Structures of a key interaction protein from the Trypanosoma brucei editosome in complex with single domain antibodies

Wu, Meiting; Park, Young-Jun; Pardon, Els; Turley, Stewart; Hayhurst, Andrew; Deng, Junpeng; Steyaert, Jan; Hol, Wim G. J.

Published in:
Journal of Structural Biology

DOI:
10.1016/j.jsb.2010.10.007

Publication date:
2011

Document Version:
Final published version

Citation for published version (APA):
Structures of a key interaction protein from the *Trypanosoma brucei* editosome in complex with single domain antibodies

Meiting Wu, Young-Jun Park, Els Pardon, Stewart Turley, Andrew Hayhurst, Junpeng Deng, Jan Steyaert, Wim G.J. Hol

**Abstract**

Several major global diseases are caused by single-cell parasites called trypanosomatids. These organisms exhibit many unusual features including a unique and essential U-insertion/deletion RNA editing process in their single mitochondrion. Many key RNA editing steps occur in ~20S editosomes, which have a core of 12 proteins. Among these, the “interaction protein” KREPA6 performs a central role in maintaining the integrity of the editosome core and also binds to ssRNA. The use of llama single domain antibodies (VHH domains) accelerated crystal growth of KREPA6 from *Trypanosoma brucei* dramatically. All three structures obtained are heterotetramers with a KREPA6 dimer in the center, and one VHH domain bound to each KREPA6 subunit. Two of the resultant heterotetramers use complementarity determining region 2 (CDR2) and framework residues to form a parallel pair of beta strands with KREPA6 – a mode of interaction not seen before in VHH domain–protein antigen complexes. The third type of VHH domain binds in a totally different manner to KREPA6. Intriguingly, while KREPA6 forms tetramers in solution adding either one of the three VHH domains results in the formation of a heterotetramer in solution, in perfect agreement with the crystal structures. Biochemical solution studies indicate that the C-terminal tail of KREPA6 is involved in the dimerization of KREPA6 dimers to form tetramers. The implications of these crystallographic and solution studies for possible modes of interaction of KREPA6 with its many binding partners in the editosome are discussed.

**1. Introduction**

Trypanosomatids are a group of unicellular eukaryotic organisms of which several members are the causative agents of major human diseases. These include: (i) *Trypanosoma brucei*, causing human trypanosomiasis, also called sleeping sickness, with ~60 million people at risk for acute infections in sub-Saharan Africa; *Trypanosoma cruzi*, responsible for Chagas disease, also called American trypanosomiasis, with ~25 million people at risk in Latin America, and (iii) various *Leishmania* species, which cause cutaneous and disseminated leishmaniasis in the tropics and subtropics with ~350 million people in danger of infection (http://www.who.int/health-topics/idindex.htm, http://www.cdc.gov/chagas/factsheet.html). The available drugs for these tropical parasitic diseases are limited by poor efficacy, toxicity, and/or increasing resistance (Croft et al., 2006; Fairlamb, 2003; Hotz et al., 2007; Tarleton et al., 2007). Vaccines have been difficult to develop for these parasites due to their ability to undergo antigenic variation and thereby evade the immune response (Smith and Parsons, 1996). Hence, there is a tremendous need for new therapeutic agents to treat the diseases caused by these pathogenic protozoa.

Trypanosomatids contain a substantial number of unusual and essential biochemical characteristics, several of which are currently being explored as targets for new chemotherapeutics (De Souza, 2002; Hammarton, 2007; Moyersen et al., 2004). One of the most remarkable features of these organisms is a very elaborate U-insertion/deletion RNA editing process of most messenger...
RNAs transcribed from genes encoded on the mitochondrial DNA of these organisms (Panigrahi et al., 2003; Schnaufer et al., 2003; Simpson et al., 2004; Stuart et al., 2005). In this editing process, a so-called pre-messenger RNA (pre-mRNA) is edited in a large series of steps according to sequence information from relatively small RNA's, called guide RNA's (gRNA's). Guided by information from many different gRNA's, many more U's are inserted than deleted by the editosome into the mRNA.

This U-insertion/deletion RNA editing process in the mitochondria of trypanosomatids requires a large number of proteins that are encoded on nuclear DNA. Once inside the mitochondrion, many of these proteins assemble into several large multi-protein complexes (Aphasizhev et al., 2003; Lukes et al., 2005; Rusche et al., 1997; Weng et al., 2008). One of these complexes is the ~20S editosome complex, hereafter called the editosome (reviewed in Simpson et al. (2004) and Stuart et al. (2005)). Recent electron microscopy studies have revealed an elongated shape of the editosome with dimensions of ~80 by ~140 by ~200 Å (Golas et al., 2009; Li et al., 2009). Evidence has been provided for the presence of three different types of editosomes that share a common core of 12 proteins (Panigrahi et al., 2006; Carnes et al., 2005, 2008) (Supplementary Fig. 1(a)). Multiple nomenclatures for the proteins in the editosome are summarized in Supplementary Fig. 1(b). The editosome core contains a large number of proteins which can be grouped as follows (Supplementary Fig. 1):

(i) The four enzymes KREX2, KREX2, KREL1 and KREL2. The enzyme KREX2 is a 3’ → 5’-exonuclease which removes U’s from the cleaved pre-mRNA (Ernst et al., 2009). KREX2 is a 3’ terminal uridylyltransferase (TUTase) adding U’s to the cleaved pre-mRNA. The crystal structure of T. brucei KREX2 revealed the structural basis of its U-specificity (Deng et al., 2005). KREL1 and KREL2 are two related RNA editing ligases which seal the mRNA after removal or addition of U’s (Deng et al., 2004).

