et al.* (2004).

The epitope on Plk-1#4 recognized by DARPin 3H10 consists of extended hydrophobic and polar interactions (Supplementary Table 2) and covers 17 residues of Plk-1 (4 Å distance cutoff). These include six arginine, one lysine and two glutamate residues. Such residues with large side chains and high conformational freedom are considered counterproductive for crystal contact formation owing to the loss of entropy when they are incorporated into crystal contacts (Derewenda, 2004). Covering or masking these surface residues by the DARPins will therefore have positively influenced crystallization.

3.7. Conformation of wild-type Plk-1

Protein kinases switch between an active conformation, in which the binding-site residues are in the correct orientation to catalyze phosphoryl transfer, and inactive conformations, in which one or several catalytically relevant residues are disorientated or in which the binding sites are blocked owing to conformational changes of loop regions (Huse & Kuriyan, 2002). Depending on the activation state, helix αC in the N-terminal lobe generally rotates in and out slightly. This movement is connected with the formation or breaking of the salt bridge between two conserved residues: Lys82 in the back wall of the ATP site and Glu101 in helix αC. Only when this salt bridge is formed is the lysine in the correct orientation to

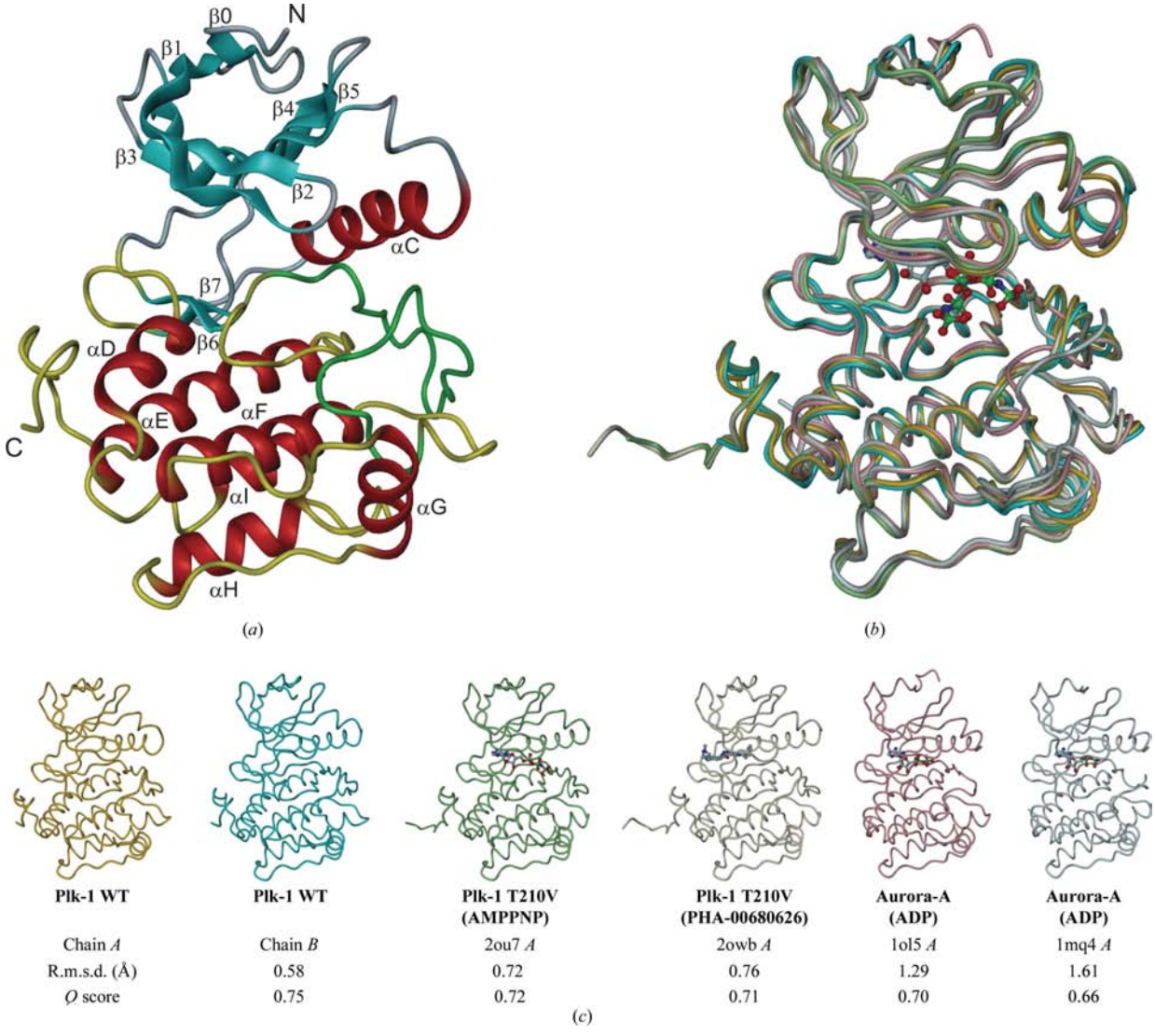

Figure 4
Three-dimensional structure of wild-type Plk-1. (*a*) Ribbon diagram of Plk-1 molecule *A* in the asymmetric unit of the Plk-1#4–DARPin 3H10 crystal structure. The C^{α} chains of the N-terminal and C-terminal lobes are coloured grey and yellow, respectively. The secondary-structure elements (SSEs) are highlighted (cyan for β-strands, red for α-helices). The activation segment, which is located in the C-terminal domain, is shown in green. The SSE numbering scheme follows that of Kothe, Kohls, Low, Coli, Cheng *et al.* (2007). (*b*) Tube diagram of the C^{α} chains of the superposed Plk-1 molecules *A* and *B* with their four closest structural homologues. The colour code is explained in (*c*). (*c*) Side-by-side display of tube diagrams of the C^{α} chains of all the molecules represented in (*b*) with the colour code used in (*b*). The values listed for r.m.s.d. and *Q* score were taken from the cross-structure statistics of a multiple three-dimensional alignment of the six structures. This figure was prepared with *DINO* (Philippsen, 2002).

