Published in final edited form as:

J Struct Biol. 2012 November ; 180(2): 362–373. doi:10.1016/j.jsb.2012.07.012.

Explorations of linked editosome domains leading to the discovery of motifs defining conserved pockets in editosome OB-folds

Young-Jun Park^a and Wim G. J. Hol^{a,*}

^aBiomolecular Structure Center, Department of Biochemistry, School of Medicine, University of Washington, Seattle, WA 98195, USA

Abstract

Trypanosomatids form a group of protozoa which contain parasites of human, animals and plants. Several of these species cause major human diseases, including *Trypanosoma brucei* which is the causative agent of human African trypanosomiasis, also called sleeping sickness. These organisms have many highly unusual features including a unique U-insertion/deletion RNA editing process in the single mitochondrion. A key multi-protein complex, called the ~20S editosome, or editosome, carries out a cascade of essential RNA-modifying reactions and contains a core of 12 different proteins of which six are the interaction proteins A1 to A6. Each of these interaction proteins comprises a C-terminal OB-fold and the smallest interaction protein A6 has been shown to interact with four other editosome OB-folds. Here we report the results of a “linked OB-fold” approach to obtain a view of how multiple OB-folds might interact in the core of the editosome. Constructs of multiple variants of linked domains in 25 expression and co-expression experiments resulted in 13 soluble multi-OB-fold complexes. In several instances, these complexes were more homogeneous in size than those obtained from corresponding unlinked OB-folds. The crystal structure of A3^{OB} linked to A6 could be elucidated and confirmed the tight interaction between these two OB domains as seen also in our recent complex of A3^{OB} and A6 with nanobodies. In the current crystal structure of A3^{OB} linked to A6, hydrophobic side chains reside in well-defined pockets of neighboring OB-fold domains. When analyzing the available crystal structures of editosome OB-folds, it appears that in five instances “Pocket 1” of A1^{OB}, A3^{OB} and A6 is occupied by a hydrophobic side chain from a neighboring protein. In these three different OB-folds, Pocket 1 is formed by two conserved sequence motifs and an invariant arginine. These pockets might play a key role in the assembly or mechanism of the editosome by interacting with hydrophobic side chains from other proteins.

Keywords

RNA editing; Editosome; Interaction proteins; Protein fusions; Sleeping sickness

© 2012 Elsevier Inc. All rights reserved.

*To whom correspondence should be addressed. Telephone: +1 (206) 685 7044; Fax: +1 (206) 685 7002; wghol@u.washington.edu.

Author Contributions

Protein production, protein characterization and crystallography: Young-jun Park

Structure analysis: Young-jun Park, Wim Hol

Manuscript writing: Young-jun Park, Wim Hol

Project coordination: Wim Hol

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Trypanosoma brucei, *Trypanosoma. cruzi* and *Leishmania* species like *L. donovani*, *L. major* and others are related unicellular eukaryotic organisms which cause a number of major diseases worldwide, such as human trypanosomiasis, also called sleeping sickness, in sub-Saharan Africa; Chagas disease in Latin America; and various forms of leishmaniasis, such as kala-azar, in tropical and subtropical regions of the world (Barrett et al., 2003; Fevre et al., 2008). There is only a limited number of therapeutics available to treat these diseases. Several of these therapeutics have serious drawbacks with regards to toxicity, are difficult to administer, and the value of some is eroded by increasing drug resistance (Croft et al., 2006; Fairlamb, 2003; Hotez et al., 2007; Tarleton et al., 2007). While recently major progress has been made in understanding the mode of cell entry and resistance mechanisms in the case of compounds used for treating human African trypanosomiasis, the mode of action of several anti-trypanosomatid drugs is still unknown or uncertain (Alsford et al., 2012). In order to arrive at new therapeutics for these diseases, understanding of essential processes in trypanosomatids at the molecular level can be a starting point for the design and discovery of new leads, in particular when such processes are absent in the human host.

One such process is the remarkable U-insertion/deletion RNA editing occurring in the mitochondria of trypanosomatids (Benne et al., 1986; Simpson et al., 2004; Stuart et al., 2005). This process starts from a pre-mRNA which is expanded and edited, in a cascade of steps using information encoded on numerous small guide RNAs (gRNAs), into a mature mRNA (Blum et al., 1990; Kable et al., 1996). Many proteins and multi-protein complexes, including MRP1/2, RBP16, REAP1, TbRGG1, the guide RNA binding complex, the MRB1 complex and the ~20S editosome are involved in this unique RNA editing process (Ammerman et al., 2008; Aphasizhev and Aphasizheva, 2011; Hashimi et al., 2009; Hernandez et al., 2010; Panigrahi et al., 2008; Simpson et al., 2004; Stuart et al., 2005; Weng et al., 2008). Crystal structures of several proteins involved in U-insertion/deletion RNA editing have been elucidated (Deng et al., 2004; Deng et al., 2005; Park et al., 2012a; Park et al., 2012b; Schumacher et al., 2006; Wu et al., 2011). A large number of RNA editing enzymes are assembled, together with other proteins, in the so-called ~20S editosome (Simpson et al., 2004; Stuart et al., 2005), hereafter called the editosome. Approximately twenty proteins have been found to be incorporated into the editosome, six of these are the so-called “interaction proteins” which share an oligonucleotide-binding (OB)-fold as C-terminal domain (Panigrahi et al., 2003; Park et al., 2012a; Park et al., 2012b; Schnauffer et al., 2003; Worthey et al., 2003). Not only is this U-insertion/deletion RNA editing process in kinetoplastids unique among living organisms known, many editosome proteins from *Trypanosoma brucei* have also been shown to be essential for parasite viability (Aphasizhev et al., 2002; Brecht et al., 2005; Drozd et al., 2002; Huang et al., 2002; Kang et al., 2004; Law et al., 2007; Panigrahi et al., 2006; Rusche et al., 2001; Salavati et al., 2006; Schnauffer et al., 2001). Hence the editosome is of great interest, both for fundamental questions regarding its mode of action as well as a potential drug target (Amaro et al., 2008; Demir and Amaro, 2012; Durrant et al., 2010; Salavati et al., 2012).

Three different distinct types of editosomes have been reported (Carnes et al., 2011; Carnes et al., 2008; Panigrahi et al., 2006). These three assemblies share a common core of twelve proteins: four RNA-binding enzymes, two proteins with an RNase III domain, and the six interaction proteins just mentioned - called A1, A2, A3, A4, A5 and A6. (For editosome protein nomenclature see e.g. (Schnauffer et al., 2010; Simpson et al., 2010; Stuart et al., 2005; Wu et al., 2011)). Comprehensive biochemical and yeast-two hybrid studies have shown that the four RNA-binding enzymes and RNase-III domain proteins are engaged in associations with interaction proteins (Park et al., 2012b; Schnauffer et al., 2010).

The six interaction proteins can be subdivided into two groups. One group, comprised by A1, A2 and A3, contain two predicted Zn-binding domains N-terminal from the OB-fold. The members of the other group, consisting of A4, A5 and A6, are distinctly smaller and none of these have a predicted Zn-binding motif. Except for A5, these proteins have been well-studied with the following highlights:

- i. A1 consists of three functional domains: the L2BD, which binds the RNA editing ligase L2; the T2BD, which binds the terminal uridylyltransferase (TUTase) T2; and, the OB domain which is interacting with A6 and is also critical for enhancing the affinity of the editing ligase L2 for RNA (Park et al., 2012b; Schnauffer et al., 2010). The crystal structure of A1^{OB}, in complex with a nanobody, has recently been determined (Park et al., 2012b);
- ii. A2 also consists of three functional domains: the L1BD, which binds the ligase L1; a second domain, which probably binds the exonuclease X2; and, the OB domain which is interacting with A6 (Schnauffer et al., 2010). It has been shown that the N-terminal zinc-binding domains of A2 are implicated in the stability of the U-deletion sub-complex and in enhancement of RNA ligase enzyme activity (Gao et al., 2010; Kang et al., 2004);
- iii. A3 also consists of three domains, with the two N-terminal zinc-binding domains being essential for RNA editing, but their specific functions are largely unknown (Guo et al., 2008; Guo et al., 2010). A C-terminal OB-fold has at least two functions: interacting with A6 and with the RNase III domain protein B5 (Guo et al., 2008; Park et al., 2012a; Schnauffer et al., 2010);
- iv. A4 is the largest of the interaction proteins without Zn-binding domains. Its C-terminal OB-fold interacts with A6 (Park et al., 2012a; Schnauffer et al., 2010), while various segments of the protein chain have been shown to bind tightly to gRNA (Kala and Salavati, 2010);
- v. A5 is so far the most mysterious of the interaction proteins. It has been shown as a component of the editosome in multiple studies (Carnes et al., 2011; Panigrahi et al., 2006; Panigrahi et al., 2009) but its function is poorly understood. Binary co-expression studies in our group have failed up to now to indicate an interaction with several other editosome proteins suggesting that A5 may only form a stable complex via binding sites on multiple proteins (unpublished data);
- vi. A6 is the smallest of the interaction proteins and actually the smallest protein of the entire editosome. Its multitude of functions is therefore quite remarkable. A6, which is essentially an OB-fold with a 19-residue N-terminus extension containing a signal peptide and a 34-residue C-terminal tail, interacts with four other OB-folds: those of A1, A2, A3 and A4 (Park et al., 2012a; Park et al., 2012b; Schnauffer et al., 2010). In addition, A6 binds to the oligo (U) tail of the gRNA (Tarun et al., 2008). The interactions of the small A6 with four other OB-folds has led to the suggestion of a “five OB-fold center” in the core of the editosome, that might perform a critical scaffold function (Law et al., 2007; Panigrahi et al., 2003; Park et al., 2012a; Park et al., 2012b; Tarun et al., 2008).

