Structure of transportin SR2, a karyopherin involved in human disease, in complex with Ran

Vicky G. Tsirkone, Katrien G. Beutels, Jonas Demeulemeester, Zeger Debyser, Frauke Christ and Sergei V. Strelkov


Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see http://journals.iucr.org/services/authorrights.html

Acta Crystallographica Section F: Structural Biology Communications is a rapid all-electronic journal, which provides a home for short communications on the crystallization and structure of biological macromolecules. Structures determined through structural genomics initiatives or from iterative studies such as those used in the pharmaceutical industry are particularly welcomed. Articles are available online when ready, making publication as fast as possible, and include unlimited free colour illustrations, movies and other enhancements. The editorial process is completely electronic with respect to deposition, submission, refereeing and publication.

Crystallography Journals Online is available from journals.iucr.org
Structure of transportin SR2, a karyopherin involved in human disease, in complex with Ran

Transportin SR2 (TRN-SR2) is a β-type karyopherin responsible for the nuclear import of specific cargoes, including serine/arginine-rich splicing factors. The protein has been implicated in a variety of human diseases, including HIV infection, primary biliary cirrhosis and limb-girdle muscular dystrophy 1F. Towards understanding its molecular mechanism, a 2.9 Å resolution crystal structure of human TRN-SR2 complexed with the small GTPase Ran has been determined. TRN-SR2 is composed of 20 α-helical HEAT repeats forming a solenoid-like fold. The first nine repeats form a ‘cradle’ for the binding of RanGTP, revealing similarities but also differences with respect to the related importin 13 complex.

1. Introduction

Compartmentalization in the eukaryotic cell allows tight regulation of biological functions such as transcription and translation, but it also necessitates active transport of complex macromolecules in and out of the nucleus via the nuclear pore complex (NPC; Hoelz et al., 2011). Such transport relies on soluble factors (importins and exportins) belonging to the karyopherin-β family (Stewart, 2007; Chook & Süel, 2011). The karyopherins shuttle their cargoes between the cytoplasm and the nucleus via a GTP-dependent process which relies on a small GTPase called Ran (Ras-related nuclear protein). Specifically, nuclear import is driven by the gradient of the GTP-bound form of Ran, which is enriched in the nucleus. After binding the cargo in the cytoplasm, the importin docks to the cytoplasmic side of the NPC and is then transferred to the other side, driven by transient interactions with nucleoporins (D’Angelo & Hetzer, 2008). Once in the nucleus, the importin–cargo complex dissociates owing to the competitive binding of RanGTP to the importin, followed by the return of the importin–RanGTP complex to the cytoplasm and GTP hydrolysis (Lee et al., 2005).

Here, we focus on human transportin SR2 (TRN-SR2), a β-type importin (Lai et al., 2000). Both TRN-SR2 and its splice variant TRN-SR1 are encoded by the ttnpo3 gene (Yun et al., 2003). A characteristic feature of TRN-SR2 is its ability to import serine/arginine-rich proteins (SR proteins), in particular splicing factors. Like other family members, TRN-SR2 is predicted to consist of 20 consecutive α-hairpin motifs (two antiparallel α-helices joined by a short linker) known as HEAT repeats (Andrade & Bork, 1995). The stacked HEAT repeats are generally believed to create a structure with considerable plasticity, which provides adaptive binding of different cargoes and regulatory proteins (Cook et al., 2007). While forming a single continuous physical domain, the HEAT-based karyopherins include distinct and sometimes overlapping functional regions. In particular, their N-terminal part binds RanGTP, whereas the C-terminal part is generally responsible for cargo binding (Lai et al., 2000).