coordinate the phosphates of ATP in a catalytically productive way. In Plk-1 molecule *A* the salt bridge is present (Lys–Glu distance 3.1 Å), which indicates an active conformation (in molecule *B* this distance is 3.8 Å, indicating a partially inactive conformation). Further indications of an active state are the conformations of the DFG motif (residues 194–196), the nucleotide-binding loop (G*X*G*XX*G motif, here ^{60}GKG-GFA^{65}) and the activation segment (^{194}DFG to APE^{221}). The DFG motif features the 'DFG-in' conformation which allows ATP binding in a catalytically productive way (Pargellis *et al.*, 2002). The nucleotide-binding loop between strands β1 and β2 (which forms the roof of the ATP site in the orientation shown in Fig. 4*a*) does not fold down into the empty ATP-binding site, as seen for example in the inactive conformation of Tie2 (PDB code 1vfr; Shewchuk *et al.*, 2000). The activation segment is in an extended conformation. In the inactive state, this segment often folds back onto the protein and blocks the substrate-binding groove, the ATP-binding site or both (Johnson *et al.*, 1996). This blockage is released when the kinase is phosphorylated at one (or more) serine, threonine or tyrosine residues in the activation segment, for example Thr210 in Plk-1 (Qian, Erikson, Li *et al.*, 1998; Qian, Erikson & Maller, 1998; Qian *et al.*, 1999).

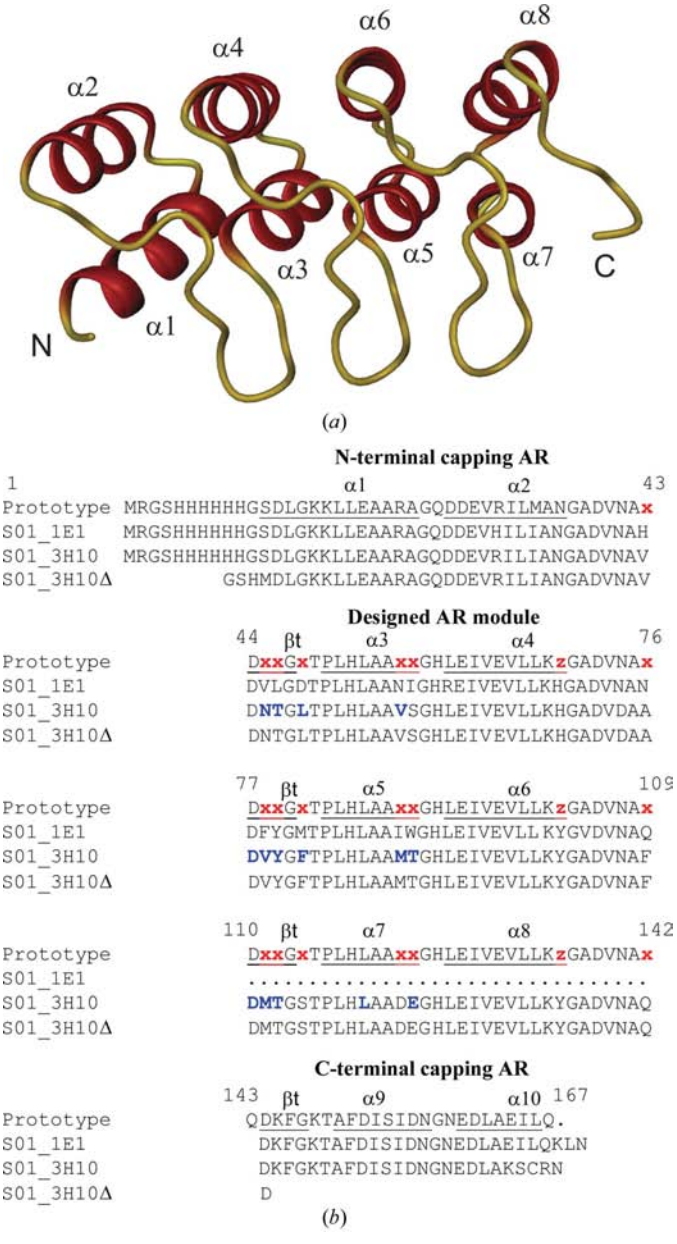

Figure 5
Three-dimensional structure of DARPin 3H10 and DARPin sequence alignment. (*a*) Ribbon diagram of DARPin 3H10 molecule *C*. The α-helices are coloured red. This figure was prepared with *DINO* (Philippsen, 2002). (*b*) Sequence alignment of a prototypical DARPin (Binz *et al.*, 2004), the four-ankyrin-repeat DARPin 1E1, the five-ankyrin-repeat DARPin 3H10 and the truncated version 3H10Δ. Positions listed as 'x' in red represent randomized residues for target binding (any amino acid except Cys, Gly or Pro); 'z' indicates randomized framework residues (His, Asn or Tyr; Binz *et al.*, 2004). Residues in blue and bold are those involved in the binding of Plk-1#4 in the crystal structure (4 Å distance).

The indications of an active kinase conformation are apparently contradicted by the conformation of Arg136 at the entrance to the ATP site near the hinge region. In the absence of a ligand, its side chain stretches up to the nucleotide-binding loop (G*X*G*XX*G motif) and hydrogen bonds to the backbone carbonyl group of Plk-1 Leu59 (Fig. 6*b*; 2.8 and 2.7 Å in molecules *A* and *B*, respectively). By doing so, it partially blocks the front entrance to the adenine subpocket. However, Arg136 is located at the circumference of the Plk-1–DARPin interface and a side-chain conformation such as that observed in the ligand-bound Plk-1 mutant structures is prevented by the presence of Val78 of DARPin 3H10 (Figs. 6*b* and 6*c*). The carbonyl O atom of this residue forms a bifurcated hydrogen bond to Arg136 (2.9 Å to N^{ε} and $N^{\eta 2}$ in molecule *A*; 2.9 Å to N^{ε} and 3.0 Å to $N^{\eta 2}$ in molecule *B*). As the DARPins were selected in the absence of substrate, they might well stabilize this conformation. On the one hand, this may be an explanation for the inhibitory effects of DARPin 3H10. On the other hand, it may also have helped to reveal a conformation that is possible in solution and which the DARPin was able to capture. Nevertheless, as mentioned above, ATP and inhibitor binding is not prevented by the complex formation.

