Molecular & Biochemical Parasitology 174 (2010) 141–144

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Short communication

Crystal structure of *Leishmania major* ADP-ribosylation factor-like 1 and a classification of related GTPase family members in this Kinetoplastid

Jennifer R. Fleming, Alice Dawson, William N. Hunter*

Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK

ARTICLE INFO

Article history:
Received 8 March 2010
Received in revised form 11 August 2010
Accepted 13 August 2010
Available online 27 August 2010

Keywords:
ADP-ribosylation factor-like
GTPase
Leishmania
Protein structure

ABSTRACT

ADP-ribosylation factor-like (ARL) proteins are small GTPases that undergo conformational changes upon nucleotide binding, and which regulate the affinity of ARLs for binding other proteins, lipids or membranes. There is a paucity of structural data on this family of proteins in the Kinetoplastida, despite studies implicating them in key events related to vesicular transport and regulation of microtubule-dependent processes. The crystal structure of *Leishmania major* ARL1 in complex with GDP has been determined to 2.1 Å resolution and reveals a high degree of structural conservation with human ADP-ribosylation factor 1 (ARF1). Putative *L. major* and *Trypanosoma brucei* ARF/ARL family members have been classified based on structural considerations, amino acid sequence conservation combined with functional data on Kinetoplastid and human orthologues. This classification may guide future studies designed to elucidate the function of specific family members.

ADP-ribosylation factor-like (ARL) proteins are a heterogeneous group of GTPase enzymes within the Ras (Rat sarcoma gene) family, and related to the ADP-ribosylation factor (ARF) group [1,2]. ARLs lack the defining ARF functions; they are not cholera toxin cofactors, cannot complement the lethal *arf−/arf2−* phenotype in *Saccharomyces cerevisiae*, and are incapable of activating phospholipase D. Functional data on ARF/ARL proteins are limited and the primitive eukaryotic Kinetoplastida may provide appropriate model systems to investigate this protein family since genomic data are available as are reagents and protocols for genetic studies.

ARL1 is conserved among the trypanosomatids [3]. In *Trypanosoma brucei*, ARL1 (*Tb*ARL1) is expressed only in the bloodstream form and RNAi knock down is lethal [4]. Both *Leishmania major* ARL1 (*Lm*ARL1) and *Leishmania donovani* ARL1 (*Ld*ARL1), which share 98% sequence identity, are myristoylated on Gly2 for localization to the kinetoplast and the flagella pocket *trans*-Golgi network where they likely contribute to intracellular trafficking and mutations have revealed that GTP-binding is required for correct localization [2].

Crystal structures of human and yeast ARLs are known [1] but not of any protist ARL. We determined the structure of *Lm*ARL1 to 2.1 Å resolution following protocols commonly applied in our laboratory. Crystallographic statistics are presented in Table 1 and detailed in Protein Data Bank entry 2X77. The model supports sequence-structure comparisons of *Lm*ARL1 with ARF/ARL proteins predicted from *L. major* and *T. brucei* genomic data [3,5] exploiting what is known of other ARF/ARL proteins [1,6]. A classification of these GTPases that may guide future investigations of family members in the Kinetoplastida is presented.

Analytical gel filtration indicated that dimeric and monomeric species of *Lm*ARL1 were present and the ratio varied in a time dependant manner due to disulfide bond formation. The dimer peak gave ordered crystals in space group $P2_1$ with two molecules in the asymmetric unit. Molecule A consists of which consists of residues 4–184, and molecule B, consisting of residues 3–48, 55–73 and 83–183. The molecules are linked via a Cys83–Cys83 disulfide bond similar to that observed in the *S. cerevisiae* orthologue, *Sc*ARL1 [7]. *Lm*ARL1 Cys83, conserved in ARL1 orthologues (data not shown), is placed on the loop linking β4 with β5 (Fig. 1A). Mutation of this cysteine in *Sc*ARL1 prevents dimerization yet the mutant protein can rescue cold sensitivity of *arl−* cells, implying that biological function is retained by the monomer and that the dimer is an artifact generated during purification. The r.m.s.d. of 165 Cα positions common to molecule A and B is 0.63 Å. No restraints were imposed on non-crystallographic symmetry so this indicates highly similar molecules and only A is discussed, unless otherwise stated.