(ii) Two proteins with an RNase III-like domain, KREPB4 and KREPB5, which are most likely critical for linking the editosome core with the two or three specific extra proteins per type of editosome (Carnes et al., 2008).

(iii) Six so-called “interaction proteins”, KREPA1 to KREPA6. These six proteins vary greatly in length, yet each contains a predicted “OB-fold” near the C-terminus (Brech et al., 2005; Drozdz et al., 2002; Kang et al., 2004; Law et al., 2007, 2008; Panigrahi et al., 2006, 2001; Salavati et al., 2006; Schnaufer et al., 2003; Worthey et al., 2003). A key interaction protein is KREPA6, the three-dimensional architecture of which is the focus of the current paper.

Trypanosomatid KREPA6 contains 164–229 amino acids depending on species (Fig. 1(a)) and is a remarkable multifunctional protein, central to the integrity of the entire editosome (Tarun et al., 2008). It interacts with four other proteins: KREPB1, KREPB2, KREPB3 and KREPB4 (Schnaufer et al., 2003, 2010) (Supplementary Fig. 1(a)). In addition, KREPA6 interacts with ssRNA (Tarun et al., 2008). KREPA6 belongs to the large single strand nucleic acid-binding superfamily of OB-folds (Worthey et al., 2003), hereafter also called SSB domains and SSB proteins (Arcus, 2002; Murzin, 1993; Theobald et al., 2003). In eukaryotes, SSB-domains are encoded on nuclear DNA. Once inside the mitochondrion, many different gRNA’s, many more U’s are inserted than deleted by the editosome into the mRNA.

So-called P-dyad of the canonical SSB tetramer is shared by the two dimers (Saikrishnan et al., 2005), but a considerable variation occurs regarding the orientation of the two dimers with respect to each other (Chan et al., 2009; DiDonato et al., 2006; Jedrzejczak et al., 2006; Raghunathan et al., 1997; Saikrishnan et al., 2003, 2005; Yang et al., 1997). An interesting case regarding multimerization is E. coli PriB (EcpPriB) which can form typical SSB-dimers (Liu et al., 2004; Lopper et al., 2004; Shoji et al., 2005) and tetramers (Shoji et al., 2005). Hence, SSB proteins have been shown to be able to exist as monomers, dimers and tetramers.

Here we describe crystal structures of T. brucei KREPA6 in complex with three different single domain antibodies (VHH domains) derived from llamas (Desmyter et al., 2002; Goldman et al., 2006; Muylleman, 2001). The use of VHH domains as crystallization chaperones was critical in obtaining well-diffracting crystals of KREPA6 in complex with VHH domains. Two of the complexes showed that CDR2 and framework 3 of the VHH domain form a β-strand which engages in parallel pairing with a β-strand of KREPA6, a mode of interaction not seen before in protein-VHH domain complexes to the best of our knowledge. Surprisingly, all three VHH domains exhibited the capacity in solution to break the KREPA6 tetramer and form (KREPA6)2:(VHH-domain)2 heterotetramers. The hypothesis that the C-terminal tail of KREPA6 might be involved in KREPA6 tetramer formation was confirmed by solution studies. The implications of these findings for the possible modes of interactions of KREPA6 with other proteins in the core of the editosome are discussed.

2. Materials and methods

2.1. KREPA6 cloning, expression and purification

The gene fragment corresponding to KREPA6 (residues 19–164; the first 18 amino acids are the predicted mitochondrial import signal (Tarun, 2008)) was PCR amplified from genomic DNA of T. brucei and cloned into pRSET vector (Invitrogen) for expression in T. brucei and cloned into pRSET vector (Invitrogen) for expression.

2.2. Single domain antibody selection, cloning, expression and purification

2.2.1. Nanobody generation

A llama was immunized six times with 330 µg of purified recombinant KREPA6/KREL1/KREPA2 protein complex over a period of 6 weeks. From the anti-coagulated blood of the immunized llama, lymphocytes were used to prepare cDNA which served as template to amplify genes coding for the variable domains of the heavy-chain antibodies. The procedures were essentially as described elsewhere (Lam et al., 2009) and summarized in Supplemental material.

2.2.2. Solid-phase ELISA

Maxisorb 96-well plates (Nunc) were coated with 100 µl purified T. brucei KREPA6 overnight at 4 °C at 1 µg/ml in sodium bicarbonate buffer pH 8.2. Residual protein binding sites in the wells
were blocked for 2 h at room temperature with 2% milk in PBS. Detection of antigen-bound nanobodies was performed with a mouse anti-haemaglutinin-decapeptide-tag (clone 16B12, BAbCO) or a mouse anti-histidine-tag (Serotec), as appropriate. Subsequent detection of the mouse anti-tag antibodies was done with an alkaline phosphatase anti-mouse-IgG conjugate (Sigma). The absorption at 405 nm was measured 15 min after adding the enzyme substrate p-nitrophenyl phosphate.

2.2.3. Nanobody cloning, expression and purification

The gene fragment containing Nb5 was subcloned into a pRSF vector (Invitrogen) with a non-cleavable C-terminal hexahistidine tag. Nb5 was expressed in E. coli BL21 (DE3). Cells were grown to an OD600 of 1.5–2 at 37 °C in terrific broth (TB) and induced with 0.5 mM IPTG at 28 °C for 3 h. Cells were harvested by centrifugation and resuspended in buffer A (20 mM Tris–HCl pH 8, 500 mM NaCl, 20 mM imidazole and 10% glycerol). Cells were lysed in a...
French Press and insoluble material was removed by centrifugation. The supernatant was passed through a Ni-NTA column (Qiagen), washed with buffer A, and eluted with buffer A containing 250 mM imidazole. Protein was applied onto a Superdex-75 column (GE Health) for size-exclusion chromatography in buffer D (20 mM HEPES, pH 7.5, 300 mM NaCl, 2 mM DTT and 10% glycerol). Nb15 was cloned, expressed and purified essentially by the same procedures as described above for Nb5. The molecular weights of the purified His6-tagged Nb5 and Nb15 nanobodies are 14.0 and 14.1 kDa, respectively.