4. Discussion

While the kinase domain of Plk-1 has been recognized as a potential drug target for some time, the first crystal structures of this important molecule were reported only recently and not for the wild-type kinase but for Plk-1 harbouring a mutation which was required for crystallization (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007). This delay may be a consequence of the fact that the kinase domain of Plk-1 has proven to be notoriously difficult to express in soluble form and in yields sufficient for crystal structure determination (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007). Only after employing the multi-parallel cloning and expression approach first described in Malawski *et al.* (2006) were we able to identify a set of constructs which could be expressed as soluble proteins and in large amounts in a variety of eukaryotic expression systems. These constructs featured domain boundaries that were

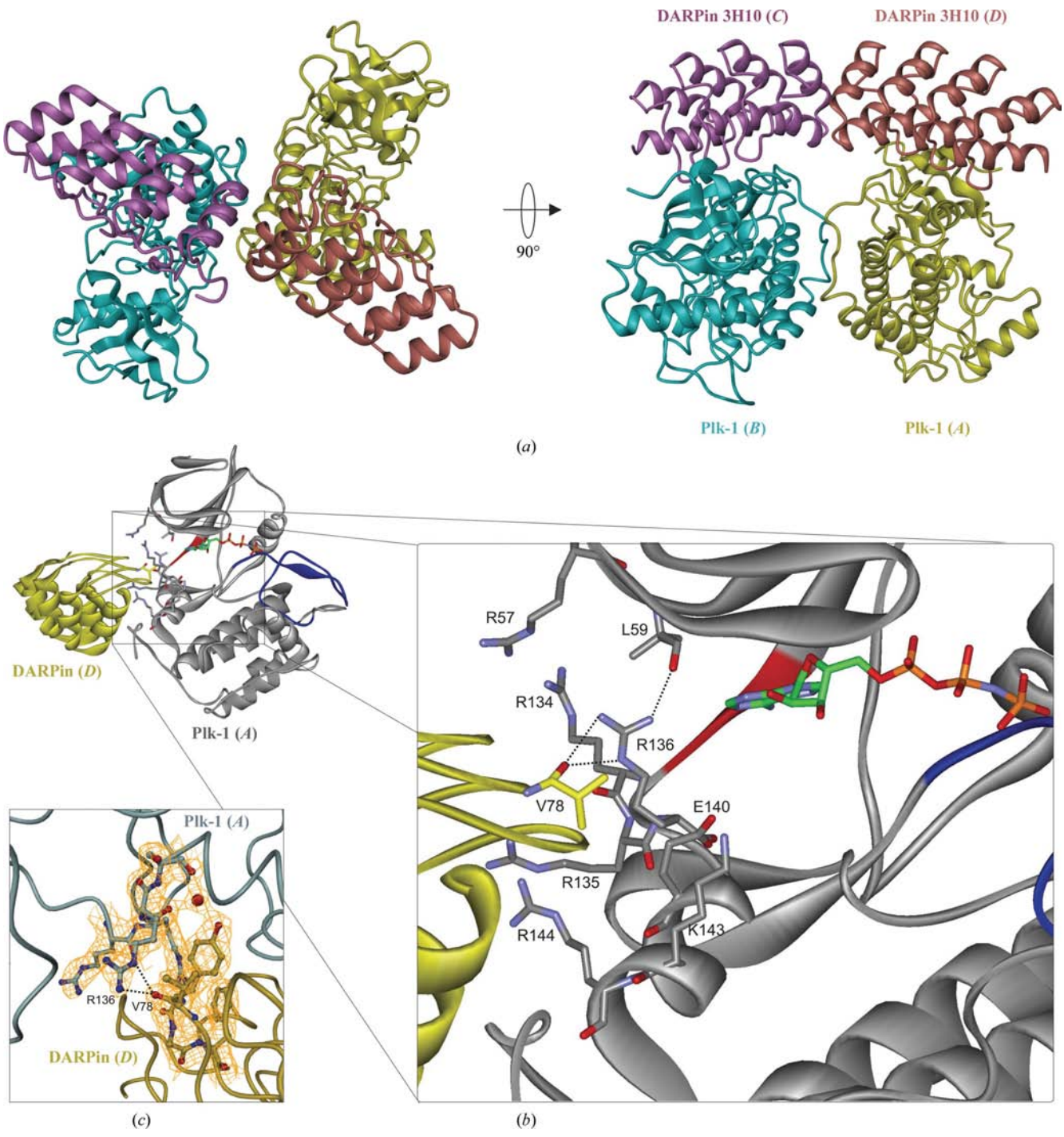

Figure 6
The Plk-1#4–DARPin 3H10 interaction. (*a*) Ribbon diagrams of the $\alpha_2\beta_2$ heterotetramer in the asymmetric unit of the crystal structure of the complex of Plk-1#4 and DARPin 3H10. In the left panel, the view is down the noncrystallographic twofold-rotation axis; in the right panel, the noncrystallographic twofold-rotation axis lies vertically on the page. (*b*) Ribbon representation of molecules *A* (Plk-1#4) and *D* (3H10), with Plk-1 shown in grey and 3H10 in yellow. For orientation, AMPPNP from PDB entry 2ou7 (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007) is superimposed and shown in atom colours; the hinge region (residues 131–133) is shown in red and the activation segment (residues 194–221) is shown in blue. Selected residues contributing to the interface are shown in stick representation. The DARPin uses its randomized residue positions, predominantly those at the tips of the short β-hairpins (see alignment in Fig. 5). The epitope on Plk-1 covers a large set of arginines, two glutamates and one lysine (see text and Supplementary Table 2). Arg136 in front of the ATP site forms hydrogen bonds to Leu59 (in the nucleotide-binding loop above the ATP site) and to DARPin residue Val78 (dotted lines), thereby partially blocking access to the ATP site. (*c*) Detailed view of the Arg136 region in Plk-1. Val78 of the DARPin pushes the side chain of Plk-1 Arg136 towards the ligand-binding pocket; the ^{134}RRRS^{137} region in Plk-1 chain *A*, the ^{78}VYGF^{81} region of DARPin chain *D* and a well ordered bridging water molecule are shown in ball-and-stick representation and the C^{α} chains for rest of the molecules are shown as tubes. The C atoms and the C^{α} chain are coloured light cyan for Plk-1 and gold for DARPin. The final $2|F_o| - |F_c|$ electron-density map is drawn at the 1.2 map r.m.s. contour level around the ball-and-stick atoms to illustrate the quality of the electron-density map. This figure was prepared with *DINO* (Philippsen, 2002) and *INSIGHT* (Accelrys Inc., San Diego, USA).

significantly extended from the expected core kinase domain. They could be purified and concentrated to levels generally sufficient for crystallization. Initial indications that the length of the core kinase domain had been underestimated were obtained from the limited proteolysis experiments, underlining the value of such experiments for recombinant expression studies (Gao *et al.*, 2005). Our results show the power of multi-parallel expression approaches in a broad variety of expression systems, particularly when aiming at the production of recombinant protein for X-ray crystallography (Malawski *et al.*, 2006).

