Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure, crystal packing and molecular dynamics of the calponin-homology domain of *Schizosaccharomyces pombe* Rng2

Chern-Hoe Wang,[a]* Mohan K. Balasubramanian[b] and Terje Dokland[c]

[a]Institute of Molecular and Cell Biology, 61 Biopolis Drive (Proteos), Singapore 138673, Singapore, [b]Temasek Life Sciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604, Singapore, and [c]Department of Microbiology, University of Alabama at Birmingham, 845 19th St South, BBRB 311, Birmingham, AL 35294, USA

Correspondence e-mail: chernhoe@imcb.a-star.edu.sg

Received 17 March 2004
Accepted 28 May 2004

PDB References: $Rng2_{1-190}$, Br-derivative, 1p2x, r1p2xsf; Hg-derivative, 1p5s, r1p5ssf.

Schizosaccharomyces pombe Rng2 is an IQGAP protein that is essential for the assembly of an actomyosin ring during cytokinesis. Rng2 contains an amino-terminal calponin-homology (CH) domain, 11 IQ repeats and a RasGAP-homology domain. CH domains are known mainly for their ability to bind F-actin, although they have other ligands *in vivo* and there are only few examples of actin-binding single CH domains. The structures of several CH domains have already been reported, but this is only the third report of an actin-binding protein that contains a single CH domain (the structures of calponin and EB1 have been reported previously). The 2.21 Å resolution crystal structure of the amino-terminal 190 residues of Rng2 from Br- and Hg-derivatives includes 40 residues (150–190) carboxyl-terminal to the CH domain that resemble neither the extended conformation seen in utrophin, nor the compact conformation seen in fimbrin, although residues 154–160 form an unstructured coil which adopts a substructure similar to dystrophin residues 240–246 in the carboxyl-terminal portion of the CH2 domain. This region wraps around the stretch of residues that would be equivalent to the proposed actin-binding site ABS1 and ABS2 from dystrophin. This distinctive feature is absent from previously published CH-domain structures. Another feature revealed by comparing the two derivatives is the presence of two loop conformations between Tyr92 and Arg99.

1. Introduction

The calponin-homology (CH) domain is found in many eukaryotic proteins. It was first identified at the amino-terminus of calponin, a protein implicated in the regulation of smooth muscle contraction through its interaction with F-actin and inhibition of the actin-activated MgATPase activity of phosphorylated myosin (Tang *et al.*, 1996). The CH domain is about 100 residues long and may appear as either a single domain or as tandem repeats.

CH domains are commonly involved in actin binding, but many other substrates are found *in vivo* (Banuelos *et al.*, 1998; Gimona *et al.*, 2002; Korenbaum & Rivero, 2002; Stradal *et al.*, 1998). CH domains known to bind F-actin usually appear in the form of tandem repeats known as actin-binding domains (ABDs). These consist of two CH domains connected *via* a short linker, of which the amino-terminal CH domain (CH1) binds F-actin to a greater extent than the carboxyl-terminal CH2 domain (Banuelos *et al.*, 1998; Corrado *et al.*, 1994; Winder *et al.*, 1995). Fimbrins are actin-bundling proteins that contain tandem ABDs (*i.e.* four CH domains). Unlike the ABDs, single CH domains bind a wide variety of ligands (Banuelos *et al.*, 1998; Gimona *et al.*, 2002; Korenbaum & Rivero, 2002; Stradal *et al.*, 1998). For example, the CH

domain of calponin binds calmodulin, caltropin and extracellular regulated kinase (ERK), but not F-actin (Leinweber *et al.*, 1999; Wills *et al.*, 1994). Calponin instead binds F-actin *via* its three calponin family repeats (CFRs), which are carboxyl-terminal to the CH domain (Gimona & Mital, 1998; Mino *et al.*, 1998). Similarly, the carboxyl-terminus of SM22 containing a single CFR is required for full F-actin affinity (Fu *et al.*, 2000).

The fission yeast *Schizosaccharomyces pombe* has three simple F-actin structures: cortical patches, cables and rings (Marks & Hyams, 1985). During the interphase, F-actin patches localize to both growing ends of the cell, whereas F-actin cables run longitudinally and are linked to F-actin patches (Arai & Mabuchi, 2002). During mitosis, the patches disappear and F-actin reappears in an actomyosin ring in the medial region of the cell attached to F-actin cables (Arai & Mabuchi, 2002).

Rng2 is an *S. pombe* protein that is associated with the spindle pole body (SPB) during interphase and mitosis and with the medial actomyosin ring during mitosis and cytokinesis (Eng *et al.*, 1998). The 1489-amino-acid Rng2 protein contains an amino-terminal CH domain, 11 IQ repeats, two coiled-coil domains, a RasGAP domain and a RasGAP carboxyl-terminal domain and is an orthologue of the human protein IQGAP1 (Fig. 1). IQGAP1 binds F-actin and calmodulin *via* residues 1–216, which contain a single CH domain (Fukata *et al.*, 1997; Ho *et al.*, 1999).

X-ray structures are known for several CH domains, including CH2 of β-spectrin (Carugo *et al.*, 1997), the ABDs from dystrophin (Norwood *et al.*, 2000), utrophin (Keep, Norwood *et al.*, 1999; Keep, Winder *et al.*, 1999) and plectin (García-Alvarez *et al.*, 2003), the first ABD (ABD1) from fimbrin (Goldsmith *et al.*, 1997) and the single CH domain from the microtubule-binding protein EB1 (Hayashi & Ikura, 2003). In addition, the NMR structure of a CH domain from calponin has been reported (Bramham *et al.*, 2002). Here, we report the crystal structure of the amino-terminal 190 residues of Rng2 (denoted $Rng2_{1–190}$), including the CH domain and the 40 residues (150–190) carboxyl-terminal to the CH domain that do not bear any sequence similarity to known proteins.

2. Materials and methods

2.1. Crystallization and structure determination

Details concerning the preparation of the protein, crystallization, data collection and structure determination have been published elsewhere (Wang *et al.*, 2003). Briefly, amino acids 1–190 were expressed as a GST-fusion protein and purified by affinity chromatography. After removal of the GST moiety, the protein was crystallized by vapour diffusion in 21% PEG 3000, 0.3 *M* calcium acetate and 0.1 *M* Tris pH 7 at 295 K. After brief fixation with 3% glutaraldehyde, a Br-derivative was prepared by soaking the crystals for 30 min in 21% PEG 3000, 25% glycerol and 1 *M* NaBr. Likewise, an Hg-derivative was made by soaking the crystals for 2 h in saturated $HgCl_2$, 40% glycerol. Diffraction data were collected under cryogenic conditions at ESRF beamline BM14 and were processed with the *HKL* package (Otwinowski & Minor, 1997; Table 1).