Interestingly, recent electron microscopy studies have revealed that there is probably only one copy of each editosome protein in multi-protein complex (Golas et al., 2009; Li et al., 2009). That stoichiometry is reinforced by the crystal structure of the A3^{OB}-A6^{OB} complex which suggests that there is one A6 per A3 (Park et al., 2012a). Given that there is very likely only one A3 per editosome given its 42 kDa size, there is probably also one copy of the ~18 kDa A6, the smallest editosome protein, in each editosome.

It is therefore of considerable interest to increase our understanding of multi-OB-fold assemblies of the editosome, and in particular to obtain additional structural insights into the interactions of A6 with other editosome proteins. A6 by itself is in a tetramer-dimer equilibrium in solution and only crystalizes with the assistance of nanobodies. Three crystal structures obtained showed the presence of A6 dimers (Wu et al., 2011). Of great interest is of course the structure of A6 in complex with other OB-folds from the editosome. A complex of A1^{OB}•A6 in solution appeared to be a tetramer which did, however, not form crystals (Park et al., 2012b). With the help of the anti-A1 nanobody A1Nb10 as a crystallization chaperone, a complex of A1^{OB}•A6•A1Nb10 could be obtained in solution, but in the crystals grown A6 appeared to be absent, and the structure of a dimer of A1^{OB} in complex with two A1Nb10 nanobodies was obtained (Park et al., 2012b). In another study, a complex of A2^{OB}•A3^{OB}•A6 with anti-A3 nanobody A3Nb14 was obtained in solution but the structure determination revealed a remarkable tetramer revealing an A3^{OB}•A6 heterodimer with each OB-fold bound to an A3Nb14 nanobody (Park et al., 2012a). This A3^{OB}•A6•(A3Nb14)₂ structure did show an extensive interface between the two OB-folds of the A3^{OB}•A6 heterodimer. These studies indicate that the interactions between A3^{OB} and A6 are strong, but those between A1^{OB} and A6, and between A2^{OB} and A6, likely quite weak.

Linking proteins together can be a way to promote complex formation since the local concentration of one protein in the neighborhood of the other is increased by the peptide linker (Glynn et al., 2009; Janda et al., 2010; Martin et al., 2005; Zhou et al., 2011). In this paper, we therefore set out to obtain a variety of well-defined multi-OB-fold complexes by linking OB-fold domains from different editosomal interaction proteins. It appeared that in several cases linked multi-OB-fold complexes could be obtained which were more homogeneous than when the same OB-folds were co-expressed individually. Crystal growth of these complexes was not trivial, however. Yet, a structure of a linked A3^{OB}-9aa-A6 protein was obtained which shows extensive interactions between the OB-folds of A3 and A6, confirming those observed in the recent structure of the A3^{OB}•A6 heterodimer in the A3^{OB}•A6•(A3Nb14)₂ complex (Park et al., 2012a).

Interestingly, in the current linked A3^{OB}•A6 crystals, pockets on the “upper” surface of A3^{OB} and A6 are occupied by side chains provided by residues of the long N-terminal extension of A3^{OB} while the same pockets are occupied by nanobody side chains in A3^{OB}•A6•(A3Nb14)₂ and (A1^{OB}•A1Nb10)₂ complexes previously solved by our group (Park et al., 2012a; Park et al., 2012b). Two sequence motifs and a completely conserved arginine surround Pocket 1 in all three editosome OB-folds with known structure. This indicates that Pocket 1 is poised to capture side chains and hence may make critical interactions with side chains of editosome proteins in the OB-center of the core of the editosome.

2. Materials and methods

2.1. Cloning of *T. brucei* multi OB-fold domains, fusions variants of A1^{OB}, A2^{OB}, A3^{OB}, A4^{OB} and A6

A variety of OB-fold fusion genes were assembled from coding regions of *T. brucei* A1^{OB}, A2^{OB}, A3^{OB}, A4^{OB} and A6 amplified from genomic DNA by PCR using appropriate primers and verified by DNA sequencing. To create selected multi OB-fold protein expression constructs, genes were linked by partially overlapping primer-based PCR. As a part of the construction of the multi OB-fold, flexible peptide linkers were incorporated between OB-folds in a single chain to allow OB-folds to form a complex in an unstrained manner. We utilized flexible GS repeat linkers (9, 14 and 19 residues), as suggested from studies on hexameric ATPase (Martin et al., 2005), and also long natural peptide linkers (34 and 36 residues) taken from the natural C- and N-terminally extended sequences of A6^{OB}

and A3^{OB}, respectively (Fig. 1). Inspection of the crystal structures of the A3^{OB}•A6•(A3Nb14)₂ shows that the N- and C- termini of the two OB-folds are separated by ~12 Å. Therefore, linkers of 9, 14 and 19 amino acids were used to connect the OB-folds with the following sequences: the 9aa linker: KLASGAGHM; the 14aa linker: KLASGAGGSEGGHM; the 19aa linker: KLASGAGGSEGGSEGGHM. Several multi-OB-fold variants were prepared with a different order and number of OB-folds and with different lengths of the linker peptide (Fig. 1).

Each gene encoding a multi OB-fold protein was cloned into a pRSF vector for expression with an N-terminal His₆-tag followed by a tobacco etch virus (TEV) protease cleavage site. The plasmids generated were transformed into *E. coli* strain BL21 (DE3) (Novagen) for expression as hexa-histidine-tagged protein products. In addition, to produce binary multi OB-fold complexes, the genes for A6, A6 fused with A2^{OB}, and A6 fused with A4^{OB}, were cloned into pACYC without any tag. Complexes containing His₆-tagged and untagged multi-OB-folds were produced using a co-transformation and co-expression strategy as described previously (Park et al., 2012a; Park et al., 2012b; Schnauffer et al., 2010).

2.2. Protein expression and purification

Transformed *E. coli* BL21 (DE3) cells were grown at 37°C in 2 L of Luria-Bertani (LB) broth containing appropriate antibiotics. Expression of the recombinant proteins was induced by addition of isopropyl β-D-thiogalactopyranoside (to 0.5 mM final concentration) to the culture at an optical density OD₆₀₀ of ~0.6. The culture was shaken at 20°C for a further 16 h. Cells were lysed by sonication on ice. Soluble proteins were separated by centrifugation (30 min, 40,000 *g*, 4°C) from the cell pellet. The soluble fraction was applied to a gravity column packed with Ni-NTA beads (QIAGEN), washed, and eluted with 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, and 250 mM imidazole. Recombinant TEV protease was added to the collected fractions after Ni-NTA purification. The protein was further purified by a second Ni-NTA column and followed by size-exclusion chromatography using a Superdex 200 10/300 GL (GE Healthcare) filtration column in 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP and 10% glycerol buffer.

Expression and solubility of 25 different fusion proteins from *E. coli* grown at 20°C in 2 L of LB broth were examined. Insoluble and soluble fractions were separated by centrifugation at 40,000*g* for 30 min in the JA20 rotor (Beckman, USA). Total cell extracts, insoluble fractions, soluble fractions and purified fractions were analyzed on 8% to 12% denaturing SDS-PAGE. The concentration of the protein was determined using the Bio-Rad Protein Assay. The oligomeric state of recombinant proteins was characterized by size exclusion chromatography (Table 1).

2.3. Crystallization of *T. brucei* A3^{OB}-linker-A6

The purified multi-OB-fold proteins described above were screened for crystal growth. All initial screening was performed using several commercially available sparse matrix crystallization kits with a Phoenix crystallization robot (Art Robbins Instruments), where 200 nanoliter volume sitting protein drops were mixed with an equivalent volume of reservoir solution. Initially, A3^{OB}-9aa-A6, A3^{OB}-14aa-A6, A6-9aa-A1^{OB} and A6-14aa-A1^{OB} proteins gave micro crystals at 10 different conditions. The best diffracting crystals were obtained from a construct that consisted of A3^{OB}, a nine-residue linker (9aa), and A6 at room temperature by vapor diffusion experiments. Crystals of A3^{OB}-9aa-A6 protein were eventually grown in 0.4 M potassium sodium tartrate tetrahydrate, belonged to space group P2₁2₁2₁, and diffracted to 2.55 Å. Crystals were cryoprotected with 30% glycerol and flash frozen in liquid nitrogen prior to data collection.

