STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN 2053-230X

Structure of an ADP-ribosylation factor, ARF1, from *Entamoeba histolytica* bound to Mg^{2+} -GDPDmitry A. Serbzhinskiy,^{a,b,*} Matthew C. Clifton,^{a,c} Banumathi Sankaran,^d Bart L. Staker,^{a,b} Thomas E. Edwards^{a,c} and Peter J. Myler^{a,b,e}^aSeattle Structural Genomics Center for Infectious Disease (SSGCID), USA, ^bSeattle Biomedical Research Institute, Seattle, WA 98109, USA, ^cBeryllium West Coast Operations, 7869 NE Day Road West, Bainbridge Island, WA 98110, USA, ^dBerkeley Center for Structural Biology, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA, and ^eDepartments of Global Health and Medical Education and Biomedical Informatics, University of Washington, Seattle, WA 98195, USA. *Correspondence e-mail: dmitry.serbzhinskiy@seattlebiomed.org

Received 1 December 2014

Accepted 6 March 2015

Edited by W. N. Hunter, University of Dundee, Scotland

Keywords: GDP-ribosylation factor; small GTPase; ARFA1; structural genomics; Seattle Structural Genomics Center for Infectious Disease; SSGCID; signaling protein.**PDB reference:** ARF1 bound to Mg-GDP, 4y0v**Supporting information:** this article has supporting information at journals.iucr.org/f

Entamoeba histolytica is the etiological agent of amebiasis, a diarrheal disease which causes amoebic liver abscesses and amoebic colitis. Approximately 50 million people are infected worldwide with *E. histolytica*. With only 10% of infected people developing symptomatic amebiasis, there are still an estimated 100 000 deaths each year. Because of the emergence of resistant strains of the parasite, it is necessary to find a treatment which would be a proper response to this challenge. ADP-ribosylation factor (ARF) is a member of the ARF family of GTP-binding proteins. These proteins are ubiquitous in eukaryotic cells; they generally associate with cell membranes and regulate vesicular traffic and intracellular signalling. The crystal structure of ARF1 from *E. histolytica* has been determined bound to magnesium and GDP at 1.8 Å resolution. Comparison with other structures of eukaryotic ARF proteins shows a highly conserved structure and supports the interswitch toggle mechanism of communicating the conformational state to partner proteins.

1. Introduction

Amebiasis is a diarrheal disease caused by the anaerobic protozoan parasite *Entamoeba histolytica*. Diarrheal diseases are a significant global health problem and a major cause of childhood mortality and morbidity. Globally, one in ten deaths of children under five years is the result of diarrheal disease (Haque *et al.*, 2003). Approximately 50 million people are infected worldwide with *E. histolytica*, with the highest incidences in Central and South America, Africa and South Asia. Although only 10% of infected people develop symptomatic amebiasis, there are still an estimated 100 000 deaths each year (Pritt & Clark, 2008).

Infection with *E. histolytica* usually results from ingestion of fecally contaminated food or water containing the infective cyst form of the parasite. Upon passage through the stomach and small intestine, excystation occurs, mobilizing the invasive trophozoites. The trophozoites colonize the intestinal mucin layer, where new cysts will eventually be produced and secreted by the host (Haque *et al.*, 2003). Once the intestinal epithelium has been invaded, extra-intestinal spread to the liver or other sites may occur. In severe or chronic cases, amebiasis results in liver abscesses and colitis. The current treatment of amebiasis is dominated by nitroimidazole drugs such as metronidazole (Bansal *et al.*, 2006). Metronidazole is a frequently prescribed medication for parasitic and bacterial infections in both human and veterinary applications. The emergence of nitroimidazole resistance in bacterial

© 2015 International Union of Crystallography

Table 1

Macromolecule-production information.