SHELXD (Sheldrick, 1998) was used to locate the positions of six anomalous scatterers in the Br-derivative by the two-wavelength anomalous diffraction (MAD) method. *SHARP* (de La Fortelle & Bricogne, 1997) was used to refine the six Br sites, from which the protein phases were derived and used to

Table 1
Data-collection, processing, phasing and refinement statistics.

Values in parentheses refer to the highest resolution shell.

Data-collection, processing and phasing statistics		
Crystal	Rng2 CH + 1 *M* NaBr	Rng2 CH + saturated $HgCl_2$
X-ray source	ESRF BM14	ESRF BM14
Wavelength (Å)	peak, 0.919087; inflection, 0.919338; remote, 0.855057	1.653
Resolution range (Å)	23–2.21 (2.29–2.21)	28–2.21 (2.29–2.21)
Unit-cell parameters		
a (Å)	31.266	30.884
b (Å)	68.836	68.666
c (Å)	39.811	35.309
$\alpha = \gamma$ (°)	90	90
β (°)	105.544	102.514
V_M (Å^3 Da^{-1})	2.3	2.0
Solvent content (%)	46.3	39.3
Observed reflections	25342	130550
Unique reflections	7347 (202)	7034 (570)
Completeness (%)	89.6 (24.8)	97.4 (81.5)
R_{merge}† (%)	5.3 (7.7)	8.4 (14.5)
$I/\sigma(I)$	20.43 (9.39)	32.70 (9.93)
Isomorphous phasing power (acentric)	1.43	—
Anomalous phasing power (acentric)	1.248	1.915
Figure of merit (acentric)	0.51757	0.50549
Refinement statistics		
Resolution range (Å)	23–2.21	21–2.21
Reflections in working set	7173	6571
Reflections in test set	369	319
R_{cryst}‡ (%)	20.7 (27.5)	19.2 (22.0)
R_{free}§ (%)	26.9 (29.3)	24.4 (30.1)
Protein atoms	1303	1303
Derivative atoms	11	4
Water atoms	157	145
R.m.s.d. bond lengths (Å)	0.008	0.007
R.m.s.d. bond angles (°)	1.0	1.2
R.m.s.d. dihedral angles (°)	19.9	23.0
R.m.s.d. improper angles (°)	1.01	1.78
Mean B factor (Å^2)	17.6	22.2

† $R_{merge} = \sum_{hkl} |I(hkl)_{obs} - \overline{I(hkl)}| / \sum_{hkl} I(hkl)_{obs}$. ‡ $R_{cryst} = \sum_{hkl} |F(hkl)_{obs} - F(hkl)_{calc}| / \sum_{hkl} F(hkl)_{obs}$. § $R_{free} = R_{cryst}$ for a random 5% of reflections not used in refinement (test set).

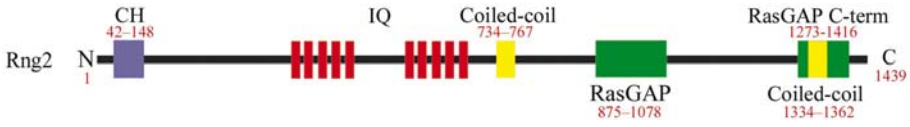

Figure 1
Domain organization within Rng2. Domain boundaries (numbers in red) were defined by performing a global search across a Pfam HMM library (Eddy, 1998).

Table 2
R.m.s.d. values of the pairwise comparison of all the CH domains using the C^{α} atoms belonging to the four core helices of each structure.

The core helices of $Rng2_{1-190}$ correspond to residues 43–54, 73–81, 99–112 and 133–144 (Fig. 2*c*). Abbreviations: DMD, dystrophin; UTRN, utrophin; PLEC, plectin; CNN, calponin; SPTB, spectrin; PLS3, fimbrin ABD1; EB1, end-binding protein 1.

	Rng2 CH	DMD CH1	UTRN CH1	PLEC CH1	CNN CH	PLEC CH2	SPTB CH2	DMD CH2	PLS3 CH1	UTRN CH2	PLS3 CH2	EB1 CH
DMD CH1	1.19	—										
UTRN CH1	1.28	0.47	—									
PLEC CH1	1.30	0.52	0.48	—								
CNN CH	1.45	1.54	1.54	1.54	-—							
PLEC CH2	1.53	1.25	1.18	1.11	1.33	—						
SPTB CH2	1.56	1.23	1.16	1.09	1.40	0.43	—					
DMD CH2	1.63	1.19	1.14	1.01	1.47	0.52	0.60	—				
PLS3 CH1	1.72	1.45	1.29	1.27	1.63	1.41	1.31	1.30	—			
UTRN CH2	1.74	1.44	1.35	1.27	1.46	0.74	0.67	0.82	1.33	—		
PLS3 CH2	1.75	1.33	1.38	1.20	1.59	1.08	1.02	0.92	1.44	1.25	—	
EB1 CH	2.28	2.25	2.36	2.33	2.15	2.19	2.12	2.32	1.84	2.11	2.33	—

trace an initial model with *wARP* (Perrakis *et al.*, 1999). A polyalanine model was built using *O* (Jones *et al.*, 1991). At this point, side chains were not added owing to the poor side-chain densities in the loops. Similarly, *SHELXD* located four anomalous scatterers in the Hg-derivative by the single-wavelength anomalous diffraction (SAD) method. *SHARP* and *wARP* were again used to refine the four Hg sites and to obtain an initial trace of the structure. The three Hg sites with the highest occupancies were used to identify the three cysteine residues, allowing the sequence to be matched to both the Br and Hg F_o maps with confidence. A Br anomalous difference Fourier map was used to locate five additional Br sites.