Importantly, TRN-SR2 has been implicated in a variety of human diseases, including human immunodeficiency virus (HIV) infection (Brass et al., 2008; Christ et al., 2008; König et al., 2008), primary biliary cirrhosis (PBC; Hirschfield et al., 2010) and limb-girdle muscular dystrophy 1F (LGMD1F; Melià et al., 2013; Torella et al., 2013). The role of TRN-SR2 in the HIV life cycle has been well
studied. Yeast two-hybrid screening initially identified TRN-SR2 as a direct binding partner of HIV integrase (Christ et al., 2008), a finding that was later confirmed by other approaches (Krishnan et al., 2010; Larue et al., 2012). Using fluorescent techniques, it has been demonstrated that TRN-SR2 is directly involved in the nuclear import of the HIV pre-integration complex, leading to the establishment of infection (Christ et al., 2008). Next, hereditary limb-girdle muscular dystrophy 1F was linked to a single nucleotide deletion in the stop codon of the tnpo3 gene. This mutation was shown to result in an extended protein of either 938 (form A) or 1018 (form B) amino acids, compared with the native TRN-SR2 of 923 residues. The mutated protein accumulates in the outer rim of the nuclear membrane, suggesting disturbed nuclear transport (Melila et al., 2013; Torella et al., 2013). Given both the fundamental aspect of the TRN-SR2-dependent nuclear import and the above-mentioned disease context, it is not surprising that this protein has recently gained interest and has been actively studied at the biochemical and cellular levels (Christ et al., 2008; De Houwer et al., 2012; Larue et al., 2012; Taltynov et al., 2013). As the next major step, here we present the atomic structure of human TRN-SR2, which was made possible by resolving its complex with Ran at 2.9 Å resolution using X-ray crystallography.

2. Materials and methods

Both full-length human TRN-SR2 and full-length human RAN mutant Q69L (which is practically incapable of GTP hydrolysis) were overexpressed in *Escherichia coli* BL21 (DE3) Rosetta pLysS cells using the pETHSUMO vector (Wecks et al., 2007). The expressed products were fusions containing an N-terminal 6×His tag, the small ubiquitin-like modifier (SUMO) protein, a SUMO hydrolyase cleavage site and the protein sequence of interest. After digesting the vector with HindIII and KpnI, the sequences encoding *Homo sapiens* TRN-SR2 (residues 3–923) and RanQ69L (residues 1–216) were inserted using the In-Fusion system (Clontech).

Bacterial cultures were grown at 24°C until an OD600nm of 4.0 was reached, and then for a further 24 h at 18°C. The cells expressing TRN-SR2 and RanQ69L were pelleted by centrifugation and resuspended in buffer A (20 mM Tris pH 7.5, 200 mM NaCl, 15 mM imidazole, 5 mM β-mercaptoethanol) and buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 4 mM MgCl2, 15 mM imidazole, 10% glycerol, 2.5 mM β-mercaptoethanol), respectively. After the addition of 5 µl Benzonase per 100 ml, the cells were sonicated on ice for 2 min with 20% amplitude and 30/10 s on/off pulses using a Branson Digital Sonifier. The procedure was repeated twice with 15 min in between. The lysates were cleared at 18 000g for 30 min at 4°C and were then applied onto an Ni2+–NTA (GE Healthcare) column in buffer A, and subsequently into a Ni2+–NTA column, whereby the flowthrough containing the cleaved protein was collected while the uncleaved protein and the protease, both containing the His tag, were retained on the column. For TRN-SR2, the buffer was subsequently exchanged to 20 mM Tris pH 7.5, 50 mM NaCl, 5 mM DTT using an Ultra-15 ultrafiltration device (Amicon), which was followed by ion-exchange chromatography on a 5 ml HiTrap Q HP column (GE Healthcare) using a linear gradient of 50 mM to 1 M NaCl. To load Ran with GTP, the protein was incubated for 30 min at room temperature with a fivefold molar excess of GTP and 20 mM EDTA. The loading was stopped by the addition of 40 mM MgCl2 followed by dialysis against 50 mM Tris, 150 mM NaCl, 5 mM DTT pH 7.5. For complex formation, TRN-SR2 and RanQ69L-GTP were mixed in a 1:2 molar ratio and incubated on ice for 2 h, followed by size-exclusion chromatography on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare) in 20 mM Tris pH 7.5, 50 mM NaCl, 5 mM DTT. Prior to crystallization, the complex was concentrated to 9.8 mg ml⁻¹.

Initial screening for crystallization of the TRN-SR2–RanQ69L-GTP complex was performed using commercial screens and a Mosquito crystallization robot (TTP Labtech). Small needles could be obtained in 400 nl sitting drops using The Protein Complex Suite (Qiagen) condition 45 [100 mM MES pH 6.5, 5% (w/v) MPD, 15% (w/v) PEG 6000] at 20°C. Subsequently, larger crystals could be obtained in 3 µl hanging drops equilibrated against an equivalent reservoir solution but with a slightly higher pH (6.7). After several rounds of microseeding, the best crystal reached dimensions of 150 × 30 × 20 µm in 10 d. X-ray diffraction data were collected on the PROXIMA 2A beamline (Synchrontron Soleil, Saint-Aubin, France) using a microfocus 10 × 5 µm beam with 0.979 Å wavelength. The crystals were briefly soaked in mother liquor supplemented with 30% MPD and flash-cooled in liquid nitrogen. To limit radiation damage to the thin crystal, five different spots on the crystal were exposed; each was used to collect a 20° rotation wedge. Data were processed with XDS (Kabsch, 2010) and scaled with AIMLESS (Evans, 2006). The crystals contained one protein complex per asymmetric unit (Table 1).