Source organism	<i>E. histolytica</i>
DNA source	<i>E. histolytica</i> HM-1:IMSS cDNA
Forward primer	GGGTCCTGGTTCGATGGGAAGTTGGTTAAGTAAA-TTACT
Reverse primer	CTTGTCGCTGCTGTTTATTATTGAGATTATCAG-CAAGCCAATCA
Cloning vector	AVA0421 (in non-expression <i>E. coli</i> cells)
Expression vector	AVA0421
Expression host	<i>E. coli</i> BL21(DE3)R3 Rosetta
Complete amino-acid sequence of the cleaved protein used for crystallization	GPMSGWSLSKLLGKKEMRILMVGLDAAGKTSIL-YKLKLGIEIVTTIPTIGFNVETVEYKNISFTVW-DVGGQDKIRPLWRHYQNTQATIFVVDSDNRD-RIGEAEREELMKMLNEDEMRNAILLVFANKHDL-PQAMSISEVTEKLGLOTIKNRKWCQTSCATN-GDGLYEGLDWLADNLK

populations heralds the potential of resistance in parasites (Bansal *et al.*, 2006; Tanwar *et al.*, 2014).

Intracellular protein trafficking is a critical component of the complex life cycle of *E. histolytica*. The small ADP-ribosylation factor (ARF) proteins are major regulators of vesicle formation in intracellular trafficking (D'Souza-Schorey & Chavrier, 2006). Deciphering the mechanisms of intracellular trafficking is an important step towards understanding the complicated lifecycle of *E. histolytica* and may lead to opportunities for identifying new treatments. In this paper, we present the crystal structure of the ADP-ribosylation factor from *E. histolytica* (EhARF1) and compare the structure with orthologous structures in both GDP-bound and GTP-bound conformations.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Cloning. Cloning, expression and purification followed standard Seattle Structural Genomics Center for Infectious Disease (SSGICD; Myler *et al.*, 2009; Phan *et al.*, 2014) protocols as described previously (Bryan *et al.*, 2011; Choi *et al.*, 2011). The 174-residue *E. histolytica* gene for putative ADP-ribosylation factor 1 (ARF1; UniProt accession code C4LXL1) was amplified from genomic DNA and cloned into an expression vector (pAVA0421) encoding an N-terminal six-histidine affinity tag followed by the human rhinovirus 3C protease cleavage sequence using ligation-independent cloning (LIC; Aslanidis & de Jong, 1990). The sequence of the entire tag is MAHHHHHHMGTLEAQTQGPSS, which is

followed by the 174-residue *E. histolytica* ARF1. Gel-extracted and column-purified PCR inserts were treated with T4 DNA polymerase (T4DNAP) in 96-well plates in the presence of dTTP. The LIC-ready inserts were mixed with LIC-ready AVA0421 vector (Table 1). The resulting LIC products then were transformed into an *Escherichia coli* host designed for the amplification of plasmid DNA and plated into 12-lane LB-Agar grids with antibiotic markers. Single colonies from each transformation were grown in LB with antibiotics in 96-well plates overnight. After harvesting, plasmid DNA was isolated from the overnight cultures using 96-well plate plasmid mini-prep kits.

2.1.2. Expression. Successfully cloned plasmid DNA was transformed into chemically competent *E. coli* BL21(DE3)R3 Rosetta cells. Several clones were then tested to choose the colony with the best ability to express the target protein. Glycerol stock and starter cultures were made using this clone. Starter cultures were grown overnight with antibiotics and used to inoculate five 2 l flasks with auto-induction medium. The cultures were grown for 6 h on a 37°C shaker followed by culture at 18°C for 60 h. Six aliquots of 1 ml each were taken, harvested by centrifugation and frozen for future analysis. Cells were harvested by centrifugation in 1 l centrifuge bottles at 4000 rev min⁻¹ in an F9 rotor at 4°C for 30 min. Pellets were resuspended in 25 ml ice-cold 1× SGPP buffer, transferred into 4× 50 ml conical tubes and re-pelleted by centrifugation at 5000 rev min⁻¹ in a swinging-bucket rotor at 4°C. Liquid was decanted and the cell pellets were flash-cooled in liquid nitrogen and stored at -80°C.

2.1.3. Purification. A four-step protocol was conducted composed of an Ni²⁺-affinity chromatography (IMAC) step, cleavage of the N-terminal His tag with 3C protease, removal of the cleaved His tag by a second round of Ni²⁺-affinity chromatography and size-exclusion chromatography (SEC). All chromatography runs were performed on an ÄKTA-purifier 10 (GE) using automated IMAC and SEC programs according to previously described procedures (Bryan *et al.*, 2011).