The two structures containing either 11 Br or four Hg atoms but devoid of water were initially refined with *CNS* (Brünger *et al.*, 1998). The refinement was continued with *SHELXL* (Sheldrick, 1998) using default geometrical and atomic displacement restraints. 157 water molecules were added by inspection of the $2mF_o - DF_c$ map until no further interpretation of the difference ($F_o - F_c$) density contoured at 3σ and -3σ was possible. Only waters within a hydrogen-bonding distance of 3.5 Å were accepted (Ippolito *et al.*, 1990). At the end of the refinement, when the process reached convergence, three cycles of tightly restrained (weight matrix 0.1) isotropic refinement with *REFMAC* (Murshudov *et al.*, 1997) were performed using all reflections, including those previously used for R_{free}, in order to obtain as accurate a model as possible (Dauter *et al.*, 2001; Table 1).

2.2. Molecular-dynamics simulation

Production-phase molecular dynamics was performed with the *AMBER* 7.0 suite of programs (Case *et al.*, 2002) and *PMEMD* (Duke & Pedersen, 2003) with a 2.0 fs time step under the isothermal–isobaric ensemble (300 K and 101.3 kPa), the ff99 force field (Wang *et al.*, 2000), the TIP3P (Jorgensen *et al.*, 1983) model for water, periodic boundary conditions, the particle mesh Ewald method (PME; Darden *et al.*, 1993) for electrostatics, a 10 Å cutoff for Lennard–Jones interactions and the use of *SHAKE* (Ryckaert *et al.*, 1977) for restricting the motion of all covalent bonds involving hydrogen. In an iterative manner, five sodium cations were placed around $Rng2_{32-190}$ near residues with the highest electrostatic potential in order to compensate for the overall negative charge of the molecule. A total of 8763 TIP3P water molecules were then added around residues 32–190 of Rng2, resulting in a box size of 72.535 × 70.410 × 72.229 Å. Temperature was maintained by the Berendsen coupling algorithm (Berendsen *et al.*, 1984) using separate τ coupling constants of 1.0 for the protein and solvent; pressure was maintained with isotropic molecule-based scaling (Berendsen *et al.*, 1984), also with a τ coupling constant of 1.0. The PME grid spacing was about 1.0 Å and was interpolated on a cubic *B*-spline, with the direct sum tolerance set to 10^{-5}. We removed the net centre of velocity every 100 ps in order to correct for the small energy drains that resulted from the use of *SHAKE*, discontinuity in the potential energy near the Lennard–Jones cutoff value and constant-pressure conditions.

For equilibration, we first minimized the solute ($Rng2_{32-190}$) using the steepest-descent method for the first 200 steps, followed by the conjugate-gradient method until the root-mean-square (r.m.s.) of the Cartesian elements of the gradient was <1.7 kJ mol^{-1} Å^{-1}. Water molecules and sodium ions were then minimized in the same way until the r.m.s. was <0.4 kJ mol^{-1} Å^{-1} and then slowly heated, while allowing them to move unrestrained for 25 ps (with a 1.0 fs time step) in order to fill in any vacuum pockets. The solute atoms alone were then minimized in the presence of ever decreasing positional restraints, thereby allowing them to slowly feel the forces of the equilibrated solvent, until the positional restraints reached zero. Finally, a temperature ramp was used to gradually raise the temperature of the whole system to 300 K over 20 ps. Coordinates from the trajectories were saved every 5 ps and analyzed using *AMBER* 7.0.

2.3. Generation of figures and tables

The r.m.s.d. values listed in Table 2 were generated with the help of *Stralign* (http://www.hgc.ims.u-tokyo.ac.jp/service/tooldoc/stralign/intro.html). Other r.m.s.d. values were calcu-

lated with the help of the *MMTSB Tool Set* (Feig *et al.*, 2001). Figs. 2, 3, 4(*a*) and 4(*b*) were prepared using the programs *POVScript+* (Fenn *et al.*, 2003) and *Raster3D* (Merritt & Bacon, 1997). The multiple structure superposition of CH domains shown in Figs. 2(*b*) and 2(*c*) was prepared using *Indonesia* (http://xray.bmc.uu.se/dennis/) using the brute-force method with a cutoff of 10 Å and a fragment length of 100 residues. The crystal contacts shown in Figs. 3(*a*) and 3(*b*) were visualized with the help of *Swiss-PDBViewer* (Guex & Peitsch, 1997).

3. Results and discussion

3.1. Rng2$_{1–190}$ structure

S. pombe Rng2$_{1–190}$ was crystallized in space group $P2_1$, with a single molecule per asymmetric unit. The structures of Br- and Hg-derivatives were determined using multi- and single-wavelength anomalous diffraction (MAD/SAD) methods (Wang *et al.*, 2003; Table 1). Interpretable electron density was observed from residues 32 to 190. The overall structure is compact and globular. 100% of the residues in both derivatives were either in the 'most favoured' or 'additional allowed' regions of Ramachandran space (Laskowski *et al.*, 1993). The overall mean *B* values are 17.51 and 19.89 Å^2 for the Br- and Hg-derivatives, respectively, in reasonable agreement with the Wilson *B* values of 21.4 and 18.3 Å^2. The respective estimated standard uncertainties (e.s.u.) for the *B* values based on maximum likelihood are 5.58 and 5.51 Å^2.

The CH domain resides between residues 40–150 and has a core of four main α-helices (a1, a3, a4, a6) connected by long loops (Fig. 2*b*). There are two short α-helices: a2 and a5. There are also three short stretches of 3_{10}-helices: h1, h2 and h3. α-Helices a3 and a6 are almost parallel to one another, while a1 and a4 are perpendicular to each other. The CH-domain fold is held together by extensive hydrophobic contacts among the α-helices in the core helical structure (Fig. 2*c*). Residues 150–190 do not have a defined secondary structure, except for the α-helical stretch (a7) between Asp167 and Gln179. This stretch wraps around helix a1, forming a substructure that resembles neither the extended conformation seen in utrophin nor the compact conformation seen in fimbrin, although residues 154–160 form an unstructured coil which adopts a substructure similar to dystrophin residues 240–246 in the carboxyl-terminal portion of the CH2 domain. The primary sequence within residues 150–190 is unique to Rng2 and cannot be found elsewhere.