2.4. Structure determination and refinement

Native data sets of A3^{OB}-9aa-A6 crystals were collected on beamline BL12-2 at the SSRL (Stanford Synchrotron Radiation Lightsource) and processed with HKL2000 (Otwinowski and Minor, 1997). The structure was solved by molecular replacement using PHASER (McCoy et al., 2007) with an A6^{OB} monomer and A3^{OB} monomer (PDB-ID: 3STB; (Park et al., 2012a)) search models. The initial structure model was improved and completed using the program Buccaneer (Cowtan, 2006) and COOT (Emsley and Cowtan, 2004) and refined with the program REFMAC5 (Murshudov et al., 1997). Simulated annealing and omit refinement were performed in Phenix (Adams et al., 2002). The quality of the crystal structures were analyzed using MolProbity (Chen et al., 2010). Crystallographic data collection and refinement statistics are shown in Table 2.

2.5. Figure preparation

Least squares analysis to determine the structural similarity between OB-fold structures was carried out using LSQKAB (Collaborative Computational Project, 1994). The disorder predictions of twelve proteins forming the core of the editosome are analyzed by DisMeta-server (<http://www-nmr.cabm.rutgers.edu/bioinformatics/disorder/>) that uses 10 disorder prediction programs to generate a consensus disorder prediction. Protein quaternary-structure analysis was performed with the PISA server (Krissinel and Henrick, 2007). Sequence alignment figures were made with Esript (Gouet et al., 1999). Electrostatic potential surfaces were calculated using APBS (Baker et al., 2001). All other figures of molecular structures were prepared using PyMOL (DeLano Scientific Research LLC).

2.5. Protein Data Bank accession numbers

Coordinates and structure factors for the crystal structure of A3^{OB}-9aa-A6 have been deposited with the Protein Data Bank under accession code 4DNI.

3. Results

3.1. Soluble linked OB-folds from editosome interaction proteins

Five different OB-fold domains (A1^{OB}, A2^{OB}, A3^{OB}, A4^{OB} and A6) were connected in various combinations and with different linkers to arrive at 11 different binary OB-fold combinations, 8 ternary combinations, and 6 quaternary combinations. Since A6 interacts with all other OB-folds, and since no strong interactions are known to occur mutually between other OB-folds, these 25 linked domain variants were expressed or co-expressed with each other, or with A6, in such a way that A6 was always present. A total of 25 expression and co-expression experiments were carried out, yielding 13 soluble protein complexes (Fig. 1).

A total of 11 of the 13 soluble combinations gave proteins with molecular weights less than ~136 kDa as evaluated by size exclusion chromatography (Table 1; Supplementary Figs. 1, 2 and 3). The results can be summarized as follows:

- i. Two soluble four-OB-combinations were obtained: A2^{OB}-A3^{OB}-14aa-A1^{OB} plus A6 and A4^{OB}-A3^{OB}-14aa-A1^{OB} plus A6. Although encouraging, these combinations gave unfortunately large aggregates (>160 kDa), and will not be discussed here further (Table 1; Supplementary Fig. 4).
- ii. Two three-OB-fold combinations, A1^{OB}-A3^{OB}-9aa-A6 and A2^{OB}-A3^{OB}-9aa-A6, yielded upon co-expression a complex of a size expected for a heterodimer containing four OB-folds (Table 1; Supplementary Fig. 2). This is half the size of the complexes obtained by co-expressing A1^{OB}, A3^{OB} and A6 as separate domains

(Table 1; Supplementary Fig. 5), or A2^{OB}, A3^{OB} and A6 without linkers (see Fig. 7D in (Schnauffer et al., 2010)).

- iii. Three other soluble three-OB-fold complexes were obtained by co-expressing A3^{OB}-9aa-A1^{OB} plus A6, A3^{OB}-14aa-A1^{OB} plus A6, and A3^{OB}-19aa-A1^{OB} plus A6. For the latter combination a mixture of complexes of different sizes was detected by sizing chromatography (Table 1; Supplementary Fig. 3). For A3^{OB}-9aa-A1^{OB} plus A6, and A3^{OB}-14aa-A1^{OB} plus A6, heterodimers of 60–68 kDa were obtained with one two-OB fusion forming a complex with one A6 domain as judged by size exclusion chromatography and SDS-PAGE analysis (Supplementary Fig. 3). When co-expressing unlinked A3^{OB}, A1^{OB} and A6 a complex of ~136 kDa was the result (Table 1; Supplementary Fig. 5).
- iv. three two-OB-fold fusions containing A6 and A1^{OB} connected by linkers of different lengths, obtained by adding 9, 14 or 19 amino acids to the C-terminal tail of A6, gave complexes consisting of (A6-9aa-A1^{OB})₂, (A6-14aa-A1^{OB})₂ and (A6-19aa-A1^{OB})₂ homodimers as supported by sizing and SDS-PAGE (Table 1; Supplementary Fig. 1 (c) and (d)). This was a more homogeneous preparation than obtained by co-expressing unlinked A6 and A1^{OB}, where a mixture of A6 homotetramers and (A1^{OB}•A6)₂ heterotetramers was formed (Table 1; Supplementary Fig. 5 (a–c)).
- v. three two-OB-fold fusions containing A3^{OB} and A6 connected by 9, 14 or 19 amino acid linkers gave complexes consisting of (A3^{OB}-9aa-A6)₂, (A3^{OB}-14aa-A6)₂ and (A3^{OB}-19aa-A6)₂ dimers as judged by sizing and SDS-PAGE (Table 1; Supplementary Fig. 1 (a) and (b)). This was smaller than the complexes obtained by co-expressing unlinked A3^{OB} plus A6, which yielded (A3^{OB}•A6)₃ hetero-hexamers (see Figure 4D in reference (Schnauffer et al., 2010)).

For all the soluble proteins obtained crystallization experiments were set up with multiple sparse matrices. Of these, A3^{OB}-9aa-A6, A3^{OB}-14aa-A6, A6-9aa-A1^{OB} and A6-14aa-A1^{OB} proteins gave hits at several different conditions. Optimization yielded diffraction quality crystals of A3^{OB}-9aa-A6.

3.2. The crystal structure of A3^{OB}-9aa-A6

The crystal structure of the A3^{OB}-9aa-A6 complex was solved to a resolution of 2.55 Å and refined to an R-factor of 21.4% and an R-free of 24.7%. Electron density maps allowed the fitting of all A3 (260–393), 9aa linker (1–9) and A6 (20–133) residues, with the exception of the C-terminal tail residues (134–164) of A6 (Table 2). The crystal structure of A3^{OB}-9aa-A6 revealed that the OB-folds of A3 and A6 form a tight complex burying 2730 Å² solvent accessible surface in their interface (Fig. 2). An unexpected feature of the structure was that the 23 residues preceding the OB-fold of A3 extend far from the globular A3^{OB} domain as a long, extended chain with well-defined electron density for 16 out of the 23 residues (Fig. 2(a)). Moreover, side chains of three residues of this N-terminal extension occupy pockets of two neighboring A3^{OB}-9aa-A6 fusion proteins in the crystal (Supplementary Fig. 6). This will be further discussed below.

There is one A3^{OB}-9aa-A6 protein per asymmetric unit, which has an extensive contact via the extended N-terminal residues of A3 with an A3^{OB} domain of a neighboring protein in the crystal lattice, burying ~3400 Å² solvent accessible surface (Supplementary Fig. 6). The nine residue linker connecting the two OB-folds is well defined in the electron density map, which is probably due to the fact that it is engaged in the contacts between the same crystal mates, burying 580 Å² of its accessible surface (Supplementary Fig. 6). It is possible that these multiple interactions between A3^{OB}-9aa-A6 proteins in the crystal explain the

formation of A3^{OB}-9aa-A6 dimers observed in size exclusion experiments (Supplementary Fig. 1).

3.3. Structural comparison with the A3^{OB}•A6 heterodimer

The crystal structure of a complex of a non-linked A3^{OB}•A6 heterodimer with two anti-A3 nanobodies, A³Nb14, has recently been determined where, remarkably, the anti-A3 nanobody formed a complex with A3^{OB}, as expected, but also with A6, which was entirely unexpected (Park et al., 2012a). Another special feature of this A3^{OB}•A6•(A³Nb14)₂ complex was that A3^{OB} and A6 form a heterodimer of different OB-folds, the first OB-fold heterodimer reported so far. The current structure of A3^{OB}-9aa-A6 offers the opportunity to see if the A3^{OB}•A6 heterodimer is formed again, and if so, how close the two heterodimers are structurally.