Thawed bacterial pellets were placed into a beaker with 200 ml lysis buffer (1× SGPP, 30 mM imidazole, 5% MOPS, 2% MgCl₂). This mixture was then sonicated for 15 min. Sonication cycles consisted of a 5 s pulse and a 10 s pause. The sonicated suspension was incubated with Benzonase nuclease (Millipore) at room temperature for 40 min and then

Figure 1

Multiple sequence alignment shows a high similarity of ARF1 among eukaryotic species. Reference sequences are from *E. histolytica* (XP_654041.1), *E. invadens* (XP_004257850.1), *Homo sapiens* (NP_001649.1), *Arabidopsis thaliana* (NP_182239.1) and *Saccharomyces cerevisiae* (NP_010089.1). Multiple sequence alignment was conducted using *Clustal Omega* (Sievers *et al.*, 2011) and the figure was generated using *BoxShade* v.3.21.

Table 2

Data-collection statistics.

Values in parentheses are for the highest of 20 resolution shells.

Ligand	GDP-Mg ²⁺
Space group	C2
Unit-cell parameters (Å, °)	$a = 123.54$, $b = 40.66$, $c = 78.95$, $\beta = 97.45$
Wavelength (Å)	1.0
Resolution range (Å)	50.0–1.80 (1.85–1.80)
No. of unique reflections	34091 (2600)
Multiplicity	4.12 (4.55)
Completeness (%)	93.4 (96.9)
CC _{1/2}	99.8 (92.2)
R _{int} (%)	6.6 (44.8)
Mean I/ σ (I)	15.97 (3.24)

Table 3

Refinement and model statistics.

Ligand	GDP-Mg ²⁺
Resolution range (Å)	19.7–1.80 (1.85–1.80)
R _{cryst} (%)	22.5
R _{free} (%)†	27.2
R.m.s.d., bonds (Å)	0.007
R.m.s.d., angles (°)	1.14
Protein atoms	2555
Nonprotein atoms	307
Mean B factor (Å ²)	25.0
Residues in favoured region (%)	98.8
Residues in allowed region (%)	1.2
PDB code	4ylg

$$\dagger R_{\text{free}} = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$$

centrifuged at 10 000 rev min^{−1} for 1 h using a Sorvall centrifuge (Thermo Scientific). Centrifuged and filtered lysate was run *via* an automated IMAC program on an ÄKTApurifier 10 (GE) using a HisTrap HP 5 × 5 ml column (GE). IMAC purification peak fractions were collected and concentrated using an Amicon purification system (Millipore). Cleavage of the N-terminal His tag was accomplished by dialysis of the target protein with His-MBP-3C protease at 4°C overnight in 3C reaction buffer. The reaction mixture was passed over a

Figure 2

ARF1 from *E. histolytica* is structurally similar to human ARF1. The structure of ARF1 from *E. histolytica* (gray) is overlaid with ARF1 from *H. sapiens* (PDB entry 1hur; orange) in the magnesium GDP-bound state (the superposition r.m.s.d. of PDB entries 4ylg and 1hur is 0.50 Å overall, calculated on all common C α atoms). Similar to the human structure of ARF1, the interswitch region is largely disordered in the magnesium GDP-bound state.

second Ni²⁺-affinity column to remove His-MBP-3C protease, uncleaved protein and cleaved His tag. The cleaved protein was recovered in the flowthrough and concentrated prior to loading onto the final SEC column (Superdex 75, GE). After an automated SEC run, the peak fractions were collected and analysed for the presence of the protein of interest using SDS-PAGE. Comparison of protein elution peaks with molecular-weight standards showed a decreased retention time on the SEC column, indicating multimerization of the protein. The peak fractions eluted in the molecular-mass range 120–210 Da, suggesting multimeric complexes of tetramers or larger (data not shown). The peak fractions were pooled, concentrated and flash-cooled in liquid nitrogen and stored at −80°C until use for crystallization.