Initially, we failed to phase the data by molecular replacement using as a search model our homology-modelled structure of the TRN-SR2–RanQ69L-GTP complex (Taltynov et al., 2013), as well as any available experimental structures of related importins or their complexes with Ran, including the complex of a closely related human importin 13 (Imp13) with yeast RanGTP (PDB entry 2xtj; Bon et al., 2010). In retrospect, this difficulty could be explained by significant differences in the overall conformation. However, it was possible to obtain the correct molecular-replacement solution, albeit with a marginal contrast, by searching with only the RanGTP from the latter structure using MOLREP (Vagin & Teplyakov, 2010). This can be rationalized by the remarkably high conservation between human and yeast Ran (90% sequence identity), even though this

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crytallographic data-collection and refinement statistics.</strong></td>
</tr>
<tr>
<td>     </td>
</tr>
<tr>
<td><strong>Space group</strong>  </td>
</tr>
<tr>
<td><strong>Unit-cell parameters (Å)</strong></td>
</tr>
<tr>
<td><strong>Resolution range (Å)</strong></td>
</tr>
<tr>
<td><strong>No. of unique reflections</strong></td>
</tr>
<tr>
<td><strong>Multiplicity</strong></td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
</tr>
<tr>
<td><strong>CC₁/²</strong>†‡</td>
</tr>
<tr>
<td><strong>Rₜₚ</strong>‡</td>
</tr>
<tr>
<td><strong>Rmerge</strong>‡</td>
</tr>
<tr>
<td><strong>Wilson B factor (Å²)</strong></td>
</tr>
</tbody>
</table>

† Half-set correlation of intensities  ‡ As output by AIMLESS (Evans, 2006). § As calculated using a ‘free’ set of 10% reflections (%).
protein makes up only 15% of the total complex mass. With the Ran molecule in place, a further MOLREP search could position a segment of the Imp13 structure from the complex containing HEAT repeats 1–5. With these two elements in place, another MOLREP run could correctly position a segment containing repeats 6–9. Such an assembled partial structure yielded a crystallographic free R factor of 0.54, which could be improved to 0.51 by automated refinement in REFMAC (Murshudov et al., 2011), but only if the ‘jelly-body’ restraints were enabled. At this point, HEAT repeats 10–13 could be approximately placed into the electron-density map, followed by further refinement and placement of additional repeats until the model was complete. After multiple cycles of manual rebuilding using Coot (Emsley et al., 2010) and refinement using PHENIX (Adams et al., 2010), a final model with good quality was obtained (Table 1). Atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.pdb.org) under accession code 4ol0. Figures were prepared with PyMOL (http://www.pymol.org/). The protein complexes were analyzed with the PISA algorithm (Krissinel & Henrick, 2007) using an online server (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html).

3. Results and discussion
Previously, we have studied the detailed biochemical properties of human TRN-SR2, and in particular characterized its complex with the catalytically inactive Q69L mutant of human Ran (Taltynov et al., 2013). For the current work, we have changed the overexpression system for both proteins, which yielded highly pure tag-free samples. As a result, after forming the TRN-SR2–RanQ69L-GTP complex, single crystals could be grown. Despite their small size, the use of a microfocus synchrotron beam and an optimized data-collection strategy allowed us to measure a complete 2.9 Å diffraction data set (Table 1). Phasing of the data only became possible via a carefully chosen molecular-replacement protocol, whereby the RanGTP molecule and a number of the N-terminal HEAT repeats of TRN-SR2 could be positioned. This was followed by gradual building of the remaining structure. Of note, we were also able to obtain large (>100 μm in each direction) crystals of human TRN-SR2 alone, but despite extensive optimization efforts these crystals diffracted X-rays to only about 6 Å resolution (O. Taltynov, V. O. Tsirkone and S. V. Strelokov, unpublished results).