Data from fission yeast showed that Rng2 localizes to the F-actin ring that forms during mitosis (Eng *et al.*, 1998). An F-actin-binding function probably exists within the first 189 residues of Rng2, since an amino-terminal fusion of GFP to

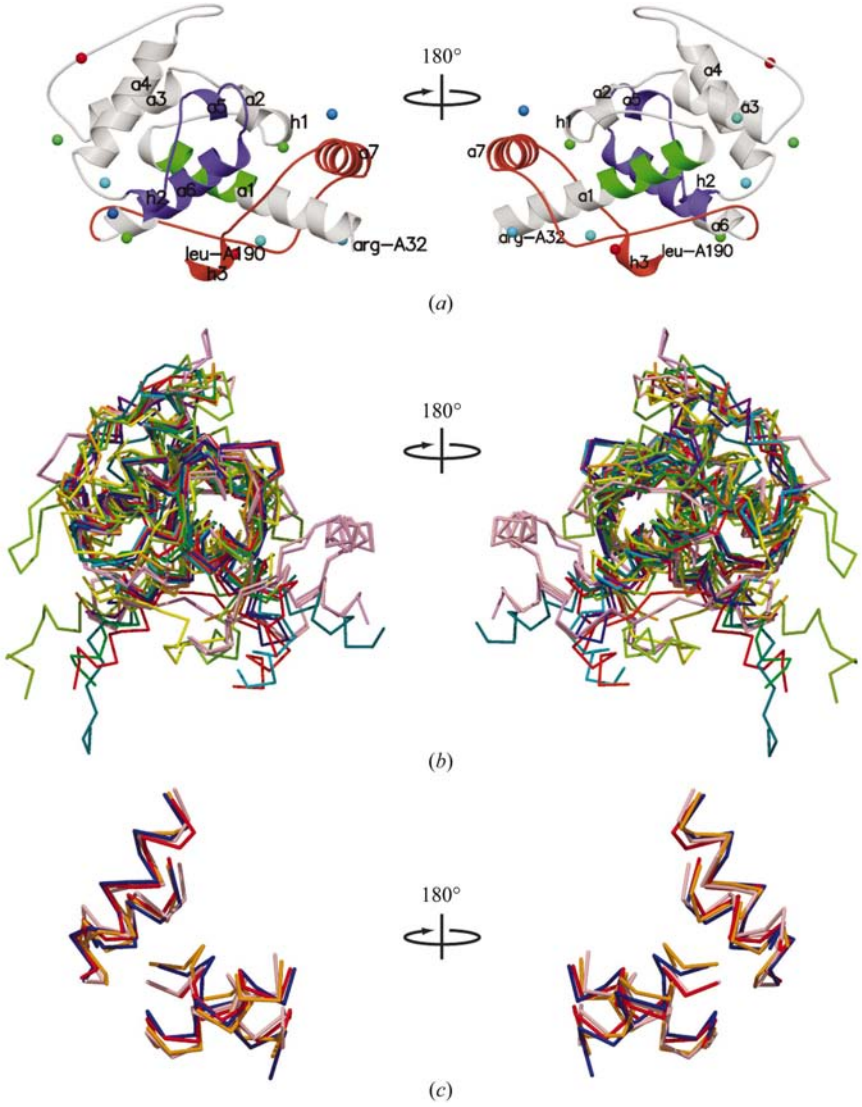

Figure 2
Structure of Rng2$_{1–190}$. (*a*) Ribbon representation of the crystal structure of Rng2$_{1–190}$ shown in two orientations with a relative rotation of 180°. Br atoms are represented by spheres coloured according to *B* factors (blue to red denoting lower to higher values). The amino- and carboxyl-terminal residues are labelled together with α-helices a1–a7 and 3_{10}-helices h1–h3. For comparison purposes, the region corresponding to actin-binding sites ABS1 (a1) and ABS2 (h2–a6) of dystrophin CH1 domain (Norwood *et al.*, 2000) are highlighted in green and blue, respectively. The 40 residues carboxyl-terminal to the CH domain (150–190) are highlighted in red. (*b*) C$^\alpha$-backbone superposition of the CH domains from the Br-derivative of Rng2 CH (pink), Hg-derivative of Rng2 CH (plum), dystrophin CH1 (red), dystrophin CH2 (orange–red), calponin CH (orange), EB1 CH (yellow), fimbrin CH1-1 (yellow–green), fimbrin CH1-2 (green–yellow), utrophin CH1 (green), utrophin CH2 (cyan), plectin CH1 (dark cyan), plectin CH2 (blue) and β-spectrin CH2 (blue–violet). (*c*) C$^\alpha$ backbone superposition of the four core α-helices belonging to the Br-derivative of Rng2$_{1–190}$ CH (pink), dystrophin CH1 (red), calponin CH (orange) and plectin CH2 (blue) domains. Orientations in (*a*), (*b*) and (*c*) are the same.

$Rng2_{1–189}$ localizes to cortical actin patches within the growing cell tips (Wachtler *et al.*, 2003). F-actin co-sedimentation assays indicate a direct interaction with residues 1–216 of human IQGAP1 (Fukata *et al.*, 1997), which contain a single CH domain. Interestingly, residues 150–190 of Rng2 wrap around the proposed ABS1 and ABS2 (Corrado *et al.*, 1994; Norwood *et al.*, 2000; Fig. 2*a*) mainly *via* hydrophobic interactions, suggesting they are not involved in binding F-actin. However, this does not rule out the possibility of a direct interaction between $Rng2_{1–190}$ and F-actin.

A superposition of the CH domains from Rng2, calponin, EB1, fimbrin, dystrophin, utrophin, β-spectrin and plectin (Fig. 2*b*) shows that most of the structural variations are concentrated at the amino- and carboxyl-termini; *i.e.* before residue 43 in helix a1 of Rng2 and beyond residue 148 in helix a6 of Rng2. The loops between helices a1–a2, a3–a4 and h2–a6 also have considerable variation consistent with the alternative loop conformation found in Rng2 residues 92–99 (Fig. 4*a*).

The C^{α} backbone of the four core helices of Rng2, containing 47 equivalent atoms, can be superimposed on those of other solved structures with up to 2.33 Å root-mean-square deviation (r.m.s.d.; Fig. 2*c* and Table 2). The closest match of Rng2 CH domain is with the CH1 of dystrophin, with an r.m.s.d. of 1.19 Å. In contrast, the second CH (CH2) domain of plastin 3/fimbrin has an r.m.s.d. of 1.75 Å when compared with Rng2. The CH1 of dystrophin, utrophin and plectin are all similar, with an r.m.s.d. of about 0.5 Å. This is consistent with their known actin-binding properties.