2.2. Crystallization, X-ray data collection and processing

Purified *E. histolytica* ARF1 (EnhiA.01533.a; 13.5 mg ml^{−1}) was screened for crystallization in 96-well sitting-drop plates against the Wizard I and II crystal screens (Rigaku Reagents). Equal volumes of protein solution (0.4 μ l) and precipitant solution were set up at 289.15 K against reservoir (80 μ l) in sitting-drop vapor-diffusion format. The final crystallization precipitant was Wizard I condition No. 42 consisting of 15% ethanol, 0.1 M Tris pH 7.0 at 289.15 K. The crystals were cryoprotected in crystallant plus 25% ethylene glycol and flash-cooled using liquid nitrogen. Data were collected at 100°C on Advanced Light Source beamline 5.0.3 using an ADSC Quantum 315 CCD detector with 1° oscillations and a crystal-to-detector distance of 220 cm. Data were reduced with XDS/XSCALE (Kabsch, 2010). Phases were obtained by molecular replacement using Phaser from the CCP4 suite of programs using PDB entry 1r8s as a search model (McCoy, 2007; Winn *et al.*, 2011; Renault *et al.*, 2003). The structure was refined using multiple cycles of phenix.refine (Afonine *et al.*, 2012) followed by manual rebuilding of the structure using Coot (Emsley *et al.*, 2010). The quality of all of the structures

Figure 3

The structure of ARF1 from *E. histolytica*. The protein forms a homodimer (light gray and dark gray). The G1 loop (also called the P-loop; red) creates a tight packing interaction surrounding the GDP and magnesium (green). Switch 1 (orange) creates the interface for the packing of the homodimer and the switch 2 region is only partially ordered (purple). There are multiple water-mediated contacts between ARF1, GDP and magnesium (water, red spheres; Mg²⁺, green spheres).

was checked using *MolProbity* (Chen *et al.*, 2010). All data-reduction and refinement statistics, as shown in Tables 2 and 3, have been deposited in the PDB (PDB entry 4ylg). Figures, overlays and electrostatic surface potentials were created using *PyMOL* (v.1.5.0.4; Schrödinger).

3. Results and discussion

3.1. Comparison of selected orthologs

ADP-ribosylation factor (ARF) proteins are ubiquitous in all eukaryotic cells and are housekeeping proteins that are essential for normal cell function. ARF1 from *E. histolytica* (*Eh*ARF1) is structurally homologous to other ARF proteins from diverse species. Overall, there is high homology in both sequence and structure in this enzyme family, suggesting strong evolutionary pressure to conserve the protein fold (Li *et al.*, 2004; Price *et al.*, 1996). The areas with the highest structural conservation include the residues responsible for binding guanine nucleotides (G-loops) as well as residues in the hydrophobic core of the protein. *Eh*ARF1 shares 77–83% identity in its amino-acid sequence with other species across different taxonomic units such as types and even kingdoms: plants, yeasts, mammals and protozoa (Fig. 1). The closest known amino-acid sequence is that from *E. invadens*, which shares about 97% identity. The human ortholog ARF1 shares 79% sequence identity with *Eh*ARF1. The structure of the GDP-bound *Eh*ARF1 subunit (PDB entry 4ylg) superimposed on that of GDP-bound human ARF1 (PDB entry 1hur; Amor *et al.*, 1994) shows significant structural similarity,

with an r.m.s.d. of 0.5 Å between C α atoms (Fig. 2). The overall architecture of the structure and substrate-binding center are almost identical.

3.2. Crystal structure of *Eh*ARF1

A 1.8 Å resolution X-ray data set was collected for the *Eh*ARF1–Mg²⁺–GDP structure. The final model has a crystallographic R_{cryst} value of 22.3% and an R_{free} value of 27.2% (Table 3). The asymmetric unit of the crystal in space group C2 is formed by two molecules oriented so that the GDP-binding site from each molecule forms crystal-packing contacts through a β -sheet hydrogen-bonding pattern (Fig. 3). Each monomer contains five α -helices, seven β -strands and four 3_{10} -helices. The amino-acid residues forming the GDP-binding site are highly conserved (Bourne *et al.*, 1991) and are predominantly located in the second α -helix and the second and third β -strands (IVTTIPT–NVETVEY; Fig. 4). Important elements in small GTPases, including the ARF family, are the so-called G-loops: 6–8 amino-acid sequences which are responsible for the functional activity of small G-proteins (Paduch *et al.*, 2001). The G-loops (especially G1 and G4) are more highly conserved in ARF family proteins compared with the secondary-structure units (Paduch *et al.*, 2001). The G1 loop, with the sequence GLDAAGKT, is located between the β 1 strand and the α 2 helix and is responsible for nucleotide binding (Figs. 3 and 4). A conserved threonine (Thr27) is known to play a key role in Mg²⁺ ion binding (Paduch *et al.*, 2001; Fig. 5). The G2 loop contains part of the second β -strand and the loop between the β 2 and β 3 strands. This functional