It appears that in the new structure of the linked A3^{OB}-9aa-A6 protein chain, the two OB-folds from different editosome interaction proteins contact each other in a similar manner as observed in the previous A3^{OB}•A6•(A³Nb14)₂ structure (Fig. 3). The pair of OB-folds in the two structures, as well as the individual OB-folds, superimpose with an root-mean-square deviation of 0.93 to 1.13 Å, i.e. the structures are highly similar. Clearly the formation of the A3^{OB}•A6 heterodimer is not dependent on the presence of nanobody A³Nb14. The largest structural differences are observed in the β4-β5 regions which extend “upwards” from the globular body of the OB-fold, with the tip of the loop connecting β4-β5 of A3 and A6 differing in the two structures by 5.9 and 9.4 Å, respectively (Fig. 3). The interface between the two OB-folds in the new A3^{OB}-9aa-A6 structure buries 2730 Å² solvent accessible surface, which is ~14 % less than in the A3^{OB}•A6•(A³Nb14)₂ complex. A major difference between the two OB-fold heterodimers is that in the new structure the N-terminal 23 residues from A3^{OB}, which were not observed in the A3^{OB}•A6•(A³Nb14)₂ complex, are full extended and wind as “guirlandes” through the crystal, interacting with two different A3^{OB}-9aa-A6 molecules (Fig. 3; Supplementary Fig. 6(a)).

3.4. Pockets 1 and 2 in three editosome OB-folds in four crystal structures

The extended N-terminus of A3^{OB}-9aa-A6 positions three side chains in specific pockets of neighboring molecules (Fig. 4; Supplementary Fig. 6). The side chain of the residue in the fusion protein equivalent to L270 of A3 (called hereafter L270^{A3}; for ease of reference across papers on editosome proteins, the residues in the fusion protein are numbered in the text according to their sequence in wild type A3 and in A6, with a superscript indicating the protein of origin) fits into a pocket which is formed by five residues from A3^{OB} of another molecule in the crystal, while the side chain of Q267^{A3} resides in a neighboring pocket formed by five residues of the same A3 OB-fold. Two A3 residues contribute to both pockets (Fig. 4(a) and (b)). These two pockets are called Pocket 1 and Pocket 2, hereafter. Another residue of the extended N-terminus of A3^{OB}-9aa-A6, L262^{A3}, occupies a pocket in the A6 OB-fold of yet another A3^{OB}-9aa-A6 molecule in the crystal. This A6 pocket is structurally equivalent to Pocket 1 in A3 mentioned above (Fig. 4(a), (b), (f) and (g)). It appears that the three residues (L262, Q267 and L270) of the N-terminal extension of A3 which reside in pockets are completely conserved among trypanosomatids (Supplementary Fig. 7).

It is of interest to analyze these two pockets in other recent crystal structures containing OB-folds from editosome interaction proteins: the A3^{OB}•A6•(A³Nb14)₂ complex (Park et al., 2012a) and the (A1^{OB}•A1Nb10)₂ complex (Park et al., 2012b). When combining this also with sequence comparisons (Supplementary Fig. 8–10) then it appears then that:

- i. in the A3^{OB}•A6•(A³Nb14)₂ complex, the side chain of F102 of A³Nb14 occupies Pocket 1 of A3^{OB} (Fig. 4(c));

- ii. in the $(A1^{OB} \bullet A1^{Nb10})_2$ complex, the side chain of V106 of $A1^{Nb10}$ occupies Pocket 1 of $A1^{OB}$ (Fig. 4(d); Supplementary Fig. 11);
- iii. in the $A3^{OB} \bullet A6 \bullet (A3^{Nb14})_2$ complex, the side chain of F102 of $A3^{Nb14}$ occupies Pocket 1 of $A6$ (Fig. 4(e));
- iv. in the $(A1^{OB} \bullet A1^{Nb10})_2$ complex, the side chain of F101 of $A1^{Nb10}$ occupies Pocket 2 of $A1^{OB}$ (Fig. 4(d); Supplementary Fig. 11).

The shapes of Pockets 1 and 2 in $A1^{OB}$, $A3^{OB}$ and $A6$ are in all cases quite similar (Fig. 4(a), (d) and (g)). The residues forming Pocket 2 are not well conserved (Supplementary Figs. 8). In contrast, the residues lining Pocket 1 in these three OB-folds include (i) the conserved Q/E-x-G-F/Y-L/V Motif I, (ii) the conserved V-x-Q Motif II, where “x” can be any residue, and (iii) an invariant arginine residue (R703 in *T. brucei* A1, R336 in *T. brucei* A3, and R73 in *T. brucei* A6) (red spheres in Supplementary Figs. 7, 8, 9 and 10). The invariant arginine makes a hydrogen bond with the conserved Q of Motif II (Fig. 4(b) and (f); Supplementary Fig. 11). The backbone of residues Q/E-x-G-F/Y of Motif I forms one wall of Pocket 1, with the side chains of the “x” pointing towards the solvent, and the side chain of the “F/Y” pointing away from the pocket. The final “L/V” of Motif I and the “V” of Motif II contribute to the hydrophobic character of the pocket, together with the aliphatic side chain atoms of the “Q/E” of Motif I and the “Q” of Motif II. The conservation of the Motifs and the invariant arginine is the more remarkable since the sequence identities of the three editosome OB-folds with known structure are quite low: between $A1^{OB}$ and $A3^{OB}$ the sequence identity is 26 %, and between $A1^{OB}$ and $A6^{OB}$ 28 %; only $A3^{OB}$ and $A6^{OB}$ share a higher sequence identity of 40 %.

4. Discussion

4.1. Oligomeric states of editosome OB-folds and linking strategy

Individual OB-folds of the interaction proteins A1 to A6 appear to be rather sticky given their tendency to aggregate into various multimeric complexes. For instance, A6 forms tetramers in solution, while A6 crystallizes as an A6 dimer with the assistance of nanobodies (Wu et al., 2011). Although the formation of an A6 tetramer is not likely to occur in the editosome, the C-terminal tail of A6 appears to be involved in tetramer formation (Wu et al., 2011). Furthermore, each OB-fold domain of A2, A3 and A4 when expressed individually forms either high molecular mass oligomers or aggregates (data not shown). Less prone to oligomerization is $A1^{OB}$ which is in monomer-dimer equilibrium in solution (Park et al., 2012b). The only exception to the multimerization tendency is $A5^{OB}$ which by itself forms a monomer, while we have also not found any strong interaction between $A5^{OB}$ and any other OB-fold domain so far (unpublished data). Unsurprisingly, mixing heterogeneous preparations of different editosome OB-folds did not yield well-defined homogeneous multi OB-fold complexes *in vitro*. But co-expression has resulted in relatively well defined multimers of pairs ($A1^{OB} \bullet A6$, $A2^{OB} \bullet A6$, $A3^{OB} \bullet A6$ and $A4^{OB} \bullet A6$) and triplets ($A2^{OB} \bullet A3^{OB} \bullet A6$, $A1^{OB} \bullet A3^{OB} \bullet A6$ and $A4^{OB} \bullet A3^{OB} \bullet A6$) of OB-folds, which were obtained using polycistronic vectors (Table 1; (Park et al., 2012a; Park et al., 2012b; Schnauffer et al., 2010)). Unfortunately, major efforts to crystallize these single, double and triple OB-fold complexes did not meet with success.

This problem could be solved to some extent by two strategies. First, the use of nanobodies as crystallization chaperones. With the help of nanobodies, crystals and structures of $A6 \bullet A6^{Nb5}$, $A6 \bullet A6^{Nb15}$, $A6 \bullet A6^{Ab2}$, $A1^{OB} \bullet A1^{Nb10}$ and $A3^{OB} \bullet A6 \bullet (A3^{Nb14})_2$ could be obtained in seven different crystal forms in total (Park et al., 2012a; Park et al., 2012b; Wu et al., 2011). The nanobodies were not only used as crystallization chaperones but also for purification (Park et al., 2012b). Second, a linking strategy as described in the current study.

This approach has as advantage a perfect stoichiometric ratio for the domains incorporated into a single polypeptide chain. Also, as mentioned in the introduction, linking domains can promote complex formation by increasing the local domain concentration. On the other hand, linkers might be too short to allow naturally occurring domain interfaces to form, or linkers might interact with domain interface surfaces and thereby prevent multimerization.

Several OB-fold complexes obtained with the new strategy appeared to have improved properties compared with corresponding experiments where separate domains were expressed (Table 1). For instance, the three A1^{OB}-linker-A6 variants tested in the current study resulted in quite homogeneous (A1-linker-A6)₂ complexes (Supplementary Fig. 1(b)), while, in contrast, when A1^{OB} and A6 were co-expressed as individual domains a mixture of (A6)₄ homotetramers and (A1^{OB}•A6)₂ heterotetramers was obtained (Supplementary Fig. 5 (a–c)). Also, expression of A3^{OB}-linker-A6 yielded two well-defined (A3^{OB}-linker-A6)₂ complexes for three different linker lengths (Table 1; Supplementary Fig. 1(a)) while co-expression of unlinked A3^{OB} and A6 gave a more complicated chromatogram in size exclusion experiments indicating a possible (A3^{OB}•A6)₃ complex (Park et al., 2012a). Also other linked multi OB-fold proteins yielded more homogeneous preparations than were obtained by co-expressing unlinked separate domains (Table 1). The crystal growth success rate of even the homogeneous linked OB-fold proteins was, however, quite low (Table 1), but did confirm the unexpected and important discovery of the first OB-fold heterodimer.