2.2.4. VHH library screening and Ab expression and purification

The construction of Nomad #1, a hyperversified semi-synthetic llama VHH phage library, was essentially carried out as previously reported (Goldman et al., 2006) as described in Supplementary material.

2.2.5. Expression of soluble VHH domains from the Nomad library

Two binders (Ab1 and Ab2) were sub-cloned into the expression vector pRSF22b (Invitrogen) to generate an expression clone producing these VHH domains with a C terminal hexahistidine tag. Cells were grown in TB media containing kanamycin. IPTG was added at a final concentration of 0.5 mM when the OD at 600 nm was between 1.8 and 2.0. The culture was incubated for an additional 3 h at 28 °C. Ab1 and Ab2 proteins were purified using a nickel affinity column followed by gel filtration on a Superdex 75 column. The molecular weight of the purified His6-tagged Ab2 VHH domain is 13.9 kDa.

2.3. Complex formation, crystallization and structure determination

For crystallization of KREPA6:Nb15, purified KREPA6 was mixed with Nb15 at a 1:1.2 molar ratio to form the complex. The KREPA6:Nb15 complexes were purified with Superdex 75 gel filtration column. The KREPA6:Nb15 complex produced crystals by the sitting drop vapor diffusion method of mixing 2 μl of the complex solution (4 mg/ml protein) with 2 μl reservoir solution containing 25% PEG2000, 0.1 M Bis–tris buffer pH 6.5, at 20 °C.

For crystal growth of KREPA6:Nb5, KREPA6 and Nb5 solutions were mixed at a 1:1 molar ratio at 4 °C for 30 min prior to setting up crystallization experiments. Crystals were obtained from vapor diffusion experiments using 200 nL of protein solution plus 200 nL of a reservoir solution containing 20% PEG4000, 10% iso-proanol and 0.1 M HEPES pH 7.5, using a Phoenix crystallization robot. Crystals appeared at 20 °C within 72 h. Just like in the case of the KREPA6:Nb15 complex, the KREPA6:Nb5 complex exhibited a molecular weight of approximately 60 kDa on the sizing column (Supplementary Fig. 2(c)).

KREPA6:Ab2 crystals were obtained by first mixing purified KREPA6 with antibody Ab2 at a 1:1.2 molar ratio to form the complex. The KREPA6:Ab2 complex was purified with a Superdex 75 gel filtration column followed by a protein concentration step (Centriprep). The KREPA6:Ab2 complex was crystallized at 20 °C by sitting drop vapor diffusion. Protein (4 mg/ml protein, in 20 mM Tris–HCl pH 7.5, 2 mM dithiothreitol, 300 mM NaCl) was combined in a 1:1 ratio with a reservoir solution containing 15% PEG3350, 0.1 M citrate pH 4.2. Thin, plate-like crystals with a maximum thickness of 20 μm appeared within hours, and grew to a maximum size of 100 by 200 μm in the two other directions within one week. As with the cases of the KREPA6:Nb15 and the KREPA6:Nb5 complexes, the KREPA6:Ab2 complex exhibited a molecular weight of approximately 60 kDa on the sizing column (Supplementary Fig. 2(d)).

KREPA6:Nb5 crystals were transferred to a precipitant solution supplemented with 25% glycerol for cryoprotection in liquid nitrogen. Data were collected on the Saturn in-house CCD detector (Rigaku) and processed by HKL2000 (Otwinowski and Minor, 1997). The crystals belong to space group C2, with cell dimensions a = 120.3 Å, b = 52.5 Å, c = 85.8 Å and β = 104.6° (Table 1) with two Nb5 and two KREPA6 monomers in the asymmetric unit. The structure of KREPA6:Nb5 was solved by molecular replacement using antibody 1I3V (Spinelli et al., 2001) as a search model for Nb5, and the structurally homologous protein EcSSb (Matsumoto et al., 2000) with PDB code 1EOQ as a search model for KREPA6.

PHASER (McCoy et al., 2007) was able to find a solution using both KREPA6 and Nb5 as a search model. Refinement (Table 1) with two Nb5 and two KREPA6 monomers in the asymmetric unit. The structure of KREPA6:Nb5 was solved by molecular replacement using antibody 1I3V (Spinelli et al., 2001) as a search model for Nb5, and the structurally homologous protein EcSSb (Matsumoto et al., 2000) with PDB code 1EOQ as a search model for KREPA6.

PHASER (McCoy et al., 2007) was able to find a solution using both the nanobody alone or using KREPA6 alone. Combining two nanobodies and one KREPA6 dimer as search model gave the best solution. Manual model building in subsequent electron density maps was performed with COOT (Emsley and Cowtan, 2004), alternating with refinement using REFMAC5 (Murshudov et al., 1997) and PHENIX (Adams et al., 2002). The final structure at 2.4 Å resolution had Rwork = 22% and Rfree = 27%. Data collection and refinement statistics are shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Crystallographic data collection and refinement statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data set</td>
<td>KREPA6:Nb15</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9795</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>(2.00–2.1)</td>
</tr>
<tr>
<td>(Outer shell) (Å)</td>
<td>(2.52–2.35)</td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
</tr>
<tr>
<td>a (Å)</td>
<td>123.4</td>
</tr>
<tr>
<td>b (Å)</td>
<td>50.4</td>
</tr>
<tr>
<td>c (Å)</td>
<td>44.1</td>
</tr>
<tr>
<td>β (°)</td>
<td>109.0</td>
</tr>
<tr>
<td>Total reflections</td>
<td>58,033</td>
</tr>
<tr>
<td>Total unique reflections</td>
<td>15,229</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.6 (76.2)</td>
</tr>
<tr>
<td>lσ (l)</td>
<td>30.6 (3.1)</td>
</tr>
<tr>
<td>Residues per ASU</td>
<td>287</td>
</tr>
<tr>
<td>Water per ASU</td>
<td>70</td>
</tr>
<tr>
<td>RКАMP Average B factors (Å²)</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Values in parentheses are for highest-resolution shell. Rwork = Σ|Fhkl|−|Фk|)/Σ|Fhkl|, and Rfree = Σ|Fhkl|−|Фk|)/Σ|Fhkl|.