Figure 4

Structural elements of ARF1 from *E. histolytica* aligned with its primary structure. Red boxes, α -helices; blue arrows, β -strands. G1–G5, G-loops. Residues directly responsible for GDP binding are marked with asterisks. Switches I and II and the interswitch region are marked individually.

region, called switch I (Pasqualato *et al.*, 2002), plays an important role in protein conformational changes. The G3 loop with the DXXG conserved motif is comprised of the C-terminal part of the β 4 strand and the following polypeptide loop with unspecified secondary structure. This region is also called switch II (Figs. 3, 4 and 6). The G4 and G5 loops are known for their ability to bind a guanine nucleotide base (Sprang, 1997). Lysine and aspartate residues in G4 are

responsible for interaction with the nucleotide, while the G5 loop is partially responsible for guanine base recognition.

3.3. Substrate recognition and comparison with GTP-bound human ARF1

ARF proteins are members of the small GTPases superfamily and bind GTP and GDP with two independent conformations. Hydrolysis of GTP to GDP and exchange of GDP for GTP rely on additional effector proteins and have been reviewed previously (Vetter & Wittinghofer, 2001). The most widely studied members of the ARF family, human ARF1–ARF6, have been shown to be active in their GTP-bound conformation, where they recruit coat proteins and form vesicles.

Because no GTP-bound structure of *Eh*ARF1 is available and the functional regions of this protein are highly conserved, we compared *Eh*ARF1–GDP with the GTP-bound structure of human ARF1 (PDB entry 1hur; Amor *et al.*, 1994). This comparison revealed significant conformational changes between the two forms (Fig. 3). The switch I residues of the GDP-bound *Eh*ARF1 fold over the back side of the molecule, similar to an open mousetrap. The switch I residues contact the opposing switch I residues in the noncrystallographic molecule, forming β -sheet hydrogen-bonding interactions. In the GTP-bound form of human ARF1 the switch I residues between β 2 and β 3 are noticeably further from the rest of the structure, forming a shape similar to a closed mousetrap, in which the switch I residues now lie over the top of the GTP. This conformation makes it impossible for the switch I residues to maintain contact with the other molecule in the asymmetric unit, thus breaking the putative dimer interface.

Figure 5

ARF1 binds to magnesium and GDP utilizing a large hydrogen-bond network. GDP is mainly bound to the individual monomer, creating hydrogen bonds (black dashes) to both of the phosphates as well as the purine ring. Thr27 and an O atom of the β -phosphate aid in binding to magnesium (green). Asn48 from the interswitch region of the adjacent monomer closes part of the GDP-binding site, creating a single hydrogen bond to the α -phosphate.

Figure 6

Comparison of ARF1 from *E. histolytica* bound to GDP (left, gray) with human ARF1 bound to GTP (PDB entry 4hmy; Ren *et al.*, 2013; right, cyan). The conformation of the P-loop (red) does not change upon GTP hydrolysis; however, there is a large conformational change in both the switch 1 (orange) and switch 2 (purple) regions. Although disordered in the GDP-bound structure, in the presence of GTP the switch 2 region adopts an α -helical conformation in the human ARF1 structure (right). Additionally, the switch 2 region that interacts with the γ -phosphate of GTP changes conformation to create a β -strand in the GDP-bound form.

Conformational changes in the switch I and switch II regions may be responsible for interaction with specific GTP/GDP-exchange proteins (GEPs) and GTPase-activating proteins (GAPs) (Donaldson & Jackson, 2000). In addition to changes on the GDP/GTP face of the protein, conformational changes are propagated to the opposite side of the molecule through a rearrangement of the central β -sheet, termed the interswitch toggle (Pasqualato *et al.*, 2002).