4.2. The A3^{OB}•A6 heterodimer

An A3^{OB}•A6 complex could previously be crystallized but required the assistance of an anti-A3 nanobody as crystallization chaperone (Park et al., 2012a). Unexpectedly, the resultant structure revealed that A3^{OB} and A6 form a heterodimer and that both A3^{OB} and A6 interacted with an anti A3 nanobody in the A3^{OB}•A6•(A³Nb14)₂ structure, another unusual structural feature of this heterotetramer (Park et al., 2012a). In the current study, we were able to crystallize A3^{OB}-linker-A6 and A6-linker-A1^{OB} with several different linker lengths in several different conditions in the absence of nanobodies (Table 2), and the crystal structure of the A3^{OB}-9aa-A6 complex was solved to a resolution of 2.55 Å. The A3^{OB}-9aa-A6 chain has excellent electron density for all structural elements, except for 31 residues at the A6 C-terminus. Overall, the A3^{OB}-9aa-A6 “OB-fold dimer” is very similar to the true A3^{OB}•A6 heterodimer observed in the A3^{OB}•A6•(A³Nb14)₂ heterotetramer (Park et al., 2012a), with a root-mean-square deviation of less than 1 Å. Also in the new, linked, A3^{OB}-9aa-A6 crystal structure OB-folds of A3 and A6 form a tight complex (Fig. 2). Not less than 38 residues of A3, and 36 residues of A6, are involved in the A3^{OB} and A6 contacts, resulting in the burial of 2730 Å² of accessible surface area at the interface.

Pairwise comparisons reveal that the A3^{OB} and A6 domains in the two complexes A3^{OB}-9aa-A6 and A3^{OB}•A6•(A³Nb14)₂ superpose very well, with only minor differences in the loops located in the β4-β5 hairpin (Fig. 3). It appears that the residues forming this hairpin have high mobility especially at the tips, i.e. D369^{A3}-R373^{A3} and Q104^{A6}-H112^{A6}, but there is no ambiguity about the tracing of these loops. This finding is consistent with the lack of electron density observed for the β4-β5 loop regions in the A6•A⁶Nb5 (PDB-ID: 3K80) and A6•A⁶Nb15 (PDB-ID: 3K7U) structures (Wu et al., 2011). Taking all this together, this demonstrates that A3^{OB} and A6 heterodimer formation is not dependent on nanobody binding.

4.3. The N-terminal extension of A3^{OB}

In the present A3^{OB}-9aa-A6 structure, we were able to add 23 extra N-terminal residues in A3, spanning T260 to V282, to A3^{OB} from our previous A3^{OB}•A6•(A³Nb14)₂ structure (Fig. 2; Supplementary Fig. 7). Interestingly, these N-terminal residues form long

extensions, with well-defined electron density, engaged in crystal contacts with A3^{OB} and A6 domains of symmetry mates (Supplementary Fig. 6). Specifically, the side chain of residue L270^{A3} resides in Pocket 1 of A3^{OB}, the side chain of Q267^{A3} in near-by Pocket 2 of A3^{OB}, and the side chain of L262^{A3} in Pocket 1 of A6 (Fig. 4). Some of the interactions between A3 extension residues and these pockets might be the driving force for the formation of A3^{OB}-9aa-A6 dimers in solution (Supplementary Fig. 1). The three residues of the A3 extension that are buried in the pockets of neighboring A3^{OB} and A6 domains, i.e. L262^{A3}, Q267^{A3} and L270^{A3}, are conserved in the orthologous proteins from other Kinetoplastida species (Supplementary Fig. 7). At present there is no evidence that these three residues N-terminal to A3^{OB} are playing a crucial role in bringing proteins together in an assembled editosome, although this might well be the case.

4.4. Conserved pockets in three editosome OB-folds

The pockets of the two OB-folds occupied by side chains in our newly solved structure suggested to analyze our recent crystal structures of *T. brucei* A6, A3^{OB} and A1^{OB} in complex with nanobodies (Park et al., 2012a; Park et al., 2012b; Wu et al., 2011). It appeared that Pockets 1 and 2 are occupied by nanobody side chains in the A6•A3^{Nb14}, A3^{OB}•A3^{Nb14} and A1^{OB}•A1^{Nb10} structures. In the four available structures containing editosome OB-folds, Pocket 1 is occupied five times in three different OB-folds by hydrophobic side chains (Fig. 4, Supplementary Fig. 11):

- i. by L270^{A3} in A3^{OB} in the current structure;
- ii. by A3^{Nb14}-F102 in A3^{OB} of the A3^{OB}•A6•(A3^{Nb14})₂ structure;
- iii. by A3^{Nb14}-F102 in A6 of the A3^{OB}•A6•(A3^{Nb14})₂ structure;
- iv. by L262^{A3} in A6 in the current structure; and,
- v. by A1^{Nb10}-V106 in A1^{OB} of the A1^{OB}•A1^{Nb10} structure.

In these structures, Pocket 2 is two times occupied by a side chain but since the residues lining this pocket are not well conserved (Fig. 4, Supplementary Fig. 11) we will not discuss Pocket 2 here further.

Comparison of the amino acid sequences defining Pocket 1 in A1^{OB}, A3^{OB} and A6 resulted in the discovery of the conserved sequence motifs I and II, and an invariant arginine (Supplementary Figs. 7, 8, 9, and 10). The observation that Pockets 1 of A1^{OB}, A3^{OB} and A6 are in three crystal structures occupied in five instances by a hydrophobic side chain, combined with the sequence conservation of the residues defining this pocket, suggests that Pockets 1 of these three OB-folds might be responsible for interactions with hydrophobic side chains of other editosome proteins during biogenesis of the editosome and/or in the assembled editosome. Interestingly, this combination of motifs I and II plus the invariant arginine is also present in the OB-fold of A5, possibly also in A4 (where the invariant arginine might be missing), but not in A2, nor in other non-editosomal OB-fold proteins (Theobald et al., 2003). Whether or not these Pockets 1 in four or five OB folds are indeed binding hydrophobic side chains from other editosome proteins in physiological conditions, and if so, which side chains of which proteins then occupy Pocket 1 in each of these OB-folds requires obviously further investigations.

4.5. Interactions of OB-folds and implication for editosome assembly

The current studies aimed at obtaining more insight into the architecture of the proposed OB-fold center in the core of the editosome. After a substantial effort (Table 1), it appeared difficult to obtain soluble homogeneous multi-OB-fold complexes and even more challenging to grow crystals containing more than two editosome OB-folds. The current

study has provided considerable support to the idea that the OB-folds of A3 and A6 form a tight complex at the heart of the editosome, but making further progress appears to be very difficult. Other approaches in our laboratory, without linking OB-folds, have indicated that the mutual affinities between several OB-folds are probably quite moderate. For instance, A6^{OB} displays likely a weak interaction to A1^{OB} since crystallizing the A1^{OB}•A6^{OB}•A1Nb10 complex resulted in the structure of A1^{OB}•A1Nb10 with a loss of A6 during crystal growth in mild conditions (Park et al., 2012b). Also, A2^{OB} has a weak interaction with A3^{OB}•A6^{OB} since crystallizing the A2^{OB}•A3^{OB}•A6•(A³Nb14)₂ complex resulted in the surprising and informative A3^{OB}•A6•(A³Nb14)₂ crystal structure with, however, the loss of A2^{OB} during crystal growth (Park et al., 2012a).

We had anticipated that combining and linking additional OB-folds would result in several well-defined, stable and crystallizable multi-OB-fold complexes but the approach followed in the current paper (Table 1), where up to five editosome OB-folds were combined, did not succeed with respect to the crystallization step, except for a complex of two OB-folds. It might be that the linkers were too long or too short. Another possibility is that all six OB-folds are required to form a compact complex, but A5^{OB} did not show any interactions with any other OB-fold (data not shown). Yet another option is that additional proteins from the editosome are needed to obtain a sufficiently stable complex with multiple OB-fold containing proteins. A disorder prediction analysis by DisMeta (<http://www-nmr.cabm.rutgers.edu/bioinformatics/disorder>) of all twelve proteins forming the core of the editosome indicates that several of these proteins may have extensive disordered regions, in particular the interaction protein A3, but also A1 and A2 (Supplementary Fig. 12). Some of these regions might be required for providing the editosome with sufficient flexibility to allow domains to move during the cascade of catalytic reactions and/or to allow translocations of RNA strands during U-insertion and deletion. Another possibility is that some of these flexible stretches of polypeptide chain wrap around several proteins in the editosome and have an important stabilizing function. A fascinating example of such an architecture has been shown to occur in the structure of U1 snRNP in association with U1 snRNA where the N-terminal polypeptide of U1-70K extends over 180 Å, and wraps around eight subunit proteins (Pomeranz Krummel et al., 2009). While we have established in the current study that A3 and A6 interact tightly with each other and form an unexpected heterodimer of OB-folds, major questions remain to be answered and further investigations are required to understand the architecture of editosome more completely.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the staff of BL12-2 beam line at SSRL for invaluable assistance with data collection. We also thank Tanya Budiarto, Sarah Wolf, Connie Yiching Lu and Stewart Turley for assistance with protein purification, linkers, and technical support; Jonathan Kay for maintaining the computing environment of the Biomolecular Structure Center; Els Pardon and Jan Steyaert for earlier studies of editosome OB-folds in complex with nanobodies; and Meiting Wu and Jungpeng Deng for contributions to earlier stages of the project. This study was funded by National Institute of Health grants (RO1 GM077418 and RO1 GM077418-04S1) to WGJH.