Rfree is defined as Rwork, but for a set of reflections not used in refinement.

(Rigaku) and processed by HKL2000 (Otwinowski and Minor, 1997). The crystals belong to space group C2, with cell dimensions a = 120.3 Å, b = 52.5 Å, c = 85.8 Å and β = 104.6° (Table 1) with two Nb5 and two KREPA6 monomers in the asymmetric unit. The structure of KREPA6:Nb5 was solved by molecular replacement using antibody 1I3V (Spinelli et al., 2001) as a search model for Nb5, and the structurally homologous protein EcSSb (Matsumoto et al., 2000) with PDB code 1EOQ as a search model for KREPA6. Values in parentheses are for highest-resolution shell. Rwork = Σ|Fhkl|−|Фk|)/Σ|Fhkl|, and Rfree = Σ|Fhkl|−|Фk|)/Σ|Fhkl|.

Rfree is defined as Rwork, but for a set of reflections not used in refinement.
(McCoy et al., 2007) with the previous model of KREPA6 and a model of Nb5 with CDR loops removed. Manual model building in subsequent electron density maps was performed with COOT (Emsley and Cowtan, 2004), alternating with refinement using REFMACS (Murshudov et al., 1997) and TLS parameters to model anisotropic displacements (Winn et al., 2001). The final structure at 2.1 Å resolution had $R_{\text{work}} = 20\%$ and $R_{\text{free}} = 26\%$. Additional statistics are shown in Table 1.

KREPA6:Ab2 crystals were cryo-protected with 30% of glycerol and flash cooled in liquid nitrogen. Diffraction data were collected at SSRF beamline BL9-2. Data were processed and reduced with HKL2000. The structure was solved by the molecular replacement method with PHASER (McCoy et al., 2007), using the KREPA6 dimer model derived from the refined KREPA6:Nb5 heterotetramer and coordinates of antibody HCV (Spinelli et al., 2001). The model of the heterotetramer in the asymmetric unit was manually improved with the program Coot (Emsley and Cowtan, 2004) alternating with coordinate and $B$-factor refinement using REFMACS (Emsley and Cowtan, 2004; Murshudov et al., 1997), using 1 TLS group per chain, and PHENIX (Adams et al., 2002). Crystallographic data collection and refinement statistics are shown in Table 1. These statistics are comparable to those of other structures solved at 3.5 Å resolution (Bubb et al., 2002; Ni et al., 2000; Stein et al., 2009).

### 2.4. Figure preparation

Pictures of molecules were prepared using PYMOL (DeLano Scientific Research LLC). Sequence alignment figures were made with Esprirt (Gouet et al., 1999).

### 2.5. Protein data bank accession numbers

Coordinates and structure factors have been deposited with accession codes 3K7U for KREPA6:Nb15, 3K80 for KREPA6:Nb5, and 3K81 for KREPA6:Ab2.

### 3. Results

#### 3.1. Overview of crystallographic results

Obtaining diffraction quality crystals of KREPA6 was a major challenge. Despite trying thousands of different conditions in crystal screens; different trypansomatid species; variations in length; mutating residues (such as, e.g. all four Cys to Ser); and adding RNAs of different lengths, no trace of crystal growth was observed. We were able to enrich CDR3 sequences for expression and co-crystallization trials with KREPA6. Eventually, one of these VHH domains, Ab2, yielded crystals which diffracted beyond 4 Å resolution.

The three structures solved of KREPA6 in complex with three different llama VHH domains (Table 1) revealed two entirely different binding modes of VHH domains to essentially the same KREPA6-dimer in the center of the complex. Since the KREPA6:Nb15 heterotetramer structure has the highest resolution, this structure will be described first, and then compared with the KREPA6:Nb5 and KREPA6:Ab2 heterotetramers.

#### 3.2. The heterotetramer formed by KREPA6 and Nb15

The KREPA6 monomer contains the canonical five β-strands of the OB fold and an α-helix between strands 3 and 4 (Fig. 1(a) and (b)) in agreement with earlier predictions (Worthey et al., 2003). Loop L12 is well defined in the eventual electron density maps but the longer loops L23 and L45 are missing clear density for residues 55–65 and 109–113, respectively. Also the N-terminal residues 19 and 20 and the C-terminal residues 135–164 of KREPA6 are not represented by well defined density. In contrast, the density for all residues of Nb15 is well defined except for the two N-terminal residues.

While there is only one KREPA6 and one Nb15 chain per asymmetric unit, two KREPA6 subunits related by a crystallographic twofold axis form a tight dimer (Fig. 1(b)) with a buried solvent accessible surface of 2030 Å² and a calculated $\Delta G$ of interaction of $-6.2$ kcal mol$^{-1}$ according to the PISA server (Krissinel and Henrick, 2007). The residues involved in intersubunit interactions (Fig. 1), include strands β1 and β1’ from the second subunit, which run anti-parallel to each other in the dimer. As a result, an extended six-stranded anti-parallel β-sheet (β6–β4–β1–β1’–β4’–β6’) is formed by the dimer, with the twofold axis, called P-axis in related structures (Saikrishnan et al., 2003), running perpendicular to the sheet. The dimer interface residues of KREPA6 are well conserved with 22 of the 25 interface residues being the same in all trypansomatids (Fig. 1(a)). The loop lengths and loop residues are also very similar in trypansomatids, therefore KREPA6 adopts very likely the same dimer structure across these species.