Given the sequence and structural conservation between GDP-bound forms of *Eh*ARF1 and human ARF1, it is likely that the GTP-bound conformation of *Eh*ARF1 will undergo similar conformational changes as the human protein. However, the exact mechanism of function of *Eh*ARF1 and its interaction with other parts of the cellular traffic regulation machinery is unknown. Further experiments and data analysis will be needed for a clearer understanding of these biological mechanisms.

Acknowledgements

The authors wish to thank the entire SSGCID team. This project has been funded in whole with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under Contract No. HHSN272201200025C. We also wish to thank the Advanced Light Source, which is supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst. D* **68**, 352–367.
- Amor, J. C., Harrison, D. H., Kahn, R. A. & Ringe, D. (1994). *Nature (London)*, **372**, 704–708.
- Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
- Bansal, D., Malla, N. & Mahajan, R. C. (2006). *Indian J. Med. Res.* **123**, 115–118.
- Bourne, H. R., Sanders, D. A. & McCormick, F. (1991). *Nature (London)*, **349**, 117–127.
- Bryan, C. M., Bhandari, J., Napuli, A. J., Leibly, D. J., Choi, R., Kelley, A., Van Voorhis, W. C., Edwards, T. E. & Stewart, L. J. (2011). *Acta Cryst. F* **67**, 1010–1014.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Choi, R., Kelley, A., Leibly, D., Nakazawa Hewitt, S., Napuli, A. & Van Voorhis, W. (2011). *Acta Cryst. F* **67**, 998–1005.
- Donaldson, J. G. & Jackson, C. L. (2000). *Curr. Opin. Cell Biol.* **12**, 475–482.
- D'Souza-Schorey, C. & Chavrier, P. (2006). *Nature Rev. Mol. Cell Biol.* **7**, 347–358.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Haque, R., Huston, C. D., Hughes, M., Hout, E. & Petri, W. A. Jr (2003). *N. Engl. J. Med.* **348**, 1565–1573.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 133–144.
- Li, Y., Kelly, W. G., Logsdon, J. M. Jr, Schurko, A. M., Harfe, B. D., Hill-Harfe, K. L. & Kahn, R. A. (2004). *FASEB J.* **18**, 1834–1850.
- McCoy, A. J. (2007). *Acta Cryst. D* **63**, 32–41.
- Myler, P. J., Stacy, R., Stewart, L., Staker, B. L., Van Voorhis, W. C., Varani, G. & Buchko, G. W. (2009). *Infect. Disord. Drug Targets*, **9**, 493–506.
- Paduch, M., Jeleń, F. & Otlewski, J. (2001). *Acta Biochim. Pol.* **48**, 829–850.
- Pasqualato, S., Renault, L. & Cherfils, J. (2002). *EMBO Rep.* **3**, 1035–1041.
- Phan, I. Q. H., Stacy, R. & Myler, P. J. (2014). *Methods Mol. Biol.* **1140**, 53–59.
- Price, S. R., Nightingale, M. S., Tsuchiya, M., Moss, J. & Vaughan, M. (1996). *Mol. Cell. Biochem.* **159**, 15–23.
- Pritt, B. S. & Clark, C. G. (2008). *Mayo Clin. Proc.* **83**, 1154–1160.
- Ren, X., Farias, G. G., Canagarajah, B. J., Bonifacino, J. S. & Hurley, J. H. (2013). *Cell*, **152**, 755–767.
- Renault, L., Guibert, B. & Cherfils, J. (2003). *Nature (London)*, **426**, 525–530.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Sprang, S. R. (1997). *Curr. Opin. Struct. Biol.* **7**, 849–856.
- Tanwar, J., Das, S., Fatima, Z. & Hameed, S. (2014). *Interdiscipl. Perspect. Infect. Dis.* **2014**, 541340.
- Vetter, I. R. & Wittinghofer, A. (2001). *Science*, **294**, 1299–1304.
- Winn, M. D. (2011). *Acta Cryst. D* **67**, 235–242.