References

- Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr*. 2002; 58:1948–1954. [PubMed: 12393927]
- Alsford S, Eckert S, Baker N, Glover L, Sanchez-Flores A, et al. High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature*. 2012; 482:232–236. [PubMed: 22278056]

- Amaro RE, Schnauffer A, Interthal H, Hol WGJ, Stuart KD, et al. Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A*. 2008; 105:17278–17283. [PubMed: 18981420]
- Ammerman ML, Fisk JC, Read LK. gRNA/pre-mRNA annealing and RNA chaperone activities of RBP16. *RNA*. 2008; 14:1069–1080. [PubMed: 18441045]
- Aphasizhev R, Aphasizheva I. Mitochondrial RNA processing in trypanosomes. *Res Microbiol*. 2011; 162:655–663. [PubMed: 21596134]
- Aphasizhev R, Sbicego S, Peris M, Jang S, Aphasizheva I, et al. Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell*. 2002; 108:637–648. [PubMed: 11893335]
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*. 2001; 98:10037–10041. [PubMed: 11517324]
- Barrett MP, Burchmore RJ, Stich A, Lazzari JO, Frasch AC, et al. The trypanosomiasis. *Lancet*. 2003; 362:1469–1480. [PubMed: 14602444]
- Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, et al. Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*. 1986; 46:819–826. [PubMed: 3019552]
- Blum B, Bakalara N, Simpson L. A model for RNA editing in kinetoplastid mitochondria: “guide” RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell*. 1990; 60:189–198. [PubMed: 1688737]
- Brecht M, Niemann M, Schluter E, Muller UF, Stuart K, et al. TbMP42, a protein component of the RNA editing complex in African trypanosomes, has endo-exoribonuclease activity. *Mol Cell*. 2005; 17:621–630. [PubMed: 15749013]
- Carnes J, Soares CZ, Wickham C, Stuart K. Endonuclease associations with three distinct editosomes in *Trypanosoma brucei*. *J Biol Chem*. 2011; 286:19320–19330. [PubMed: 21474442]
- Carnes J, Trotter JR, Peltan A, Fleck M, Stuart K. RNA editing in *Trypanosoma brucei* requires three different editosomes. *Mol Cell Biol*. 2008; 28:122–130. [PubMed: 17954557]
- Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010; 66:12–21. [PubMed: 20057044]
- Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Cryst*. 1994; D50:760–763.
- Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr*. 2006; 62:1002–1011. [PubMed: 16929101]
- Croft SL, Sundar S, Fairlamb AH. Drug resistance in leishmaniasis. *Clin Microbiol Rev*. 2006; 19:111–126. [PubMed: 16418526]
- Demir O, Amaro RE. Elements of nucleotide specificity in the *Trypanosoma brucei* mitochondrial RNA editing enzyme RET2. *J Chem Inf Model*. 2012; 52:1308–1318. [PubMed: 22512810]
- Deng J, Schnauffer A, Salavati R, Stuart KD, Hol WGJ. High resolution crystal structure of a key editosome enzyme from *Trypanosoma brucei*: RNA editing ligase 1. *J Mol Biol*. 2004; 343:601–613. [PubMed: 15465048]
- Deng J, Ernst NL, Turley S, Stuart KD, Hol WGJ. Structural basis for UTP specificity of RNA editing TUTases from *Trypanosoma brucei*. *EMBO J*. 2005; 24:4007–4017. [PubMed: 16281058]
- Drozd M, Palazzo SS, Salavati R, O'Rear J, Clayton C, et al. TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J*. 2002; 21:1791–1799. [PubMed: 11927563]
- Durrant JD, Hall L, Swift RV, Landon M, Schnauffer A, et al. Novel naphthalene-based inhibitors of *Trypanosoma brucei* RNA editing ligase 1. *PLoS Negl Trop Dis*. 2010; 4:e803. [PubMed: 20808768]
- Emsley P, Cowtan K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D - Biological Crystallography*. 2004; 60:2126–2132.
- Fairlamb AH. Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol*. 2003; 19:488–494. [PubMed: 14580959]

- Fevre EM, Wissmann BV, Welburn SC, Lutumba P. The burden of human African trypanosomiasis. *PLoS Negl Trop Dis*. 2008; 2:e333. [PubMed: 19104653]
- Gao G, Rogers K, Li F, Guo Q, Osato D, et al. Uridine insertion/deletion RNA editing in Trypanosomatids: specific stimulation in vitro of *Leishmania tarentolae* REL1 RNA ligase activity by the MP63 zinc finger protein. *Protist*. 2010; 161:489–496. [PubMed: 20138580]
- Glynn SE, Martin A, Nager AR, Baker TA, Sauer RT. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell*. 2009; 139:744–756. [PubMed: 19914167]
- Golas MM, Bohm C, Sander B, Effenberger K, Brecht M, et al. Snapshots of the RNA editing machine in trypanosomes captured at different assembly stages in vivo. *EMBO J*. 2009; 28:766–778. [PubMed: 19197238]
- Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*. 1999; 15:305–308. [PubMed: 10320398]
- Guo X, Ernst NL, Stuart KD. The KREPA3 zinc finger motifs and OB-fold domain are essential for RNA editing and survival of *Trypanosoma brucei*. *Mol Cell Biol*. 2008; 28:6939–6953. [PubMed: 18794366]
- Guo X, Ernst NL, Carnes J, Stuart KD. The zinc-fingers of KREPA3 are essential for the complete editing of mitochondrial mRNAs in *Trypanosoma brucei*. *PLoS One*. 2010; 5:e8913. [PubMed: 20111718]
- Hashimi H, Cicova Z, Novotna L, Wen YZ, Lukes J. Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase. *RNA*. 2009; 15:588–599. [PubMed: 19228586]
- Hernandez A, Madina BR, Ro K, Wohlschlegel JA, Willard B, et al. REH2 RNA helicase in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding. *J Biol Chem*. 2010; 285:1220–1228. [PubMed: 19850921]
- Hotez PJ, Molyneux DH, Fenwick A, Kumaresan J, Sachs SE, et al. Control of neglected tropical diseases. *N Engl J Med*. 2007; 357:1018–1027. [PubMed: 17804846]
- Huang CE, O'Hearn SF, Sollner-Webb B. Assembly and function of the RNA editing complex in *Trypanosoma brucei* requires band III protein. *Mol Cell Biol*. 2002; 22:3194–3203. [PubMed: 11940676]
- Janda CY, Li J, Oubridge C, Hernandez H, Robinson CV, et al. Recognition of a signal peptide by the signal recognition particle. *Nature*. 2010; 465:507–510. [PubMed: 20364120]
- Kable ML, Seiwert SD, Heidmann S, Stuart K. RNA editing: a mechanism for gRNA-specified uridylation insertion into precursor mRNA. *Science*. 1996; 273:1189–1195. [PubMed: 8703045]
- Kala S, Salavati R. OB-fold domain of KREPA4 mediates high-affinity interaction with guide RNA and possesses annealing activity. *RNA*. 2010; 16:1951–1967. [PubMed: 20713467]
- Kang X, Falick AM, Nelson RE, Gao G, Rogers K, et al. Disruption of the zinc finger motifs in the *Leishmania tarentolae* LC-4 (=TbMP63) L-complex editing protein affects the stability of the L-complex. *J Biol Chem*. 2004; 279:3893–3899. [PubMed: 14604987]
- Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007; 372:774–797. [PubMed: 17681537]
- Law JA, O'Hearn S, Sollner-Webb B. In *Trypanosoma brucei* RNA editing, TbMP18 (band VII) is critical for editosome integrity and for both insertional and deletional cleavages. *Mol Cell Biol*. 2007; 27:777–787. [PubMed: 17101787]
- Li F, Ge P, Hui WH, Atanasov I, Rogers K, et al. Structure of the core editing complex (L-complex) involved in uridine insertion/deletion RNA editing in trypanosomatid mitochondria. *Proc Natl Acad Sci U S A*. 2009; 106:12306–12310. [PubMed: 19590014]
- Martin A, Baker TA, Sauer RT. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature*. 2005; 437:1115–1120. [PubMed: 16237435]
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658–674. [PubMed: 19461840]
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 1997; 53:240–255. [PubMed: 15299926]