The two different yeast systems *S. cerevisiae* and *P. pastoris* revealed interesting differences in the recombinant expression levels of Plk-1 kinase-domain constructs. For expression of intracellular soluble proteins *S. cerevisiae* appears to be a more suitable host than *P. pastoris*. In this study, insect cells were finally used for expression because the yield was superior to *S. cerevisiae* and because the structure could be solved by molecular replacement. Had molecular replacement not worked then *S. cerevisiae* would have been selected, as this system is amenable to selenomethionine labelling for anomalous phasing techniques (Turnbull *et al.*, 2005).

Seven Plk-1 constructs expressed in insect cells could be purified and concentrated to levels that are generally sufficient for crystallization, but all failed to crystallize. In order to modify the surface of Plk-1, we generated a set of designed ankyrin-repeat proteins (DARPins) which bind to Plk-1 with high affinity and specificity. DARPins have been generated for a number of protein targets (Stumpp & Amstutz, 2007) and have recently been successfully used to facilitate or enable the crystallization of their target proteins (Binz *et al.*, 2004; Amstutz *et al.*, 2005; Schweizer *et al.*, 2007; Zahnd *et al.*, 2007; Sennhauser *et al.*, 2006).

Plk-1 construct #4 was chosen as this construct had behaved best in multiple purification runs. We generated five DARPins specific for this Plk-1 construct by ribosome-display selection. Both binding to Plk-1#4 and the simultaneous binding of an ATP-site inhibitor could be demonstrated for all five. This is an important prerequisite to ensure that derived crystals will allow incorporation of inhibitors to enable structure-guided drug design. Crystallization screens with mixtures of Plk-1#4 and the various DARPins produced crystals with DARPins 3H10, 2F9 and 1G7. The initial crystals with 3H10 were largest and showed the best morphology. They were optimized and used for data-collection and structure determination.

Based on the structure of Plk-1#4–3H10, we rationalized the behaviour of the different-length variants during expression. The two shortest constructs (#9 and #10) and the even shorter initial kinase core construct (Fig. 1) could not be expressed either in insect cells or in *S. cerevisiae*, while all longer constructs were expressed well in these two systems. The structure reveals that Plk-1 features an additional N-terminal β-strand (β0) which had not been expected during construct design. Strand β0 comprises residues 42–46 and is preceded by three residues (39–41) which are ordered in the crystal structure. The shortest constructs, #9 and #10, start at residues 43 and 48, which translates into a shortening or the complete removal of this β-strand. The dramatically reduced expression rates indicate that interfering with this β-strand significantly affects the folding properties and/or the stability of the resulting Plk-1 variants. Accordingly, the proteases used in limited proteolysis studies did not touch strand β0 but cut well before at residues 13 and 33. The C-terminus in the structure of Plk-1#4 ends in a loop region, with residue 323 as the final ordered residue in the density. This stably folded section is retained in the eight longer and well expressed constructs, but is slightly shortened in constructs #9 (ending at 321) and #10 (ending at 316). In the insoluble initial minimal kinase-domain construct (50–312) both the N-terminal strand β0 and the C-terminal loop region are deleted. As already seen during the construct design of protein kinase MK2 (Malawski *et al.*, 2006), parallel generation and testing of a large set of length variants turned out to be crucial in identifying a suitable construct for structural biology.