The asymmetric unit of the TRN-SR2–RanQ69L-GTP crystals contains a single copy of the complex. The entire TRN-SR2 molecule is well defined in the electron-density map, with the exception of three linkers located between HEAT repeats 5 and 6 (residues 216–220), the two helices of HEAT 14 (residues 598–605) and the two helices of HEAT 20 (residues 884–888). The C-terminal section (residues 180–216) of RanQ69L is also not visible, which is in line with earlier predictions that this part is disordered (Vetter et al., 1999; Nilsson et al., 2002). Previously, C-terminally truncated Ran had been used for the crystallization of the Imp13–Ran complex (Bono et al., 2010).

Our crystal structure confirms that TRN-SR2 is a solenoid-type protein consisting of 20 HEAT repeats, corroborating this hallmark feature of the karyopherin-β family (Andrade & Bork, 1995). Each of the repeats is principally composed of two antiparallel α-helices connected through a loop, resulting in an ‘α-hairpin’. The first α-helix of each repeat (called helix A) is always located on the outside of the solenoid, while the second α-helix (B) is on the inside (Fig. 1). The axes of helices A and B, although principally antiparallel, usually form a small crossing angle (about −15°, left-handed). Exceptionally, the C-terminal HEAT repeat 20 capping the open edge consists of three helices.

Neighbouring α-hairpins corresponding to HEAT repeats 2–9 are roughly parallel to each other (while the very first repeat is significantly rotated). Consequently, the overall conformation of the N-terminal part of TRN-SR2 corresponds to a sector of a straight solenoid (Fig. 1a). The same is valid for the C-terminal part (HEAT repeats 10–20). However, there is a pronounced rotation of HEAT 10 with respect to HEAT 9, which are connected by a particularly short linker. As a result, the overall shape of TRN-SR2 is that of a right-handed screw rather than of a straight solenoid (Fig. 1b).

The core of the Ran molecule is an SH3 domain fold consisting of six β-strands. This core is supplemented by three α-helices and two surface loops referred to as Switch I (residues 37–45) and Switch II.

Figure 1

(a, b) Overall view of the TRN-SR2–RanQ69L-GTP complex in two perpendicular orientations. Each HEAT of TRN-SR2 is shown in a different colour, with α-helices as cylinders. Ran is represented by a sand-coloured surface. Both helices A and B of HEAT 1 are labelled. The three helices A, B and C of HEAT 20 are also indicated. For all other repeats, only the outer helix A is labelled for clarity.
Both loops are known to undergo extensive conformational changes upon GTP binding (Stewart et al., 1998).

The binding of Ran to TRN-SR2 requires GTP and results in a stable complex with an apparent $K_d$ of $4.7 \pm 1.2 \text{nM}$, as we have shown in vitro using the AlphaScreen protein–protein interaction assay (Taltynov et al., 2013). Indeed, our crystal structure reveals that formation of the complex buries a very considerable part (22.9%, 2081 Å$^2$) of the total surface area of RanGTP. The complex is stabilized in a synergistic manner by both hydrophobic interactions and polar contacts (Fig. 3). The latter interactions are represented by specific hydrogen bonds (to side groups or the main chain) and salt bridges across the interface (Figs. 2 and 4). The N-terminal part of TRN-SR2 clearly presents the specific site of RanGTP binding. Indeed, HEAT repeats 1–9 make a ‘cradle’ roughly matching the size of the globular Ran protein, albeit slightly larger. Correspondingly, the binding of Ran mostly involves the $B$ (inner) helices of HEAT repeats 1–4 and 7–9, with some additional contacts made by repeats 5 and 6 (Figs. 2 and 3). On the Ran side, the exposed Switch II loop is involved in important interactions with HEAT 1 and HEAT 2, while Switch I makes further contacts with both HEAT 1 and HEAT 18 (Fig. 2). In addition, the N-terminal tips of helices $B$ from repeats 15, 17 and 18 also make some contacts with the bound Ran (Fig. 3). It is likely that these latter interactions are not essential for RanGTP binding but rather occur simply because of the curved shape of the TRN-SR2 molecule. At the same time, the expected site of cargo binding is located in the C-terminal part of TRN-SR2 (Lai et al., 2014).