- Otwinowski, Z.; Minor, W. DENZO. Processing of X-ray diffraction data collected in oscillation mode. In: Sweet, CWJCaRM., editor. *Methods in Enzymology*. Academic Press; New York: 1997. p. 307-326.
- Panigrahi AK, Ernst NL, Domingo GJ, Fleck M, Salavati R, et al. Compositionally and functionally distinct editosomes in *Trypanosoma brucei*. *RNA*. 2006; 12:1038–1049. [PubMed: 16611942]
- Panigrahi AK, Schnauffer A, Ernst NL, Wang B, Carmean N, et al. Identification of novel components of *Trypanosoma brucei* editosomes. *RNA*. 2003; 9:484–492. [PubMed: 12649499]
- Panigrahi AK, Zikova A, Dalley RA, Acestor N, Ogata Y, et al. Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex. *Mol Cell Proteomics*. 2008; 7:534–545. [PubMed: 18073385]
- Panigrahi AK, Ogata Y, Zikova A, Anupama A, Dalley RA, et al. A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics*. 2009; 9:434–450. [PubMed: 19105172]
- Park YJ, Pardon E, Wu M, Steyaert J, Hol WGJ. Crystal structure of a heterodimer of editosome interaction proteins in complex with two copies of a cross-reacting nanobody. *Nucleic Acids Res*. 2012a; 40:1828–1840. [PubMed: 22039098]
- Park YJ, Budiarto T, Wu M, Pardon E, Steyaert J, et al. The structure of the C-terminal domain of the largest editosome interaction protein and its role in promoting RNA binding by RNA-editing ligase L2. *Nucleic Acids Res*. 2012b in press. 10.1093/nar/gks1369
- Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature*. 2009; 458:475–480. [PubMed: 19325628]
- Rusche LN, Huang CE, Pillar KJ, Hemann M, Wirtz E, et al. The two RNA ligases of the *Trypanosoma brucei* RNA editing complex: cloning the essential band IV gene and identifying the band V gene. *Mol Cell Biol*. 2001; 21:979–989. [PubMed: 11158286]
- Salavati R, Moshiri H, Kala S, Shateri Najafabadi H. Inhibitors of RNA editing as potential chemotherapeutics against trypanosomatid pathogens. *International Journal for Parasitology: Drugs and Drug Resistance*. 2012; 2:36–46.
- Salavati R, Ernst NL, O'Rear J, Gilliam T, Tarun S Jr, et al. KREPA4, an RNA binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA*. 2006; 12:819–831. [PubMed: 16601201]
- Schnauffer A, Ernst NL, Palazzo SS, O'Rear J, Salavati R, et al. Separate insertion and deletion subcomplexes of the *Trypanosoma brucei* RNA editing complex. *Mol Cell*. 2003; 12:307–319. [PubMed: 14536071]
- Schnauffer A, Panigrahi AK, Panicucci B, Igo RP Jr, Wirtz E, et al. An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science*. 2001; 291:2159–2162. [PubMed: 11251122]
- Schnauffer A, Wu M, Park YJ, Nakai T, Deng J, et al. A protein-protein interaction map of trypanosome ~20S editosomes. *J Biol Chem*. 2010; 285:5282–5295. [PubMed: 20018860]
- Schumacher MA, Karamooz E, Zikova A, Trantirek L, Lukes J. Crystal structures of *T. brucei* MRP1/ MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell*. 2006; 126:701–711. [PubMed: 16923390]
- Simpson L, Aphasizhev R, Gao G, Kang X. Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U-insertion/deletion RNA editing. *RNA*. 2004; 10:159–170. [PubMed: 14730014]
- Simpson L, Aphasizhev R, Lukes J, Cruz-Reyes J. Guide to the nomenclature of kinetoplastid RNA editing: a proposal. *Protist*. 2010; 161:2–6. [PubMed: 19945343]
- Stuart KD, Schnauffer A, Ernst NL, Panigrahi AK. Complex management: RNA editing in trypanosomes. *Trends Biochem Sci*. 2005; 30:97–105. [PubMed: 15691655]
- Tarleton RL, Reithinger R, Urbina JA, Kitron U, Gurtler RE. The challenges of Chagas Disease-- grim outlook or glimmer of hope. *PLoS Med*. 2007; 4:e332. [PubMed: 18162039]
- Tarun SZ Jr, Schnauffer A, Ernst NL, Proff R, Deng J, et al. KREPA6 is an RNA-binding protein essential for editosome integrity and survival of *Trypanosoma brucei*. *RNA*. 2008; 14:347–358. [PubMed: 18065716]
- Theobald DL, Mitton-Fry RM, Wuttke DS. Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct*. 2003; 32:115–133. [PubMed: 12598368]

- Weng J, Aphasizheva I, Etheridge RD, Huang L, Wang X, et al. Guide RNA-binding complex from mitochondria of trypanosomatids. *Mol Cell*. 2008; 32:198–209. [PubMed: 18951088]
- Worthey EA, Schnauffer A, Mian IS, Stuart K, Salavati R. Comparative analysis of editosome proteins in trypanosomatids. *Nucleic Acids Res*. 2003; 31:6392–6408. [PubMed: 14602897]
- Wu M, Park YJ, Pardon E, Turley S, Hayhurst A, et al. Structures of a key interaction protein from the *Trypanosoma brucei* editosome in complex with single domain antibodies. *J Struct Biol*. 2011; 174:124–136. [PubMed: 20969962]
- Zhou Z, Feng H, Zhou BR, Ghirlando R, Hu K, et al. Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. *Nature*. 2011; 472:234–237. [PubMed: 21412236]

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version.

Name	Multi-OB folds constructs	Number of OB-fold domains	Solubility
A3-9aa-A6		2	High
A3-14aa-A6		2	High
A3-19aa-A6		2	High
A6-9aa-A1		2	High
A6-14aa-A1		2	High
A6-19aa-A1		2	High
A3-9aa-A1 + A6		2 + 1	Low
A3-14aa-A1 + A6		2 + 1	Low
A3-19aa-A1 + A6		2 + 1	Low
A1-A3-9aa-A6		3	Low
A2-A3-9aa-A6		3	Low
A4-A3-9aa-A6		3	Insoluble
A3-9aa-A6-A1		3	Insoluble
A3-9aa-A6-A2		3	Insoluble
A3-9aa-A6-A4		3	Insoluble
A2-A3-14aa-A1 + A6		3 + 1	Low
A4-A3-14aa-A1 + A6		3 + 1	Low
A2-A3-14aa-A1 + A6-A4		3 + 2	Insoluble
A4-A3-14aa-A1 + A6-A2		3 + 2	Insoluble
A1-A3-9aa-A6-A2		4	Insoluble
A1-A3-9aa-A6-A4		4	Insoluble
A2-A3-9aa-A6-A4		4	Insoluble
A2-A3-9aa-A6-A1		4	Insoluble
A4-A3-9aa-A6-A2		4	Insoluble
A4-A3-9aa-A6-A1		4	Insoluble

Fig. 1. Linked OB-folds and expression results

A schematic representation of the multi-OB-fold domain constructs is shown in the second column with the A1 in green, A2 in blue, A3 in red, A4 in purple, A6 in yellow and flexible GS repeat linkers in grey. OB-fold domains and connecting linkers are shown as boxes and solid lines, respectively. Lengths of linker peptides are indicated on the top of solid lines. Recombinant proteins of both fused and separate multi-OB-fold domains were characterized by size exclusion chromatography, as summarized in Table 1. The symbol + indicates that two proteins were co-expressed from different plasmids (see Methods).

Fig. 2. Overall structure of *T. brucei* A3^{OB}-9aa-A6

(a) Overall structure of A3^{OB}-9aa-A6. The structure of the A3^{OB}-aa9-A6 is shown as a ribbon diagram with the A3^{OB} domain, A6^{OB} domain, linker 9aa and N-terminally extended sequences of A3^{OB} depicted in red, yellow, grey and pink, respectively. Three residues of the long N-terminal extension of A3^{OB} interact with Pockets 1 and 2 shown in (b).

(b) Electrostatic surface representation of the A3^{OB}-9aa-A6 structure. The A3^{OB}-9aa-A6 structure is rotated $\sim 90^\circ$ with respect to panel (a). The electrostatic potential surface of A3^{OB}-9aa-A6 was calculated using APBS (Baker et al., 2001). Regions with potentials above $+25 \text{ k}_\text{B} \text{Te}_\text{c}^{-1}$ and below $-25 \text{ k}_\text{B} \text{Te}_\text{c}^{-1}$ are shown in blue and red, respectively. The locations of A3^{OB} and A6^{OB} pockets are indicated with lines in gold, respectively.