The structure shows extensive interactions between Plk-1#4 and DARPin 3H10 which involve a number of polar inter-

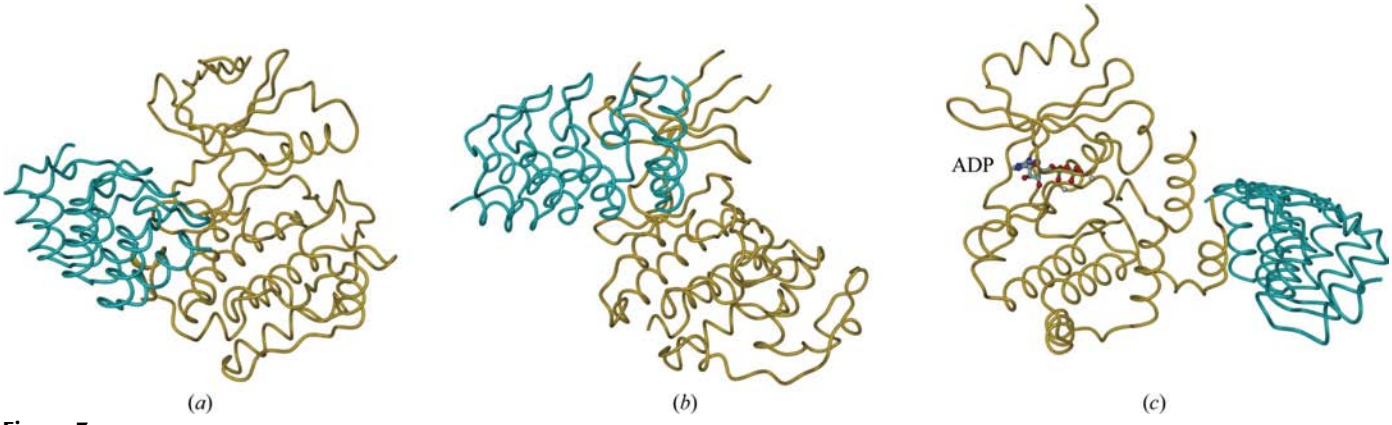

Figure 7
Comparison of Plk-1–3H10 with further kinase–ankyrin-repeat (AR) protein complexes. (*a*)–(*c*) Tube diagrams showing the overall folds of kinases in complex with AR proteins: (*a*) Plk-1–DARPin 3H10; (*b*) Cdk6–P19INK4D inhibitor complex (Russo *et al.*, 1998; chains *A* and *B* from PDB entry 1bi8); (*c*) APH–AR inhibitor complex (Kohl *et al.*, 2005; chains *A* and *B* from PDB entry 2bkk). In all panels the kinase is coloured gold and the AR protein is coloured cyan. The ADP in APH is shown in ball-and-stick representation (C, N, O and P atoms are coloured grey, blue, red and green, respectively) and is a reference for the ATP site. This figure was prepared with *DINO* (Philippsen, 2002).

actions including several salt bridges (Supplementary Table 2). However, characterization by ITC revealed that the overall interaction is endothermic and is apparently driven by entropy alone. In the structure, the C-terminal capping repeat of DARPin 3H10 was not defined. This observation indicated that either the interaction with Plk-1#4 or the close tail-to-tail interaction with a neighbouring DARPin molecule in the crystal lattice triggered a partial unfolding of the capping repeat. Since unfolding reactions are endothermic by nature, we speculated that the potentially exothermic binding reaction between 3H10 and Plk-1#4 was neutralized by the partial unfolding of the C-terminal capping repeat of the DARPin. In order to test this hypothesis, DARPin variant 3H10Δ in which the C-terminal residues 142–167 had been deleted (Fig. 5*b*) was generated and titrated under identical conditions as full-length 3H10. As expected, the binding was now characterized by exothermic heat changes (Fig. 2*c*). Compared with full-length 3H10, the interaction of Plk-1 with 3H10Δ showed a dramatic increase in binding enthalpy of more than 33 kJ mol^{-1}.

Judging from the structure, it is conceivable that either the close tail-to-tail interaction of the two DARPin molecules in the asymmetric unit required the unfolding of the C-terminal capping repeat or that the interaction with Plk-1 triggered the unfolding as the fifth repeat would have clashed with Plk-1. The ITC data indicated that the absence of the C-terminal capping repeat dramatically improved binding to Plk-1 already in solution, excluding a crystal contact as the driving force for unfolding. Interestingly, DARPin 3H10 emerged from the selection procedure with a frame-shift mutation in the C-terminal capping repeat which mutated the five C-terminal residues ^{163}EILQA^{167} to ^{163}KSCRN^{167} (Fig. 5*b*). This mutation may have destabilized the complete C-terminal repeat and may thus have been selected as favourable for Plk-1 recognition. The removal of the C-terminal capping repeat in 3H10Δ then completed the optimization process started by the serendipitous frame-shift mutation.

Packing analysis of the Plk-1–DARPin crystal form suggested that the DARPin–DARPin tail-to-tail interaction represented the main driving force for the formation of the heterotetramer and thus indirectly for the formation of stable crystal nuclei, even though the N-terminal capping repeat was disordered in the structure. This reinforces the potential role of DARPins as protein cocrystallization tools in difficult cases in which the target protein alone does not crystallize or produces poor-quality crystals. On the other hand, the difficulties in obtaining crystals of the Plk-1#4–DARPin 3H10 complex with good diffracting properties (see §2.9) may be related to the disorder of the C-terminal ankyrin repeat. Trials using DARPin 3H10Δ are currently under way.

The epitope recognized by 3H10 contains a bipartite nuclear localization sequence (^{134}RRRSLLELHKRRK^{146}) characterized by two sets of basic residues (Taniguchi *et al.*, 2002). 3H10 forms hydrogen bonds with Arg135, Arg136 and Arg144 and is in van der Waals contact with Arg134 and Lys143 of this motif. In addition, there is a bidentate salt bridge to Arg313 and van der Waals contacts to Arg57, Glu140 and Glu112. Clusters of charged and mobile surface residues such as Lys, Glu and Arg are considered to be counter-productive for crystallization as packing into a crystal contact would result in an unfavourable loss of entropy (Derewenda, 2004). The DARPin probably enabled crystallization of Plk-1 by masking this surface patch. It may thus be considered a substitute for Lys/Glu-to-Ala surface mutations which in other cases facilitated crystallization of difficult crystallization targets (for a kinase, see Munshi *et al.*, 2003). In addition to masking this unproductive surface patch, the DARPin contributed new and productive crystal contacts.