Superposition of the Rng2 CH domain with the calponin CH domain reveals a relatively large r.m.s.d. of 1.45 Å. The largest deviation is found between the CH domains of Rng2 and end-binding protein 1 (EB1; Hayashi & Ikura, 2003), with an r.m.s.d. of 2.33 Å. It is apparent from Table 2 that the CH domains from Rng2, calponin and EB1 are quite dissimilar to each other and do not cluster together like CH1 and CH2 domains. The Rng2 CH domain is more similar to the F-actin-binding CH1 domains, whereas the calponin CH domain is more similar to CH2 domains. In contrast, the EB1 CH domain is dissimilar to the rest of the CH domains. This agrees with the different functional roles that they play. For instance, Rng2 CH-domain localization correlates with cortical actin patches in fission yeast (Wachtler *et al.*, 2003) and the CH domain from its mammalian orthologue IQGAP binds F-actin and calmodulin (Ho *et al.*, 1999), calponin CH domain binds calmodulin, caltropin and ERK (Wills *et al.*, 1994; Leinweber *et al.*, 1999) and EB1 associates with microtubule filaments (Hayashi & Ikura, 2003).

The Rng2 CH domain has three 3_{10}=MMTSB Tool Set-helices (h1–h3), which are uniformly three residues long and thus only one turn in length (Fig. 2*b*). The dipoles are less well aligned here compared with regular α-helices. The C═O groups are hydrogen bonded to crystallographic waters and

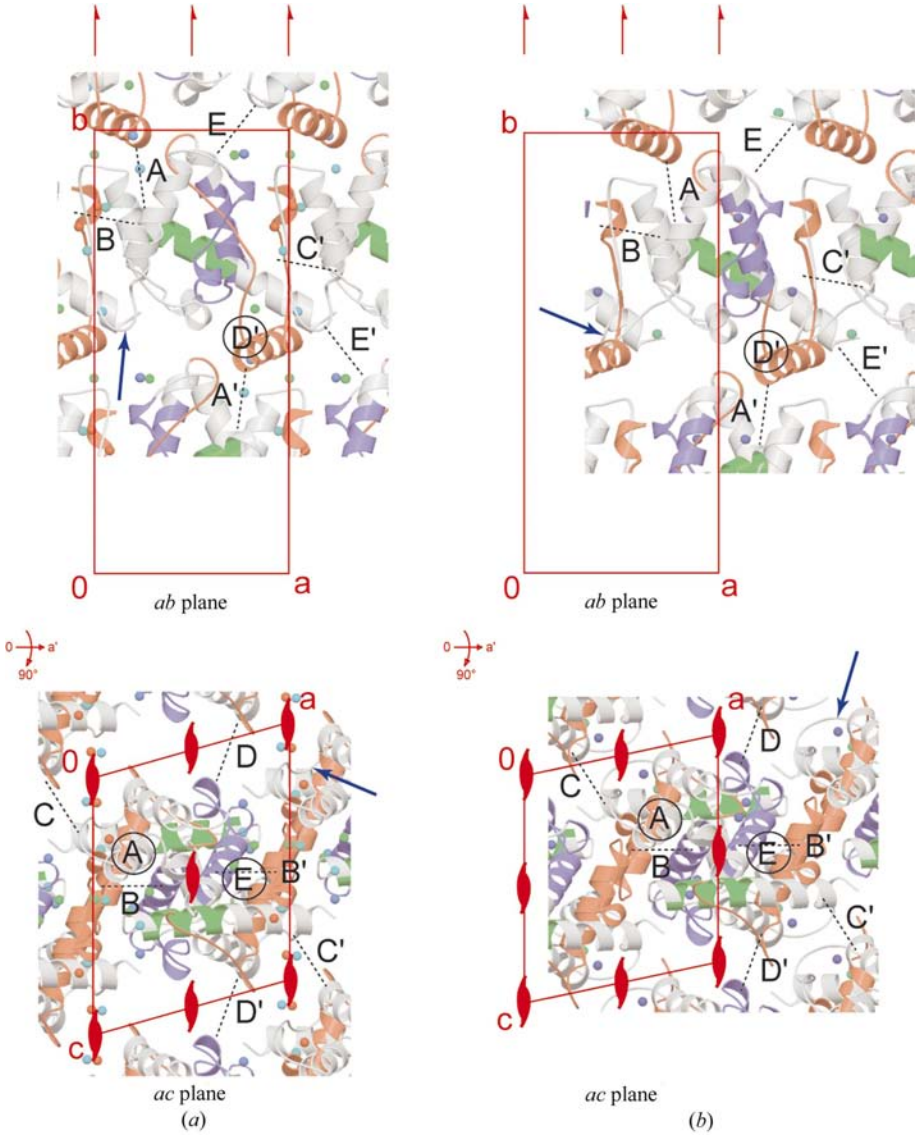

Figure 3
Intermolecular contacts in a $P2_1$ crystal of $Rng2_{1–190}$. (*a*) Molecular packing projected from the *c* axis (upper) and *b* axis (lower) directions for the Br-derivative. The unit cell ($P2_1$) and its related symmetry elements are shown in red. The macrobonds, labelled *A*–*E* are represented by dashed lines for bonds parallel to the projection plane and circles for bonds perpendicular to the projection plane. The colouring scheme of the ribbon model of Rng2 follows that of Fig. 2(*a*). (*b*) As (*a*) for the Hg-derivative. The MAD and SAD data sets for the Br- and Hg-derivatives, respectively, were independently solved. This accounts for the different origins which are related by half a unit translation along the *a* axis. Blue arrows point at the loop between residues 92–99 of Rng2.

point away from the helix axis. The 3_{10}-helices have less favourable side-chain packing and backbone conformation and are usually found in solvent-accessible areas. In the case of Rng2, all the 3_{10}-helices are solvent-accessible. They link random coils or helices together at abrupt angles. This is most clearly seen in helix h2 (Fig. 2*a*).

It has been reported that the crystal structure of proteins subjected to light cross-linking with glutaraldehyde were identical to those without such treatment (Fitzpatrick *et al.*, 1994). Thus, no data were collected from a native crystal. No glutaraldehyde cross-links were observed in the electron-density map. They form randomly and should not occur in high densities if the crystals were cross-linked briefly (Fitzpatrick *et al.*, 1993).