Fig. 3. Comparison of A3^{OB} and A6^{OB} dimers from the A3^{OB}-9aa-A6 and A3^{OB}•A6•(A³Nb14)₂ structures

Backbone superpositions of the A3^{OB} and A6^{OB} domains from the current A3^{OB}-9aa-A6 and A3^{OB}•A6•(A³Nb14)₂ structures (PDB-ID: 3STB; (Park et al., 2012a)) are shown. The A3^{OB} domain of A3^{OB}-9aa-A6 is in red; the A6^{OB} domain of A3^{OB}-9aa-A6 in yellow; the A3^{OB} domain of A3^{OB}•A6•(A³Nb14)₂ in blue, and the A6^{OB} domain of A3^{OB}•A6•(A³Nb14)₂ in green. The core structures of the anti-parallel β -barrel in both fused and separate dimers are almost identical with a root-mean-square deviation of 0.93 Å.

Fig. 4. Pockets of A3^{OB}, A6^{OB} and A1^{OB} with interacting residues in four different crystal structures

Pocket 1 (left in each panel) and Pocket 2 (right) are defined in Fig. 2(b). The pockets of A3^{OB}, A1^{OB} and A6^{OB} are shown with their electrostatic surface charge. The interacting residues, from either another OB-fold or from a nanobody (Nb), occupying the OB-fold pockets are labeled. The electrostatic potential surface of each OB-fold domain was calculated using APBS (Baker et al., 2001). Regions with potentials above $+25 \text{ k}_\text{B} \text{Te}_\text{c}^{-1}$ and below $-25 \text{ k}_\text{B} \text{Te}_\text{c}^{-1}$ are shown in blue and red, respectively.

(a – b) Amino acids occupying A3^{OB} Pockets 1 and 2. Pocket 1 and Pocket 2 of A3^{OB} in the current A3^{OB}-9aa-A6 structure occupied by, respectively, A3-L270 and A3-Q267 from another A3^{OB}-9aa-A6 molecule in the crystal, shown in both surface and stereo views. The orientation in panel (b) is the same as in panel (a). The dashed line indicates a hydrogen

bond between R336 and Q313. Motif-I, Motif-II and the invariant arginine (R) residues are labeled in red, blue and orange, respectively.

(c) A^3Nb14 -F102 in $A3^{OB}$ Pocket 1. The $A3^{OB}$ Pocket 1 of the $A3^{OB}\cdot A6\cdot (A^3Nb14)_2$ structure (PDB-ID: 3STB; (Park et al., 2012a)) bound with nanobody residue F102 shown in surface view.

(d) A^1Nb10 -V106 and A^1Nb10 -F101 in $A1^{OB}$ Pockets 1 and 2. The $A1^{OB}$ Pockets 1 and 2 of the $A1^{OB}\cdot A^1Nb10$ structure (PDB-ID: 4DKA; (Park et al., 2012b)) bound with nanobody residues V106 and F101, respectively, shown in surface view. The corresponding stereo view of these A1 pockets with interacting nanobody residues is shown in Supplementary Fig. 11.

(e) A^3Nb14 -F102 in $A6^{OB}$ Pocket 1. The $A6^{OB}$ Pocket 1 of the $A3^{OB}\cdot A6\cdot (A^3Nb14)_2$ structure (PDB-ID: 3STB; (Park et al., 2012a)) bound with nanobody residue F102 is shown in surface view.

(f – g) $A3$ -L262 in $A6^{OB}$ Pocket 1. The A6 Pocket 1 of the current $A3^{OB}$ -9aa-A6 structure bound with A3-L262 shown in both surface and stereo views. The stereo view in panel (f) has the same orientation as panel (g). The dashed line indicates a hydrogen bond between R73 and Q46. Motif-I, Motif-II and invariant arginine (R) residues are labeled in red, blue and orange, respectively.

Summary for solution properties of multi-OB-fold fusions proteins and complexes

Table 1

Linked multi-OB-folds construct					Separate multi-OB-folds construct					
Name of Construct (Mw) ^a	Number of OB-fold domains	S200 SEC Analysis			Oligomerization State	Supplementary Figure	S200 SEC analysis		Oligomerization State	Reference, or Supplementary Figure
		(Mw) ^a	(Ve) ^b	(Ve) ^b						
A3 ^{OB} -9aa-A6 (33 KDa)	2	~68 KDa	14.2		(A3 ^{OB} -9aa-A6) ₂	Sup. Fig. 1(a)				
A3 ^{OB} -14aa-A6 (33 KDa)	2	~68 KDa	14.2		(A3 ^{OB} -14aa-A6) ₂	Sup. Fig. 1(a)				
A3 ^{OB} -19aa-A6 (34 KDa)	2	~68 KDa	14.2		(A3 ^{OB} -19aa-A6) ₂	Sup. Fig. 1(a)	~83 KDa	13.9	(A3 ^{OB} -A6) ₃	Fig. 4D in Ref ^c
A6-9aa-A1 ^{OB} (32 KDa)	2	~50 KDa	14.7		(A6-9aa-A1 ^{OB}) ₂	Sup. Fig. 1(b)				
A6-14aa-A1 ^{OB} (32 KDa)	2	~50 KDa	14.7		(A6-14aa-A1 ^{OB}) ₂	Sup. Fig. 1(b)				
A6-19aa-A1 ^{OB} (32 KDa)	2	~50 KDa	14.7		(A6-19aa-A1 ^{OB}) ₂	Sup. Fig. 1(b)	~60 KDa	S75 SEC ^d ; 9.9	(A6) ₄ and (A1 ^{OB} -A6) ₂	Sup. Fig.5(a)
A3 ^{OB} -9aa-A1 ^{OB} •A6 (48 KDa)	2 + 1	~60 KDa	14.4		A3 ^{OB} -9aa-A1 ^{OB} •A6	Sup. Fig. 3(a)				
A3 ^{OB} -14aa-A1 ^{OB} •A6 (48 KDa)	2 + 1	~68 KDa	14.2		A3 ^{OB} -14aa-A1 ^{OB} •A6	Sup. Fig. 3(b)	~136 KDa	13.1	(A3 ^{OB} •A1 ^{OB} •A6) ₂	Sup. Fig. 5(d)
A3 ^{OB} -19aa-A1 ^{OB} •A6 (49 KDa)	2 + 1	~60 KDa ~136 KDa	14.4 13.1		A3 ^{OB} -19aa-A1 ^{OB} •A6 (A3 ^{OB} -19aa-A1 ^{OB} •A6) ₂	Sup. Fig. 3(d)				
A1 ^{OB} -A3 ^{OB} -9aa-A6 (48 KDa)	3	~45 KDa	14.9		A1 ^{OB} -A3 ^{OB} -9aa-A6	Sup. Fig. 2(a)	~136 KDa	13.1	(A3 ^{OB} •A1 ^{OB} •A6) ₂	Sup. Fig. 5(d)
A2 ^{OB} -A3 ^{OB} -9aa-A6 (46 KDa)	3	~49 KDa	14.7		A1 ^{OB} -A3 ^{OB} -9aa-A6	Sup. Fig. 2(c)	~110 KDa	13.5	(A2 ^{OB} •A3 ^{OB} •A6) ₂	Fig. 7D in Ref ^c
A2 ^{OB} -A3 ^{OB} -14aa-A1 ^{OB} •A6 (61 KDa)	3 + 1	> 160 KDa	11.2		(A2 ^{OB} -A3 ^{OB} -14aa-A1 ^{OB} •A6) ₃	Sup. Fig. 4(b)				
A4 ^{OB} -A3 ^{OB} -14aa-A1 ^{OB} •A6 (62 KDa)	3 + 1	> 160 KDa	11.2		(A4 ^{OB} -A3 ^{OB} -14aa-A1 ^{OB} •A6) ₃	Sup. Fig. 4(a)				

Note:

^a molecular weight (kDa),

^b elution volume (ml),

^cReference: (Schnauder et al., 2010),

^dA1^{OB}•A6 was analyzed by Superdex 75 10/300 GL (GE Healthcare) filtration column in 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP and 10% glycerol buffer.

Table 2

Crystallographic data collection and refinement statistics.

	Native
Data collection	
Space group	P2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	57.7, 59.4, 100.7
α , β , γ (°)	90, 90, 90
Wavelength (Å)	0.9795
Resolution (Å)	40-2.55
<i>R</i> _{merge} (%)	10 (55)
<i>I</i> / σ <i>I</i>	14.2 (1.6)
Completeness (%)	95.5 (70.7)
Redundancy	5.9 (3.7)
Refinement	
Resolution (Å)	40-2.55
No. reflections	11273
<i>R</i> _{work} / <i>R</i> _{free} (%)	21.4 / 24.7
No. atoms	
Protein	2027
Ligand/ion	0
Water	14
<i>B</i> -factors (Å ²)	
Protein	57.9
Water	42.7
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.981
Ramachandran plot ^a	
Preferred (%)	96.08
Allowed (%)	3.14
Outliers (%)	0.78

Values in parentheses are for the highest resolution shell.

^aRamachandran plot analysis by Molprobity (Chen et al., 2010)