KREPA6 and Nb15 form a $\alpha_2\beta_2$-type heterotetramer of ~60 kDa in the crystals, with ~16 kDa per KREPA6 subunit and ~14 kDa per Nb15 nanobody. A KREPA6 dimer occurs in the center of the tetramer. Each KREPA6 subunit interacts with one separate Nb15 chain, without any contacts between the two Nb15 subunits in the complex (Fig. 2(a)). The presence of a heterotetramer in the crystals is in agreement with solution studies (Supplementary Fig. 2(a) and (b)). The interface of Nb15 with T. brucei KREPA6 includes a parallel pair of β-strands: β6 from KREPA6 and βC’ from Nb15 (Fig. 2(a) and (b)). Strand βC’ comprises four residues (58–61) from framework 3 and residue L57 from CDR2. Therefore, CDR2 and framework 3 residues are the main elements of Nb15 responsible for the interactions with KREPA6, with 290 and 303 Å² buried surface, respectively (Fig. 2(b) and (c)). The contacts made by CDR3 are more limited with only 131 Å² buried surface area, while CDR1 does not contribute to the binding to KREPA6. A total of 17 residues from KREPA6 and 17 residues from Nb15 are engaged in these interactions (Figs. 1(a), 2(b), and 3(a)), burying 1390 Å² solvent accessible surface with a calculated interaction $\Delta G$ of $-8$ kcal mol$^{-1}$ according to the PISA server (Krissinel and Henrick, 2007). Regarding KREPA6, the C-terminal strand β6 provides most of the contacts while 11 additional residues, distributed over $\alpha_1$, $\alpha_{24}$, and $\alpha_{25}$, interact with Nb15 (Figs. 1(a) and 2(c)).
3.3. The structure of KREPA6 in complex with Nb5

The amino acid sequences of Nb5 and Nb15 are 76% identical for 122 residues. The framework residues display a total of 13 differences out of 90 amino acids while CDR1 differs at 4 positions out of 8 amino acids, CDR2 at 2 positions out of 7, and CDR3 at 13 out of 19 (Fig. 3(a)). In spite of these substantial sequence differences, the crystals, both with space group C2, of the KREPA6:Nb5 and KREPA6:Nb15 complexes are very similar. If the c-axis of the KREPA6:Nb15 complex is multiplied by a factor of two then all cell dimensions differ less than 2.5 Å from those of the KREPA6:Nb5 complex, with almost the same β-angle (Table 1). The overall three-dimensional structures of the two nanobodies are also very similar, including the CDR2 loops. CDR2 of Nb5 engages residues SRSGL (residues 52–57) to contact KREPA6 while CDR2 of Nb15 uses residues SWTGGL (Fig. 3(a)). Regarding the two CDR2 residues which differ in side chains, both R53 and S54 of Nb5 form hydrogen bonds with E82 of KREPA6 (Fig. 3(a)). This is only slightly different from the interactions in the KREPA6:Nb15 complex where W53 of Nb15 makes weak hydrophobic interactions with L78 of KREPA6, while T54 of Nb15 forms a hydrogen bond with E82 of KREPA6. The five residues of framework 3 right next to CDR2 form strand βC” which runs parallel to strand β6 of KREPA6 (Figs. 2 and 3).

In contrast with the CDR2s, the CDR3 loops in these two nanobodies adopt substantially different conformations (Supplementary Fig. 4). CDR3 of Nb15 uses residues GDYV (residues 102–105) to contact KREPA6 and CDR3 of Nb5 residues GLGSHE (Fig. 3(a)). In both heterotetramers, CDR3 only buries ~130 Å² of solvent accessible surface in its contacts with the protein antigen. The observation that neither of these very different CDR3 loops are engaged in crystal contacts (not shown), and neither appear to alter significantly the orientation of these two VH4 domains with respect to the KREPA6 subunit bound, may explain why both complexes crystallize in similar crystal lattices, despite considerable differences in sequence, length and conformation of the CDRs in the two nanobodies.

There is an entire heterotetramer in the asymmetric unit of the KREPA6:Nb5 crystals. The two KREPA6:Nb5 pairs per heterotetramer superimpose with an rms deviation of 0.8 Å for 217 Cα
atoms, indicating that the two halves are very similar. The KREPA6 dimers are essentially indistinguishable in the KREPA6:Nb15 and KREPA6:Nb15 heterotetramers with an rms deviation of 0.6 Å for 196 equivalent Cα atoms. The orientation of the Nb5 chain with respect to the KREPA6 monomer differs by a mere 3.8° compared to the orientation of Nb15 with respect to KREPA6, yielding very similar heterotetramers (Fig. 3(b)).

### 3.4. The structure of KREPA6 in complex with Ab2

The VHH domains Ab1 and Ab2 obtained from screening of the Nomad #1 semi-synthetic library were overexpressed, purified and tested in co-crystallization experiments (Supplementary Fig. 5). This yielded eventually crystals of Ab2 in complex with KREPA6 which diffracted to 3.4 Å resolution. Ab2 is 65.6% identical in amino acid sequence compared to Nb15, with substantial differences in the CDRs: 4 out of 8 residues differ in CDR1, 4 out of 6 residues in CDR2, and 12 out of 17 residues in CDR3 (Fig. 3(a)). Utilizing the KREPA6 dimer determined in the previous two structures, and a homology model of Ab2, the structure of the (KREPA6:Ab2)_2 complex could be determined by the molecular replacement method (see Section 2). There appeared to be a heterotetramer in the asymmetric unit with the KREPA6 dimer in the center of the complex. The heterotetramer in the crystals is in agreement with solution studies which revealed that adding Ab2 VHH domains to KREPA6 homotetramers yields (KREPA6:Ab2)_2 heterotetramers (Supplementary Fig. 2(d)).