During the crystal structure determination of mutant Plk-1, Kothe and coworkers also identified a soluble and well behaved Plk-1 construct which subsequently did not crystallize (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007). The construct employed (residues 13–345) was identical to our construct #1 (which in our hands yielded slightly lower amounts than constructs #3–#8). To overcome the crystallization bottleneck, Kothe and coworkers also turned to modifying the surface and the successful mutant (Thr210Val) subsequently indeed turned out to contribute a crystal contact.

It is noteworthy that a number of natural ankyrin-repeat (AR) proteins have evolved to recognize and inhibit kinases. Examples are the cell-division kinase inhibitors p16INK4a and p19INK4d (Russo *et al.*, 1998; Brotherton *et al.*, 1998). DARPins have also already been successfully generated against other kinases, such as the bacterial kinase APH (Amstutz *et al.*, 2005; Kohl *et al.*, 2005) and a mitogen-activated protein kinase (Amstutz *et al.*, 2006). Fig. 7

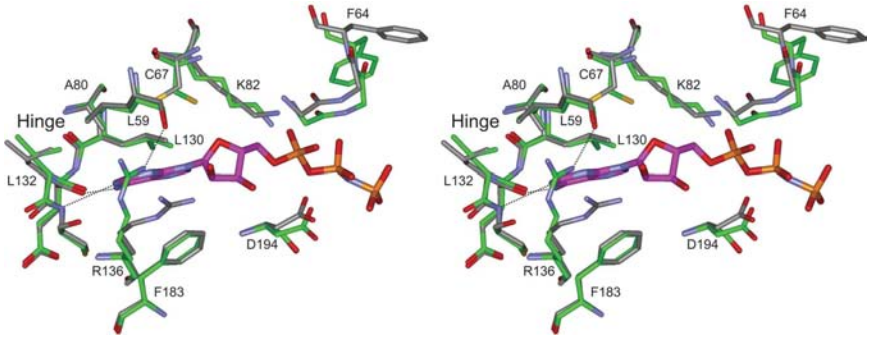

Figure 8
Superimposition of the active sites of apo wild-type Plk-1 (chain *A*) and Plk-1 T210V in complex with AMPPNP (PDB entry 2ou7). A stereo representation of the selected active site is shown with C atoms shown in grey for Plk-1–AMPPNP, in green for apo Plk-1 and in magenta for AMPPNP. The nucleotide forms two hydrogen bonds to the hinge region, which are depicted as dotted lines. Both structures show a very similar active-site conformation. Differences are the slightly more open nucleotide-binding loop with Phe64 at its tip, a different side-chain conformation for Cys67 at the roof of the ATP site and the conformation of Arg136 at the lower entrance of the ATP pocket. In the apo form, the side chain of Arg136 forms a hydrogen bond to Leu59 at the upper entrance (dotted line). Leu130 is the gatekeeper residue, Lys82 is the conserved ATP-site lysine residue and Asp194 of the DFG motif marks the start of the activation segment.

illustrates that the different AR proteins recognize different surface epitopes of the respective kinases. While the natural AR kinase inhibitors act by directly binding to the N-terminal lobe, the APH inhibitor acts as an allosteric inhibitor which binds away from the active site in the C-terminal lobe. In the study presented here, we did not intend to generate DARPins as Plk-1 inhibitors, but as surface modifiers to enable crystal growth. Nevertheless, all selected Plk-1-specific DARPins turned out to also partially inhibit Plk-1 kinase activity. The crystal structure revealed that the partial inhibition seen with 3H10 is probably related to the stabilization of Arg136 at the entrance of the ATP site. However, as ATP can still bind to the complex, Arg136 must be able to adopt a different side-chain conformation and to allow access to this site at least temporarily.

Protein kinase inhibitors typically bind to the ATP-binding site and occasionally into subpockets adjacent to this site (Noble *et al.*, 2004). The ATP site in apo wild-type Plk-1#4 superimposes very well with the ATP site of Plk-1 T210V in complex with AMPPNP (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007). Fig. 8 shows that the only significant differences are a slightly more open conformation of the nucleotide-binding loop, with Phe64 at the tip of this loop adopting different conformations, and different side-chain conformations of residue Cys67 above the ATP molecule. Together with the observed ability to bind ATP and inhibitors in the presence of 3H10, this similarity confirms the validity of our approach of using DARPins as crystallization facilitators with the ultimate goal of supporting structure-guided drug design.

Protein kinases play important regulatory roles in signal transduction pathways. They are therefore themselves tightly regulated in their activity and many different mechanisms exist which ensure that a particular kinase is only active when required (reviewed in Huse & Kuriyan, 2002). Activation involves phosphorylation of the kinase in its activation segment either by autophosphorylation or by an upstream kinase. For full activation, Plk-1 is phosphorylated at Thr210 (Qian, Erikson, Li *et al.*, 1998; Qian, Erikson & Maller, 1998; Qian *et al.*, 1999). Mass spectrometry and the electron-density maps confirmed that Plk-1#4 as used in this study was not phosphorylated. Nevertheless, the employed construct was shown to be active in an enzymatic assay and the conformation found in the crystal structure also suggests that Plk-1 crystallized in its active state. The Lys82–Glu101 salt bridge to helix αC is formed, the normal 'DFG-in' conformation is observed and the activation segment is in an extended conformation. Therefore, it appears that unphosphorylated Plk-1 is already in a conformation similar to the active conformation. Consistent with this assumption, the specific activity of the Plk-1 T210D mutant (mimicking phosphorylation) increased only fourfold relative to wild-type Plk-1 (Kothe, Kohls, Low, Coli, Cheng *et al.*, 2007).

The two Plk-1 mutant structures reported by Kothe, Kohls, Low, Coli, Cheng *et al.* (2007) also feature the active conformation, despite the fact that the deactivating mutant Thr210Val was employed for crystallization. Thus, all three available Plk-1 crystal structures indicate that the kinase domain of Plk-1 may well be in its active conformation by default. *In vivo*, it is possibly kept in an inactive state only *via* intramolecular interaction with the PBD, which was shown to reduce the activity of the kinase domain (Elia, Rellos *et al.*, 2003). Furthermore, productive activity is also thought to be achieved only upon recruitment of Plk-1 to its cellular targets and substrates, which is again mediated by the PBD (Lowery *et al.*, 2007).