**Figure 2**
Molecular detail of the interaction between TRN-SR2 and RanGTP. (a) Contacts of the main Ran-binding domain (HEAT repeats 1–9, intense blue). TRN-SR2 residues forming salt bridges and hydrogen bonds to Ran are shown in green with side chains as sticks. The Ran molecule is shown as a yellow ribbon, with the Switch I and Switch II regions highlighted in orange. The bound GTP is shown in red. (b) A close-up view of some interacting residues of TRN (green) and Ran (yellow) and the corresponding weighted $2|F_o| - |F_c|$ electron-density map at the 1$\sigma$ level, presented in stereo. The salt bridges Glu152/Glu153(TRN)–Arg110(Ran) and Asp247/TRN)–Lys141(Ran) are shown with dashed lines.
Since cargo release inside the nucleus is coupled to RanGTP binding, there must also be an interplay between the two processes at the structural level. Such interplay is facilitated by the solenoid shape of the TRN-SR2 molecule, since the two binding sites are located in proximity and may be sterically overlapping. Furthermore, the HEAT-based structure is clearly prone to conformational plasticity, as observed before for various crystallographic complexes of other importins (Lee et al., 2006; Forwood et al., 2010). This plasticity is likely to play an important role, whereby a particular binding partner could come in contact with just the N-terminal part or the C-terminal part, or with both of them as in the crystal structure presented here.

Finally, we have compared the structure of the TRN-SR2–RanGTP complex with the corresponding complex of Imp13, a β-karyopherin with 22.6% amino-acid identity and 38% similarity to...
The latter complex includes Ran from the yeast *Saccharomyces cerevisiae* rather than human Ran (90% sequence identity). In line with the considerable homology between TRN-SR2 and Imp13, the overall HEAT-based structure of both proteins is similar (Fig. 5). However, the Imp13 solenoid wraps around the Ran globule in a somewhat less tight way than TRN-SR2, so that HEAT repeats 7 and 8 of Imp13 make only a few contacts with Ran, in contrast to our TRN-SR2 complex (Fig. 3). Moreover, there is some conservation but also considerable differences in the specific interactions (hydrogen bonding and salt bridges) that RanGTP makes with either of the two proteins. In particular, superposition of the two complexes reveals that as few as only four salt bridges (involving three Ran residues Asp77, Arg110 and Arg166) and three hydrogen bonds (involving Ran residues 47, 78 and 79) are structurally preserved (Fig. 4). Of note, the Ran-binding domain of these two importins (HEAT repeats 1–9) is somewhat more conserved (amino-acid sequence identity 26.0%) than the rest of the molecule (sequence identity of 20.1%).

As this paper was being written up, Maertens et al. (2014) reported the crystal structures of human TRN-SR2 on its own and in two complexes, including the RanGDP complex (PDB entry 4c0q). Compared with our structure, the latter structure was determined in a distinct space group $P1$ with two complexes per asymmetric unit at 3.42 Å resolution. Together with a better resolution, our structure additionally reveals several surface loops (residues 110–194, 354–362, 625–633 and 672–676). The overall conformation of TRN-SR2 in the two crystal structures is very similar, especially for the N-terminal part, which shows an identical mode of RanGTP binding. For a detailed comparison of the two structures, we have aligned them by HEAT repeats 1–15. Such a superposition yields an r.m.s. deviation of only 1.2 Å for all Cα/C11 atoms of HEATs 1–15. However, the C-terminal parts of the TRN-SR2 molecule (repeats 16–20) in the two structures gradually deviate more significantly. This appears to be owing to a different packing in the two crystals and to intrinsic flexibility of the TRN-SR2 solenoid. The largest differences in the Cα positions (up to 7 Å) are within the last two HEATs (19 and 20).

Determination of the atomic structure of TRN-SR2 will provide a basis for a fundamentally better understanding of its functioning at the molecular level. As outlined in §1, besides its involvement in other human diseases, TRN-SR2 plays a key role in HIV infection. Indeed, the nuclear import of the HIV pre-integration complex can be seen as the bottleneck for HIV infection, as it is an essential prerequisite for the integration of the viral genome into the host cell, which ultimately leads to persistent infection. We believe that targeting the interaction between TRN-SR2 and HIV integrase may be a promising avenue for new drug development, complementing existing therapies.

We thank Stephen Weeks for helpful advice on cloning and protein expression, Stephanie De Houwer for discussions and Maria-Despoina Charavgi for valuable advice on crystallography. This work was supported by the Research Foundation Flanders (FWO) grant G0665.12 (to SVS and FC). Research at the Laboratory of Molecular structural communications
Virology and Gene Therapy is also funded by the Belgian IAP BelVir.

References


Virology and Gene Therapy is also funded by the Belgian IAP BelVir.