The Ab2 domains contact the KREPA6 dimer in a totally different manner than both Nb15 and Nb5 (Fig. 4(a) and (b)). The interactions of Ab2 occur on the opposite face of the dimer compared to Nb15 and Nb5. Ab2 binds to the long loops L45 and the tip of the loop L12 of KREPA6. A total of 17 Nb15 residues, from CDR2, CDR3 and framework 3 residues, are engaged in contacts with KREPA6 (Fig. 1(a)). In contrast to the two previous structures, CDR3 is of major importance for the Ab2-antigen interaction. Framework and CDR3 residues are the main elements of Ab2 responsible for the interactions with KREPA6, with 191 and 438 Å² buried surface, respectively. Although all three complexes crystallized in space group C2, the crystal form and crystal packing of the KREPA6:Ab2 heterotetramers is entirely different than in the KREPA6:Nb15 and KREPA6:Nb5 crystals. In the latter two cases, the KREPA6 dimer is

---

**Fig. 3.** Comparison of the KREPA6:Nb15 and KREPA6:Nb5 heterotetramers. (a) Sequence alignment of Nb15, Nb5 and Ab2. The secondary structure elements correspond to the crystal structure of Nb15. The numbering on top of the alignment corresponds to the continuous numbering present in the PDB file 3K7U, the numbering on the bottom of the alignment corresponds to the standard IMGT numbering for antibodies and related proteins (Lefranc, 2005). The conserved cysteines that form the intra-molecular disulfide bridge are highlighted in blue. CDR1 is highlighted in green, CDR2 in purple and CDR3 in red. Triangles indicate Nb15 contact residues with KREPA6, circles indicate Nb5 contact residues with KREPA6, and diamonds indicate Ab2 contact residues with KREPA6. (b) Superposition of KREPA6:Nb15 and KREPA6:Nb5. The KREPA6:Nb5 structure is superimposed onto the KREPA6:Nb15 structure by using only the KREPA6 dimers for calculating the superposition operation. The Nb5 and Nb15 domains differ 3.8° in orientation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
surrounded by nanobodies and is essentially not in contact with other KREPA6 dimers (Supplementary Fig. 6(a)). In the KREPA6:Ab2 crystals, contacts between helices of adjacent KREPA6 dimers are generated by a crystallographic twofold symmetry axis such that layers of KREPA6 dimers are obtained which traverse the entire crystal (Fig. 4(c), Supplementary Fig. 6(b)). Approximately 570 Å² is buried in this helix–helix interface which engages two hydrophobic and five hydrophilic residues from each KREPA6 monomer.

3.5. Comparison of KREPA6 with other OB fold proteins

A homologous structure search using the DALI server showed that the E. coli single-strand DNA-binding protein EcSSB (PDB code 1EYG (Raghunathan et al., 2000)) and the homologous protein HpSSB from Helicobacter pylori (PDB code 2WV9 (Chan et al., 2009)) are the closest structural homologs of KREPA6. The DALI Z-score is 13.4 for 84 equivalent pairs of Cα atoms with 24% amino acid identity between single subunits of T. brucei KREPA6 and EcSSB. For KREPA6 and HpSSB subunits, the comparison yields a DALI Z-score of 12.6 for 98 equivalent Cα atom pairs with 20% amino acid identity. Not only the monomers but also the dimers of KREPA6 and EcSSB are quite similar with an rms deviation of 1.3 Å for 163 equivalent Cα atoms. The equivalent numbers for the HpSSB dimer versus KREPA6 dimer comparison are 1.4 Å for 160 Cα atoms. Clearly, the same dimer organization, with the β1 strands of adjacent neighbors running anti-parallel to each other creating a six-stranded β-sheet spanning the dimer, is obtained despite the relatively low sequence identity shared by these three SSB proteins.

3.6. Trypsinolysis creates KREPA6 dimers

The fact that the Ab2 VH domain could split the KREPA6 tetramer into KREPA6 dimers (Fig. 4(a)), led us to hypothesize (see also Section 4) that the C-terminal tail of KREPA6 could be involved in tetramerization. We therefore investigated the effect of limited proteolysis by trypsin on KREPA6 multimerization, since the location of Arg and Lys residues along the KREPA6 chain is such that trypsin would be able to cleave off the C-terminal tail but not...
the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)). The combination of size exclusion chromatography, SDS–PAGE analysis and Ni-NTA affinity chromatography (Fig. 5) shows that the treatment of trypsin changes the ~60 kDa tetramer of full-length, N-terminally His6-tagged, KREPA6 into a ~30 kDa dimer of KREPA6 subunits shortened by ~30 residues. This dimer still interacts with the Ni-NTA column indicating the N-terminal tail (Fig. 1(a)).