In conclusion, we have determined the first structure of the wild-type kinase domain of Plk-1. The structure extends the experimental basis for future docking and cocrystallization studies with this pharmacologically important target. We showed that DARPin binding still allows ATP access to the nucleotide-binding site, rendering this crystal form useful for the generation of inhibitor co-complexes *via* soaking or cocrystallization. Crystallization was only possible after the identification of a suitable construct employing a parallel cloning and expression approach and after generation of a DARPin specific to Plk-1. DARPin 3H10 was most likely to have facilitated crystallization by a combination of contributing new crystal contacts and masking an Arg/Lys/Glu surface cluster which otherwise would have been entropically unfavourable for crystal packing.

Note added in proof: While this paper was under review, a further Plk-1 crystal structure was published (Kothe, Kohls, Low, Coli, Rennie *et al.*, 2007).

The authors would like to thank Martina Huber and Martina Schäfer for helpful discussions, Patrik Forrer for support and discussions, Mario Mann, Alexandra Einfeldt, Carola Moldenhauer, Anne Sparmann, Norbert Otto, Anja Wegg, Marina Isernhagen, Petra Helfrich and particularly Antje Haensel and Sandra Lange for excellent technical support, Volker Badock for mass spectrometry, Ricardo Coelho for initial help with the Plk-1–DARPin cocrystallization, and Stephanie Monaco and Elspeth Gordon (ESRF, Grenoble) for collecting the ID23-2 and ID29 diffraction data.

References

Amstutz, P., Binz, H. K., Parizek, P., Stumpp, M. T., Kohl, A., Grütter, M. G., Forrer, P. & Plückthun, A. (2005). *J. Biol. Chem.* **280**, 24715–24722.

Amstutz, P., Koch, H., Binz, H. K., Deuber, S. A. & Plückthun, A. (2006). *Protein Eng. Des. Sel.* **19**, 219–229.

Andrews, C. W., Cheng, M., Davis-Ward, R. G., Drewry, D. H., Emmitte, K. A., Hubbard, R. D., Kunts, K. W., Linn, J. A., Mook, R. A., Smith, G. K. & Veal, J. M. (2004). Patent WO 2004/014899.

Barr, F. A., Sillje, H. H. & Nigg, E. A. (2004). *Nature Rev. Mol. Cell Biol.* **5**, 429–440.

Bayliss, R., Sardon, T., Vernos, I. & Conti, E. (2003). *Mol. Cell*, **12**, 851–862.

Bearss, D. J., Vankayalapati, H. & Grand, C. L. (2006). Patent PCT/US2006/016423.

Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grütter, M. G. & Plückthun, A. (2004). *Nature Biotechnol.* **22**, 575–582.

Brotherton, D. H., Dhanaraj, V., Wick, S., Brizuela, L., Domaille, P. J., Volyanik, E., Xu, X., Parisini, E., Smith, B. O., Archer, S. J.,