Structural comparisons, using molecule A in DALI [8] identify that human ARF1 has the highest similarity to *Lm*ARL1 with an r.m.s.d. of 0.64 Å for superposition of 174 Cα atoms and indicate strong conservation of structure within the ARF/ARL protein family. The fold of *Lm*ARL1 is similar to other Ras family members, which display a β-sheet core of six strands surrounded by five α helixes;

Abbreviations: *Lm*, *Leishmania major*; *Tb*, *Trypanosoma brucei*; ARF, ADP-ribosylation factor; ARL, ADP-ribosylation factor-like; Ras, rat sarcoma gene; Sar, secretion-associated and Ras-related.

* Corresponding author. Tel.: +44 1382 385745; fax: +44 1382 385764.
E-mail address: w.n.hunter@dundee.ac.uk (W.N. Hunter).

0166-6851/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2010.08.002

Table 1
Crystallographic statistics.

Space group	$P2_1$
Unit cell dimensions (Å)	a = 35.3, b = 78.9, c = 58.0 β = 93.1°
Resolution range (Å)	78.9–2.1
Completeness (%)	99.5 (99.1)[a]
$<I/\sigma(I)>$, Mosaicity (°)	11.5 (2.4), 0.3
No. reflections measured/unique	69,671/18,516
Redundancy	3.8 (3.8)
Rmerge (%)[b]	6.8 (49)
Rwork (%)[c], Rfree (%)[d]	21.4, 28.6
Protein residues (chain A/B)	185/173
Ligands	2 GDP, 2 Mg^{2+}, 82 waters
r.m.s.d. from ideal geometry	
Bond lengths (Å)/angles (°)	0.016/1.815
Thermal parameters (B, Å^2)	
Wilson plot	30.1
Mean B all atoms	33.7
Protein atoms (chain A/B)	34.1/33.7
Mg^{2+} ions	12.9/14.2
GDP (chain A/B)	28.3/27.0
Water molecules	34.1
Ramachandran plot	
Favored/allowed/outliers (%)	95.2/3.9/0.9

[a] Values in parentheses refer to the highest resolution bin (2.21–2.10 Å).
[b] Rmerge = $\Sigma h \Sigma i ||(h,i) - <I(h)> \Sigma h \Sigma i I(h,i)$.
[c] Rwork = $\Sigma hkl ||Fo| - |Fc|| / \Sigma |Fo|$, where Fo is the observed structure-factor amplitude and the Fc is the structure-factor amplitude calculated from the model.
[d] Rfree is the same as Rwork except only calculated using a subset, 5%, of the data that are not included in any refinement calculations.

in *Lm*ARL1 a seventh β-strand is observed (Fig. 1A). ARF/ARL proteins undergo conformational changes upon nucleotide binding. They possess two segments, switch I and switch II, which direct contacts with guanine-activating proteins and guanine-exchange factors and a phosphate-binding segment (P-loop), which interacts with nucleotide phosphates (Fig. 1A). ARLs differ from other Ras family members in that they display an interswitch toggle effect whereby upon binding GTP the interswitch region, those residues positioned between switches I and II undergo a conformational change and displace the amino terminal α-helix into a position to interact with membranes [1].

In *Lm*ARL1 the switch I region consists of residues Gly43–Val54 and switch II comprises Gly72–Tyr84. Strands β3 and β4 together with the turn linking them form the interswitch section (Fig. 1A). Like other ARL:GDP complexes, α1, on the protein surface, is positioned between α2, α5 and the interswitch region [1,7]. The GDP binding site is formed between α2 and β1 and the loops linking β5–α3 and β6–α4 (Fig. 1A). Residues interacting with the ligand are shown in Fig. 1B. The P-loop, with the Walker A motif GxxxxGK(S/T), coordinates the phosphates in combination with Ser35, Asn30, and Ala31. A Mg^{2+}, coordinated by four water molecules, OG1 Thr34 and GDP β-phosphate, occupies one end of the active site. Here, Glu57 and Asp70 contribute to binding the hydrated Mg^{2+} by water-mediated interactions (Fig. 1B). Octahedral coordination of the Mg^{2+} ion stabilizes the switch I and interswitch regions in a conformation which allows α1 to adopt a retracted position.

Guanine binds at the β6–α4 region of the active site making van der Waals contacts with Lys130 on one side and Met152 from a symmetry related molecule in the crystal lattice on the other (data not shown). GDP N7 forms a hydrogen bond with a buried water molecule that also hydrogen bonds with Gly32, Asn129 (not shown) and Ser163. The amide of Ser163 and hydroxyl of Ser162 form hydrogen bonds with GDP O6 whilst Asp132 accepts hydrogen bonds donated from N1 and N2.