3.2. Intermolecular contacts

In the monoclinic crystal of $Rng2_{1–190}$, each molecule is surrounded by ten symmetry-related molecules within an intermolecular distance of 4 Å (Figs. 3*a* and 3*b*). The direct intermolecular contacts and those mediated *via* crystallographic waters are named macrobonds (Oki *et al.*, 1999), with symmetrically equivalent macrobonds forming the same interatomic interactions. There are five groups of macrobonds, labelled *A*–*E*. The contacting molecules between macrobond *A* are related by a twofold screw axis along the *b* axis. Those between macrobond *B* are related by a unit translation of the repeating lattice unit along the *a* axis. Macrobonds *A* and *B* consist of numerous non-covalent interactions. There is a potential salt bridge and a potential water-mediated hydrogen bond between the ζ-amino group of Arg70 and the γ-carboxyl group of Asp151. However, the temperature factor of the water molecule is large (49.65 Å^2), suggesting local disorder and a weaker hydrogen bond. The ζ-amino group of Arg70 is linked *via* a potential salt bridge to the δ-carbonyl group of Gln67. A bromide anion with a *B* factor of 24.54 Å^2 forms a potential hydrogen bond to the ζ-amino group of Arg175, which is doubly hydrogen bonded to the carbonyl group of Leu152. There is a nonpolar cluster (>5 Å) of residues Ile172, Leu152, Phe82, Phe110 and Ile114 on a background of specific

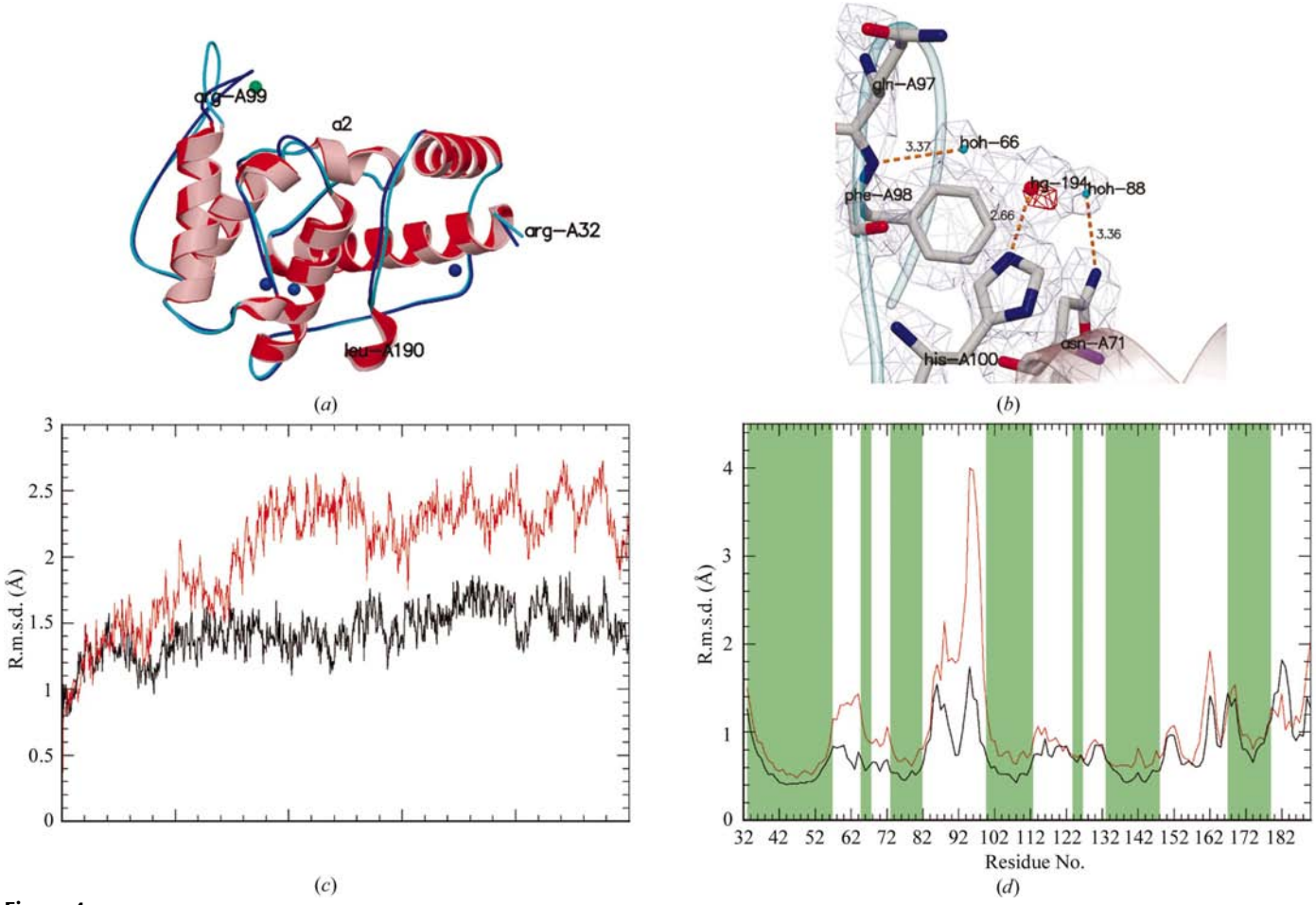

Figure 4
Comparison between Br- and Hg-derivatives of $Rng2_{32–190}$. (*a*) Ribbon representation of the Br-derivative (red helices and blue coils) superimposed on the Hg-derivative (pink helices and cyan coils). Hg atoms are shown as spheres coloured according to their *B* factors. (*b*) Environment around the Hg atom near the loop consisting of residues Tyr92–Arg99 in the Hg derivative. The $2mF_o - DF_c$ map is shown contoured at 1σ. The Hg anomalous map (red) is superimposed onto the electron-density map. (*c*) Mass-weighted C^α r.m.s.d. of $Rng2_{32–190}$ as a function of time over the entire trajectory of a 5 ns molecular-dynamics simulation for $Rng2_{32–190}$ in the Br-derivative (red) and Hg-derivative (black). (*d*) Main-chain (N, C^α, C and O) r.m.s.d. for each residue as a function of time during the same MD. α-Helical sections of the structure are highlighted in green.

intermolecular hydrogen bonds and salt bridges. The inter-residue distances are longer than the usual distance of 3.7 Å (Matsuura & Chernov, 2003) and are unlikely to be van der Waals contacts.

The contacting molecules between macrobond *C* are related by a unit translation along the *c* axis followed by another unit translation along the *a* axis. There is a direct hydrogen bond between the ζ-amino group of Arg99 and the δ-carbonyl group of Glu161. There is also another direct hydrogen bond between the α-imino group of Tyr92 and the γ-carbonyl group of Asn162.