Serrano, M., Brenner, S. L., Blundell, T. L. & Laue, E. D. (1998). *Nature (London)*, **395**, 244–250.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
Cheng, K. Y., Lowe, E. D., Sinclair, J., Nigg, E. A. & Johnson, L. N. (2003). *EMBO J.* **22**, 5757–5768.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davis-Ward, R., Mook, R. A., Neeb, M. J. & Salovich, J. M. (2004). Patent PCT/US2004/004197.
Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
Eckerdt, F., Yuan, J. & Strebhardt, K. (2005). *Oncogene*, **24**, 267–276.
Elia, A. E., Cantley, L. C. & Yaffe, M. B. (2003). *Science*, **299**, 1228–1231.
Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J. & Yaffe, M. B. (2003). *Cell*, **115**, 83–95.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fancelli, D. *et al.* (2005). *J. Med. Chem.* **48**, 3080–3084.
Gao, X., Bain, K., Bonanno, J. B., Buchanan, M., Henderson, D., Lorimer, D., Marsh, C., Reynes, J. A., Sauder, J. M., Schwinn, K., Thai, C. & Burley, S. K. (2005). *J. Struct. Funct. Genomics*, **6**, 129–134.
Golsteyn, R. M., Mundt, K. E., Fry, A. M. & Nigg, E. A. (1995). *J. Cell Biol.* **129**, 1617–1628.
Hamanaka, R., Smith, M. R., O'Connor, P. M., Maloid, S., Mihalic, K., Spivak, J. L., Longo, D. L. & Ferris, D. K. (1995). *J. Biol. Chem.* **270**, 21086–21091.
Huse, M. & Kuriyan, J. (2002). *Cell*, **109**, 275–282.
Johnson, L. N., Noble, M. E. & Owen, D. J. (1996). *Cell*, **85**, 149–158.
Kohl, A., Amstutz, P., Parizek, P., Binz, H. K., Briand, C., Capitani, G., Forrer, P., Plückthun, A. & Grütter, M. G. (2005). *Structure*, **13**, 1131–1141.
Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Plückthun, A. & Grütter, M. G. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 1700–1705.
Komander, D., Kular, G. S., Schuttelkopf, A. W., Deak, M., Prakash, K. R., Bain, J., Elliott, M., Garrido-Franco, M., Kozikowski, A. P., Alessi, D. R. & van Aalten, D. M. (2004). *Structure*, **12**, 215–226.
Kothe, M., Kohls, D., Low, S., Coli, R., Cheng, A. C., Jacques, S. L., Johnson, T. L., Lewis, C., Loh, C., Nonomiya, J., Sheils, A. L., Verdries, K. A., Wynn, T. A., Kuhn, C. & Ding, Y. H. (2007). *Biochemistry*, **46**, 5960–5971.
Kothe, M., Kohls, D., Low, S., Coli, R., Rennie, G. R., Feru, F., Kuhn, C. & Ding, Y. H. (2007). *Chem. Biol. Drug Des.* **70**, 540–546.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lee, K. S., Yuan, Y. L., Kuriyama, R. & Erikson, R. L. (1995). *Mol. Cell. Biol.* **15**, 7143–7151.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Leung, G. C., Hudson, J. W., Kozarova, A., Davidson, A., Dennis, J. W. & Sicheri, F. (2002). *Nature Struct. Biol.* **9**, 719–724.
Lowery, D. M., Clauser, K. R., Hjerrild, M., Lim, D., Alexander, J., Kishi, K., Ong, S. E., Gammeltoft, S., Carr, S. A. & Yaffe, M. B. (2007). *EMBO J.* **26**, 2262–2273.
McInnes, C., Mezna, M. & Fischer, P. M. (2005). *Curr. Top. Med. Chem.* **5**, 181–197.
Malawski, G. A., Hillig, R. C., Monteclaro, F., Eberspaecher, U., Schmitz, A. A., Crusius, K., Huber, M., Egner, U., Donner, P. & Müller-Tiemann, B. (2006). *Protein Sci.* **15**, 2718–2728.
Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. & Sali, A. (2000). *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325.
Munshi, S., Hall, D. L., Kornienko, M., Darke, P. L. & Kuo, L. C. (2003). *Acta Cryst.* D**59**, 1725–1730.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Noble, M. E., Endicott, J. A. & Johnson, L. N. (2004). *Science*, **303**, 1800–1805.
Nowakowski, J., Cronin, C. N., McRee, D. E., Knuth, M. W., Nelson, C. G., Pavletich, N. P., Rogers, J., Sang, B. C., Scheibe, D. N., Swanson, R. V. & Thompson, D. A. (2002). *Structure*, **10**, 1659–1667.
Pargellis, C., Tong, L., Churchill, L., Cirillo, P. F., Gilmore, T., Graham, A. G., Grob, P. M., Hickey, E. R., Moss, N., Pav, S. & Regan, J. (2002). *Nature Struct. Biol.* **9**, 268–272.
Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Philippsen, A. (2002). *DINO*. http://www.dino3d.org.
Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* D**59**, 1131–1137.
Qian, Y. W., Erikson, E., Li, C. & Maller, J. L. (1998). *Mol. Cell. Biol.* **18**, 4262–4271.
Qian, Y. W., Erikson, E. & Maller, J. L. (1998). *Science*, **282**, 1701–1704.
Qian, Y. W., Erikson, E. & Maller, J. L. (1999). *Mol. Cell. Biol.* **19**, 8625–8632.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.
Roussel, A., Fontecilla-Camps, J. C. & Cambillau, C. (1990). *Acta Cryst.* A**46**, C66–C67.
Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D. & Pavletich, N. P. (1998). *Nature (London)*, **395**, 237–243.
Santamaria, A., Neef, R., Eberspächer, U., Eis, K., Husemann, M., Mumberg, D., Prechtl, S., Schulze, V., Siemeister, G., Wortmann, L., Barr, F. A. & Nigg, E. A. (2007). *Mol. Biol. Cell*, **18**, 4024–4036.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.
Schwede, W., Schulze, V., Eis, K., Buchmann, B., Briem, H., Siemeister, G., Bömer, U. & Parczyk, K. (2003). Patent PCT/EP2003/004450.
Schweizer, A., Roschitzki-Voser, H., Amstutz, P., Briand, C., Gulotti-Georgieva, M., Prenosil, E., Binz, H. K., Capitani, G., Baici, A., Plückthun, A. & Grütter, M. G. (2007). *Structure*, **15**, 625–636.
Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O. & Grütter, M. G. (2006). *PLoS Biol.* **5**, e7.
Shewchuk, L. M., Hassell, A. M., Ellis, B., Holmes, W. D., Davis, R., Horne, E. L., Kadwell, S. H., McKee, D. D. & Moore, J. T. (2000). *Structure*, **8**, 1105–1113.
Steegmaier, M., Hoffmann, M., Baum, A., Lenart, P., Petronczki, M., Krssak, M., Gurtler, U., Garin-Chesa, P., Lieb, S., Quant, J., Grauert, M., Adolf, G. R., Kraut, N., Peters, J. M. & Rettig, W. J. (2007). *Curr. Biol.* **17**, 316–322.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Strebhardt, K. & Ullrich, A. (2006). *Nature Rev. Cancer*, **6**, 321–330.
Stumpp, M. T. & Amstutz, P. (2007). *Curr. Opin. Drug Discov. Devel.* **10**, 153–159.
Takai, N., Hamanaka, R., Yoshimatsu, J. & Miyakawa, I. (2005). *Oncogene*, **24**, 287–291.
Taniguchi, E., Toyoshima-Morimoto, F. & Nishida, E. (2002). *J. Biol. Chem.* **277**, 48884–48888.
Turnbull, A. P., Kummel, D., Prinz, B., Holz, C., Schultchen, J., Lang, C., Niesen, F. H., Hofmann, K. P., Delbruck, H., Behlke, J., Muller, E. C., Jarosch, E., Sommer, T. & Heinemann, U. (2005). *EMBO J.* **24**, 875–884.
Wan, P. T., Garnett, M. J., Roe, S. M., Lee, S., Niculescu-Duvaz, D., Good, V. M., Jones, C. M., Marshall, C. J., Springer, C. J., Barford, D. & Marais, R. (2004). *Cell*, **116**, 855–867.
Weerdt, B. C. van de & Medema, R. H. (2006). *Cell Cycle*, **5**, 853–864.
Zahnd, C., Wyler, E., Schwenk, J. M., Steiner, D., Lawrence, M. C., McKern, N. M., Pecorari, F., Ward, C. W., Joos, T. O. & Plückthun, A. (2007). *J. Mol. Biol.* **369**, 1015–1028.