The amino acid conservation of the *Leishmania* ARF/ARL proteins has been mapped onto *Lm*ARL1. The most highly conserved residues constitute the core of the protein structure and the nucleotide-binding site (Supplementary data, Figs. S1 and S2).

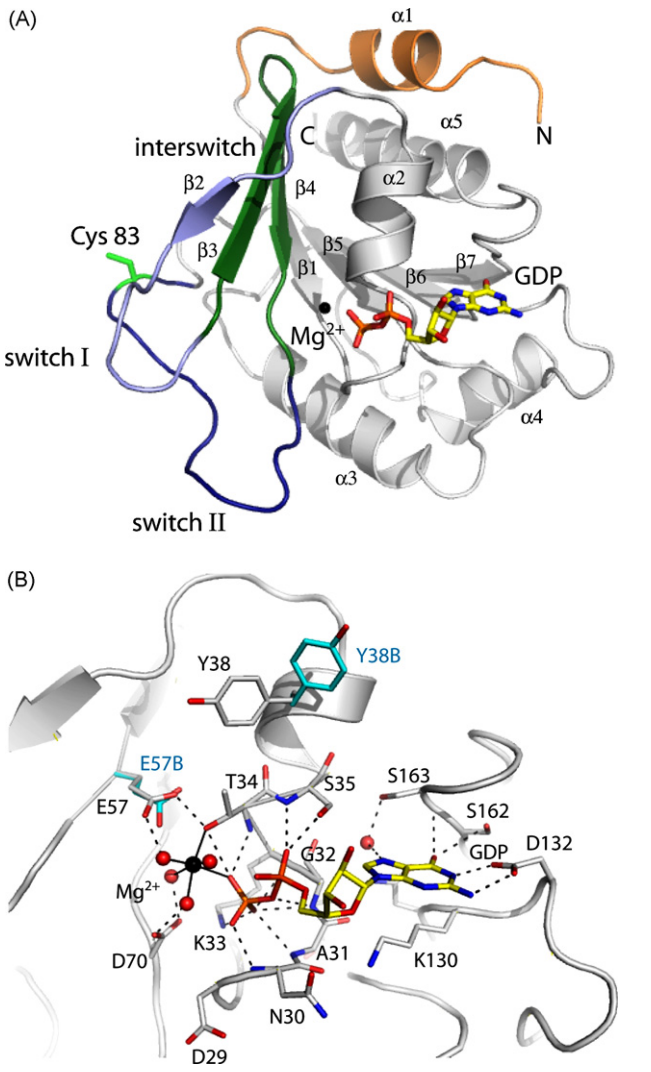

Fig. 1. (A) Ribbon diagram of the *Lm*ARL1:GDP:Mg^{2+} complex. The N-terminal α1 is orange, switch I light blue, interswitch region green, switch II dark blue and Mg^{2+} is depicted as a black sphere. GDP is shown as sticks and colored according to atom type: C yellow, N blue, O red and P orange. The side chain of Cys83, which is involved in a disulfide bond, is shown in sticks and colored green. The N and C termini and elements of secondary structure are labeled. (B) The GDP binding site. Side chains are sticks colored according to atom type: C A grey, N blue and O red. Red spheres are water molecules. Dashed lines represent potential hydrogen bonds, thin black lines coordination to Mg^{2+} ion. The positions of Glu57 and Tyr38 from molecule B, after least squares superposition, are shown with cyan C atoms for comparison. Figure prepared using PYMOL [19].

Of the fourteen residues described above with a role in binding GDP:Mg^{2+} we note that seven (Gly32, Lys33, Thr34, Asp70, Asn129, Lys130 and Asp132) are strictly conserved in eleven (of twelve) putative ARF/ARL sequences encoded in the *L. major* genome ([5], Supplementary data, Fig. S1). Ser162 is strictly conserved in seven sequences, and replaced by cysteine in another three. Asn30 and Ala31 are strictly conserved in eight and ten of the other sequences respectively. Ser35 is strictly conserved in four sequences or replaced by threonine in the remainder. Glu57, part of the interswitch region, is strictly conserved in five sequences with an aspartate in another. Residues not involved in GDP binding and located on the protein surface show poor conservation and are likely involved in complex formation with the membrane and different effectors, thus providing functional specificity to the ARF/ARL family.