Macrobond *D* consists of contacts made by molecules related by a unit translation along the *c* axis. Molecules between macrobond *E* are related by a lattice translation of one unit along the *a* axis followed by a twofold screw axis along the *b* axis. It has a unique occurrence of a potential bromide-mediated hydrogen bond between the α-imino group of Pro117 and the ζ-amino group of Arg178.

The total numbers of direct hydrogen bonds in the Br-derivative with a length of less than 3.5 Å (Ippolito *et al.*, 1990) between amino-acid residue atoms are 5, 2, 3, 0 and 0, respectively, for macrobonds *A*, *B*, *C*, *D* and *E*. The numbers of bound water molecules involved in the intermolecular hydrogen bonds are 24, 24, 8, 0 and 0, respectively. In line with the relatively high solvent content of protein crystals, water-mediated hydrogen bonds are clearly responsible for holding the crystal together. It is interesting to note that the longest dimension of the unit cell extends in the *b* axis and involves macrobonds *A* and *A*′, which have the most direct hydrogen bonds and potential water-mediated hydrogen bonds. The favourable formation of the periodic bond chain (PBC) *A*–*A*′ results in a tendency towards elongation in this direction (Hartman, 1973). PBCs refer to the intermolecular interactions or bonds occurring periodically across a crystal that contribute to its growth habit. This was indeed observed during our crystal screening process, where we encountered many conditions that resulted in crystals with a fibrous morphology. The *c* axis and *a* axis have about the same length and involve macrobonds *C* and *D*, and *B*, respectively.

The Hg-derivative shows a similar arrangement of intermolecular contacts. This is to be expected since the crystals were derivatized using crystals grown under the same conditions. The total numbers of direct hydrogen bonds in the Hg-derivative with a length of less than 3.5 Å between amino-acid residue atoms are 8, 6, 6, 6 and 1, respectively, for macrobonds *A*, *B*, *C*, *D* and *E*. The numbers of bound water molecules involved in the intermolecular hydrogen bonds are 13, 20, 9, 20 and 4, respectively. The significant increase in the number of direct and water-mediated hydrogen bonds in macrobond *D* relative to the Br-derivative correlates with a significant decrease in unit-cell parameters along the *c* axis.

3.3. Comparison of Br- and Hg-derivatives of $Rng2_{1-190}$

The Br- and Hg-derivatives of $Rng2_{1-190}$ have an r.m.s.d. of 2.14 Å between all 159 residues and 1.37 Å between the C^{α} atoms. This is mainly owing to residues Tyr92–Arg99, which form a loop that points away from helix a2 in the Hg-derivative but points towards helix a2 in the Br-derivative (Fig. 4*a*). The r.m.s.d. between C^{α} atoms fell to 0.52 Å after the loop was omitted from the calculation. The largest deviation among the backbone atoms was seen at Leu96, with an overall r.m.s.d. of 11.53 Å. This contrasts with a comparison of the orthorhombic and tetragonal crystal forms of lysozyme, which have r.m.s.d.s that do not exceed 2 Å (Oki *et al.*, 1999).

There are three cysteine residues between residues 32 and 190 of Rng2. In the Hg-derivative, they are all found to be associated with an Hg atom within 3.2 Å. They have occupancies between 45 and 64%, low temperature factors not exceeding 9.51 $Å^2$ and show strong density exceeding 5σ in the $2mF_o - DF_c$ map (Read, 1986; Fig. 4*b*). The fourth Hg atom (Fig. 4*b*; Hg194) is responsible for the maintaining the loop in an 'open' conformation relative to the Br-derivative. The residues surrounding the fourth Hg atom are Asn71 (4.36 Å relative to C^{β}), Phe98 (3.69 Å to C^{ζ}) and His100 (2.66 Å to N^{ε}). The bond length between His100 N^{ε} and Hg194 is short (2.66 Å), suggesting the presence of a metal–ligand bond in some of the protein molecules in the crystal. This is consistent with Hg194 having a low occupancy of 30% and a high temperature factor of 35.7 $Å^2$, together with weak electron density not exceeding 2σ. The Hg anomalous map (Fig. 4*b*, red density) shows a weak peak and confirms its presence.

In order to compare the crystal structure of both derivatives, a molecular-dynamics (MD) simulation was performed for the aqueous solution system of $Rng2_{32-190}$ in the absence of Br and Hg atoms. After the initial relaxation procedure for 45 ps (not shown), the system representing the conformation seen in the Hg-derivative reached a stationary state within 50 ps after starting the MD simulation, whereas the Br-derivative took about 1.9 ns (Fig. 4*c*). Furthermore, the Hg-derivative had an r.m.s.d of 1.5 Å, whereas the Br-derivative had an r.m.s.d. of 2.5 Å over the 5 ns simulation time. Although both derivatives had some conformational differences in the loop between residues 92–99, both structures remained stable and retained their differences around residues 92–99 throughout the simulation. The average r.m.s.d. of each residue over the 5 ns simulation relative to the starting point is presented in Fig. 4(*d*). It shows that the loops have more motion than the α-helices. The largest motions are seen in the loop between Tyr92–Arg99 of the Br-derivative of Rng2, whereas the Hg-derivative remained relatively stable. The higher r.m.s.d. seen in the Br-derivative relative to the Hg-derivative is consistent both with the slightly higher *B* factors ($\sim$32-51 $Å^2$) at Asn94 of the Br-derivative, compared with the Hg-derivative ($\sim$19-24 $Å^2$); and with funnel diagrams depicting the energetic drive to the native state (Pande *et al.*, 1998). The crystallographic and MD data both suggest that residues 92–99 switch between the Br- and Hg-derivative conformations, and that Hg194 had trapped them in the alternative conformation. The 40 residues (150–190) carboxyl-terminal to the CH domain maintained their conformation throughout the simulation. The 31 residues at the amino-terminus are disordered and were not included in the simulation.

We thank Stephen Wong Yeow Whye and Wong Siew Chin from the Bioinformatics Institute http://www.bii.a-star.edu.sg for allowing us to use their Linux cluster for the molecular-dynamics simulation.

References

Arai, R. & Mabuchi, I. (2002). *J. Cell Sci.* **115**, 887–898.