4. Discussion

4.1. KREPA6 dimer and tetramer

While KREPA6 forms tetramers in solution, a surprising result of our studies is that KREPA6 forms dimers in the center of all three heterotetramer structures solved (Figs. 2(a), 3(b), and 4(a)). The dimers are similar to those seen in other SSBS proteins. However, in many SSBS proteins these dimers form tetramers, with the β-sheet surfaces of the dimers facing each other (Raghunathan et al., 1997; Theobald et al., 2003; Yang et al., 1997) (Supplementary Fig. 7). If the hydrophobic β-sheet surfaces of KREPA6 indeed face each other in the solution tetramer, as would be the case if KREPA6 forms a canonical tetramer of SSBS folds with D2 symmetry (Supplementary Fig. 7), then the Nb15 and Nb5 binding sites overlap with the position of the second KREPA6 dimer in the putative architecture of the KREPA6 homotetramer and hence a KREPA6 tetramer cannot remain intact and simultaneously bind to Nb5 or Nb15. The affinities of these nanobodies for the KREPA6 dimer must be higher than of the KREPA6 dimers for each other.

In contrast, the binding site of Ab2 near the “top surface” (Fig. 4(b)) is such that no steric clash would occur between the second KREPA6 dimer of a tetramer and the Ab2 VH domains bound. The observation that Ab2 disrupts nevertheless the KREPA6 homotetramer formed in solution (Supplementary Fig. 2(a) and (d)) and forms heterotetramers with the KREPA6 dimer formed in the center (Fig. 4) is therefore intriguing. One possible explanation might be that slight conformational changes, not discernable with the 3.4 Å structure of the KREPA6:Ab2 complex, induced by the binding of Ab2 to KREPA6, might be responsible for the disruption of the KREPA6 dimer–dimer interface upon (KREPA6:Ab2)2 heterotetramer formation. Another possibility is that the C-terminal tail of KREPA6, which extends 33 residues beyond the last residue visible in our structures (Fig. 1(a)), is involved in stabilizing the KREPA6 tetramer. Ab2 binding to the “upper KREPA6 dimer” (see Supplementary Fig. 7 for a depiction of a hypothetical KREPA6 tetramer) could possibly interfere with the interactions of the C-terminal tails from the “lower” dimer subunits with the upper dimer thereby breaking up the KREPA6 tetramer into two KREPA6 dimers each bound to Ab2. The latter hypothesis was tested by trypsin treatment of the KREPA6 tetramer since there are no Arg or Lys residues in the N-terminal region of the KREPA6 chain (Fig. 1). Since trypsinolysis created KREPA6 dimers (Fig. 5), it is likely that C-terminal residues are involved in forming KREPA6 tetramers.

4.2. KREPA6 binding by the VHH domains

In the three heterotetrameric structures solved, very different segments of the VH domains are involved in binding KREPA6. The complexes with Nb5 and Nb15 are very similar to each other, with CDR2 and framework 3 as major agents interacting with KREPA6. From an examination of all available VHH domain–antigen structures (Koide, 2009; Wu and Hol, unpublished), it appears that CDR2 is only in 2 out of 32 cases involved in substantial interactions with the protein antigen. These are the α-amylase:AMD9-nanobody (Desmyter et al., 2002) and GFP:nanobody complexes (Kirchhofer et al., 2010). However, in the KREPA6:Nb5 and KREPA6:Nb15 cases, a parallel pair of β-strands is formed between nanobody and protein antigen, a feature absent in the α-amylase:AMD9 and GFP:nanobody structures. The mode of interaction by CDR2 and framework 3 of Nb5 and Nb15 with the protein antigen KREPA6 is therefore unique among VH domain–protein antigen complexes with known structure.

Ab2 uses CDR3 as a major component of the interaction surface with KREPA6. This mode of binding with CDR3 as a prime interaction segment has also been observed in cases of other proteins which were difficult to crystallize, such as the EpsI:EpsJ:nanobody complex (Lam et al., 2009) and the peri-GspD:nanobody complex (Korotkov et al., 2009). The formation of the KREPA6:Nb15 and KREPA6:Nb5 heterodimers is different from all these with a major role for CDR2 and framework 3. Clearly, the modes of interaction of VH domains with protein antigens can be highly variable.

4.3. Crystal growth promotion by VHH domains

There are striking differences in the way VH domains assist in crystal growth. In the crystals of the complexes of KREPA6 with Nb15 and Nb5, the VH domains are making essentially all crystal contacts and no contacts occur between KREPA6 dimers (Supplementary Fig. 6(a)). In the complex of the KREPA6 dimer with Ab2, the arrangement of heterotetramers in the crystals is entirely different. In addition to many contacts between Ab2 subunits (Supplementary Fig. 6(b)), contacts between KREPA6 helices of adjacent KREPA6 subunits from different heterotetramers occur (Fig. 4(c)). We did a similar experiment with two other cases of protein:nanobody structures solved in our laboratory, then it appears that (i) in the EpsI:EpsJ:Nb crystals, the VH domains are forming layers between layers of EpsI:EpsJ heterodimers (Lam et al., 2009) and (ii) in the peri-GspD:Nb crystals, the main effect of crystal growth promotion is the formation of a dimer of VH domains via a pair of anti-parallel β-strands in the center of this Nb dimer. This nanobody dimer leads to a (peri-GspD:Nb)2 heterotetramer in which the flexible linker between the N1 and N2 domains of GspD is kept in a fixed conformation (Korotkov et al., 2009). Clearly, the way in which VH domains can promote crystal formation is highly variable, indicating that nanobodies are versatile tools to crystallize recalcitrant proteins.