Banuelos, S., Saraste, M. & Djinovic Carugo, K. (1998). *Structure*, **6**, 1419–1431.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). *J. Chem. Phys.* **81**, 3684–3690.

Bramham, J., Hodgkinson, J. L., Smith, B. O., Uhrin, D., Barlow, P. N. & Winder, S. J. (2002). *Structure*, **10**, 249–258.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Carugo, K. D., Banuelos, S. & Saraste, M. (1997). *Nature Struct. Biol.* **4**, 175–179.

Case, D. A. *et al.* (2002). *AMBER* 7. University of California, San Francisco, USA.

Corrado, K., Mills, P. L. & Chamberlain, J. S. (1994). *FEBS Lett.* **344**, 255–260.

Darden, T., York, D. & Pedersen, L. (1993). *J. Chem. Phys.* **98**, 10089–10092.

Dauter, Z., Li, M. & Wlodawer, A. (2001). *Acta Cryst.* D**57**, 239–249.

Duke, R. E. & Pedersen, L. G. (2003). *PMEMD* 3. University of North Carolina, Chapel Hill, USA.

Eddy, S. R. (1998). *Bioinformatics*, **14**, 755–763.

Eng, K., Naqvi, N. I., Wong, K. C. Y. & Balasubramanian, M. K. (1998). *Curr. Biol.* **8**, 611–621.

Feig, M., Karanicolas, J. & Brooks, C. L III. (2001). *MMTSB Tool Set*, MMTSB NIH Research Resource. The Scripps Research Institute, La Jolla, CA, USA.

Fenn, T. D., Ringe, D. & Petsko, G. A. (2003). *J. Appl. Cryst.* **36**, 944–947.

Fitzpatrick, P. A., Ringe, D. & Klibanov, A. M. (1994). *Biochem. Biophys. Res. Commun.* **198**, 675–681.

Fitzpatrick, P. A., Steinmetz, A. C., Ringe, D. & Klibanov, A. M. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 8653–8657.

Fu, Y., Liu, H. W., Forsythe, S. M., Kogut, P., McConville, J. F., Halayko, A. J., Camoretti-Mercado, B. & Solway, J. (2000). *J. Appl. Physiol.* **89**, 1985–1990.

Fukata, M., Kuroda, S., Fujii, K., Nakamura, T., Shoji, I., Matsuura, Y., Okawa, K., Iwamatsu, A., Kikuchi, A. & Kaibuchi, K. (1997). *J. Biol. Chem.* **272**, 29579–29583.

García-Alvarez, B., Bobkov, A., Sonnenberg, A. & de Pereda, J. M. (2003). *Structure*, **11**, 615–625.

Gimona, M., Djinovic Carugo, K., Kranewitter, W. J. & Winder, S. J. (2002). *FEBS Lett.* **513**, 98–106.

Gimona, M. & Mital, R. (1998). *J. Cell Sci.* **111**, 1813–1821.

Goldsmith, S. C., Pokala, N., Shen, W., Fedorov, A. A., Matsudaira, P. & Almo, S. C. (1997). *Nature Struct. Biol.* **4**, 708–712.

Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.

Hartman, P. (1973). *Crystal Growth: An Introduction*, edited by P. Hartman, ch. 14. Amsterdam: North-Holland.

Hayashi, I. & Ikura, M. (2003). *J. Biol. Chem.* **278**, 36430–36434.

Ho, Y. D., Joyal, J. L., Li, Z. & Sacks, D. B. (1999). *J. Biol. Chem.* **274**, 464–470.

Ippolito, J. A., Alexander, R. S. & Christianson, D. W. (1990). *J. Mol. Biol.* **215**, 457–471.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. & Klein, M. L. (1983). *J. Chem. Phys.* **79**, 926–935.

Keep, N. H., Norwood, F. L. M., Moores, C. A., Winder, S. J. & Kendrick-Jones, J. (1999). *J. Mol. Biol.* **285**, 1257–1264.

Keep, N. H., Winder, S. J., Moores, C. A., Walke, S., Norwood, F. L. & Kendrick-Jones, J. (1999). *Structure*, **7**, 1539–1546.

Korenbaum, E. & Rivero, F. (2002).*J. Cell Sci.* **115**, 3543–3545.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leinweber, B. D., Leavis, P. C., Grabarek, Z., Wang, C. L. & Morgan, K. G. (1999). *Biochem. J.* **344**, 117–123.

Marks, J. & Hyams, J. S. (1985). *Eur. J. Cell. Biol.* **39**, 27–32.

Matsuura, Y. & Chernov, A. A. (2003). *Acta Cryst.* D**59**, 1347–1356.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Mino, T., Yuasa, U., Nakamura, F., Naka, M. & Tanaka, T. (1998). *Eur. J. Biochem.* **251**, 262–268.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Norwood, F. L. M., Sutherland-Smith, A. J., Keep, N. H. & Kendrick-Jones, J. (2000). *Structure*, **8**, 481–491.

Oki, H., Matsuura, Y., Komatsu, H. & Chernov, A. A. (1999). *Acta Cryst.* D**55**, 114–21.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pande, V. S., Grosberg, A. Y., Tanaka, T. & Rokhsar, D. S. (1998). *Curr. Opin. Struct. Biol.* **8**, 68–79.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). *J. Comput. Phys.* **23**, 327–341.

Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.

Stradal, T., Kranewitter, W., Winder, S. J. & Gimona, M. (1998). *FEBS Lett.* **431**, 134–137.

Tang, D. C., Kang, H. M., Jin, J. P., Fraser, E. D. & Walsh, M. P. (1996). *J. Biol. Chem.* **271**, 8605–8611.

Wachtler, V., Rajagopalan, S. & Balasubramanian, M. K. (2003). *J. Cell Sci.* **116**, 867–874.

Wang, C. H., Walsh, M., Balasubramanian, M. K. & Dokland, T. (2003). *Acta Cryst.* D**59**, 1809–1812.

Wang, J., Cieplak, P. & Kollman, P. A. (2000). *J. Comput. Chem.* **21**, 1049–1074.

Wills, F. L., McCubbin, W. D., Gimona, M., Strasser, P. & Kay, C. M. (1994). *Protein Sci.* **3**, 2311–2321.

Winder, S. J., Hemmings, L., Maciver, S. K., Bolton, S. J., Tinsley, J. M., Davies, K. E., Critchley, D. R. & Kendrick-Jones, J. (1995). *J. Cell Sci.* **108**, 63–71.