4.4. Interacting OB folds in the editosome

KREPA6 is related to the C-terminal OB domains of five other interaction proteins from the editosome (Supplementary Fig. 1) (Worthy et al., 2003). The pair wise sequence identity of KREPA6 with the OB folds of these five proteins ranges from 18.5% for the OB fold from KREPA4 to 39.8% for the OB fold of KREPA3 (Table 2). Since biochemical and yeast two-hybrid studies have indicated that KREPA6 can interact with the OB folds of KREPA1, KREPA2, KREPA3 and KREPA4 (Schnauer et al., 2003, 2010), an interesting question is: what light can the structure of our KREPA6 structures shed on the organization of the OB folds in the core of the editosome? A difficulty is that the stoichiometry of the protein in the editosome is as yet unknown although it is generally assumed that there is one copy of each protein present (Golas et al., 2009; Li et al., 2009). However, given the fact that the KREPA6 subunit is only ~16 kDa there could be multiple KREPA6 subunits in the editosome in view of the uncertainty in the total mass of the particle.
Hence, the core of the editosome might contain a KREPA6 tetramer analogous to the canonical SSB tetramers (Supplementary Fig. 7). Another possibility, which seems more likely than a tetramer given the ease by which VHH domains can dissociate the KREPA6 tetramer into dimers (Figs. 2(a), 3(b), and 4(a), Supplementary Fig. 2), is that a KREPA6 dimer as observed in our crystal structures is present in the core of the editosome. OB folds from the four other interaction proteins interacting with KREPA6 would then contact this dimer in a number of different ways. One possibility is that the KREPA6 dimer utilizes its β-surface to interact with other components. 

Fig. 5. Trypsinolysis converts the KREPA6 tetramer into a dimer. (a) Trypsinolysis of KREPA6. 0.8 mg/ml purified KREPA6 was incubated with trypsin (500:1 w/w) for the indicated times in 20 mM Tris pH 8.0, 300 mM NaCl, 1 mM DTT and 10% glycerol buffer. The result was analyzed by 8–16% SDS–PAGE and visualized by Coomassie staining. Lane 2 shows the undigested KREPA6 protein; lanes 3–5 trypsinized KREPA6. Protein standards are shown in lane 1. Predicted cleavage sites by trypsin are indicated in Fig. 1 (see Fig. 1(a)). (b) Gel filtration chromatography. Gel filtration over a Superdex 200 sizing column was performed on 0.8 mg/ml trypsinized KREPA6 with an N-terminal His-tag (left panel) and without an N-terminal His-tag (right panel) as indicated. 0.8 mg/ml purified KREPA6 was incubated for 1 h with trypsin (500:1 w/w). The buffer was 300 mM NaCl, 20 mM Tris–HCl, pH 7.5, 10% glycerol and 1 mM dithiothreitol. Chromatographic absorbance traces at 280 nm are shown for molecular standards (gray), untreated KREPA6 (black), and trypsinized KREPA6 (red). The elution position and molecular mass of the calibration standards are indicated. The elution volumes of His6-KREPA6 and KREPA6 are nearly identical, as are those of their trypsinized counterparts. These volumes correspond with oligomerization states of tetramers (with a MW of ~60 kDa; labeled “T”) for the untreated proteins and dimers (with a MW of ~30 kDa; labeled “D”) after trypsinolysis. (c) Analysis of gel-filtration fractions by Ni-NTA chromatography and SDS–PAGE. The major peak fractions of KREPA6 proteins were subjected to Ni-NTA His tag pull-down followed by 8–16% SDS–PAGE analysis. Lanes 2–4: full length His6-KREPA6; lanes 5–7: trypsinized His6-KREPA6; lanes 8–10: trypsinized untagged KREPA6 (as a control showing that KREPA6 has little affinity for the Ni-NTA column without a His6-tag). FT: flow-through; W: wash; and E: elute. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
with the β-surface of one or more other OB-folds resulting possibly in a heterotetramer, resembling the canonical homotetramer SSB in overall architecture (Supplementary Fig. 7). Additional contacts of a KREPA6 dimer with OB folds in the editosome might involve loops at the “top” the dimer, in a manner similar to that observed in the KREPA6:Ab2 complex (Fig. 4(a)). Yet another, or additional, possibility is that the C-terminal tails of KREPA6 are involved in interacting with OB folds of other editosome proteins. However, the C-terminal tail of KREPA6 in the editosome could also interact with non-OB fold proteins since in many SSB proteins the function of the C-terminal tail is to bring proteins and multi-protein complexes in proximity of the DNA to which the SSB proteins are attached (Bruck et al., 2002; Genschel et al., 2000; Glover and McHenry, 1998; Handa et al., 2001; Kelman et al., 1998; Savvides et al., 2004; Umezu and Kolodner, 1994; Witte et al., 2003). Finally, heterodimers of the OB folds of KREPA6 and e.g. KREPA3 or KREPA5 may occur in the assembled editosome in view of the high degree of sequence identity of interface residues of the OB folds of KREPA3 and KREPA5 with that of KREPA6 (Table 2), although so far no interactions between KREPA6 and KREPA5 have been reported. Clearly, many possibilities exist and further studies are needed to establish the actual mode of assembly of these important SSB domains in the editosome.

Author Contributions

Protein production and crystallography: Meiting Wu, Young Park, and Stewart Turley. 

Nanobody generation and cloning: Els Pardon and Jan Steyaert. 

Semi-synthetic library generation and screening: Andrew Hayhurst and Young Park. 

Structure analysis: Meiting Wu, Young Park, and Wim Hol. 

Manuscript writing: Meiting Wu, Young Park, Els Pardon, Jan Steyaert, and Wim Hol. 

Project coordination: Wim Hol. 

Acknowledgments

We like to thank Claudia Roach for assistance with molecular biology procedures, Ellen Goldman for help with the NOMAD library, the staff of SSRL (Stanford) for expert support during data collection, and Katleen Willibal for the selection, screening and recloning of the nanobodies. We thank Achim Schnauer, Ken Stuart and Xueming Guo for stimulating discussions. Financial support was provided by the Howard Hughes Medical Institute (HHMI), by grants R01GM077418 and R01GM077418-04S1 from NIH to WCH, and by the Belgian Government under the framework of the Interuniversity Attraction Poles (I.A.P. P6/19).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2010.10.007.

References


