Structural analysis of Dis3l2, an exosome-independent exonuclease from *Schizosaccharomyces pombe*

Hui Lv,[a,b] Yuwei Zhu,[a,b] Yu Qiu,[a,b] Liwen Niu,[a,b] Maikun Teng[a,b]* and Xu Li[a,b]*

[a]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Science, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China, and [b]Key Laboratory of Structural Biology, Chinese Academy of Science, Hefei, Anhui 230026, People's Republic of China. *Correspondence e-mail: mkteng@ustc.edu.cn, sachem@mail.ustc.edu.cn

After deadenylation and decapping, cytoplasmic mRNA can be digested in two opposite directions: in the 5′–3′ direction by Xrn1 or in the 3′–5′ direction by the exosome complex. Recently, a novel 3′–5′ RNA-decay pathway involving Dis3l2 has been described that differs from degradation by Xrn1 and the exosome. The product of the *Schizosaccharomyces pombe* gene SPAC2C4.07c was identified as a homologue of human Dis3l2. In this work, the 2.8 Å resolution X-ray crystal structure of *S. pombe* Dis3l2 (*Sp*Dis3l2) is reported, the conformation of which is obviously different from that in the homologous mouse Dis3l2–RNA complex. Fluorescence polarization assay experiments showed that RNB and S1 are the primary RNA-binding domains and that the CSDs (CSD1 and CSD2) play an indispensable role in the RNA-binding process of *Sp*Dis3l2. Taking the structure comparison and mutagenic experiments together, it can be inferred that the RNA-recognition pattern of *Sp*Dis3l2 resembles that of its mouse homologue rather than that of the *Escherichia coli* RNase II–RNA complex. Furthermore, a drastic conformation change could occur following the binding of the RNA substrate to *Sp*Dis3l2.

1. Introduction

Exonucleases are large enzymes that participate in RNA processing, which plays important roles in many life processes. Exonucleases can alter RNA levels by digesting aberrant RNA molecules, thereby contributing to RNA surveillance. Additionally, some precursors of RNA require trimming by exonucleases during their maturation process (Ibrahim *et al.*, 2008; Houseley & Tollervey, 2009; Schoenberg & Maquat, 2012).

The exosome is the most significant complex involved in the cleavage of RNA chains; the chains are cleaved repeatedly from the 3′ to 5′ ends one nucleotide at a time (Schilders *et al.*, 2006; Hernández *et al.*, 2006; Mitchell *et al.*, 1997). Recently, a great deal of effort has been focused on studying the structural and functional details of the exosome. The common cores of eukaryotic and prokaryotic exosomes are conserved during evolution. In eukaryotic species such as yeast, the core of the exosome consists of Exo-9, which is composed of nine subunits. Six subunits of Exo-9 assemble into a hexameric barrel containing an RNase PH domain that has structural homology to the bacterial phosphorolytic nuclease. The other three subunits, which contain S1 and KH domains, form a ring lying on the top face of the barrel. The six RNase PH-like subunits
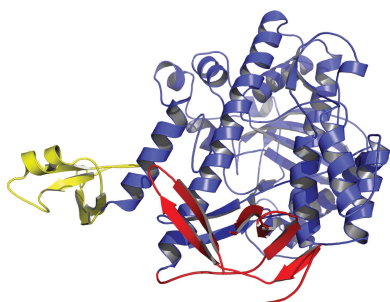
are inactive in yeast and humans, where they form a channel to transport the RNA substrate; in contrast, their counterparts in archaeal exosomes still possess catalytic activity (Liu *et al.*, 2006; Dziembowski *et al.*, 2007). In eukaryotic species, the tenth subunit, Dis3/Rrp44, which is a member of the RNase II-like exonuclease family, interacts with the components of the barrel body to form the Exo-10 catalytic complex (Lorentzen *et al.*, 2005; Büttner *et al.*, 2005; Liu *et al.*, 2006, Vanacova & Stefl, 2007).

Recent findings show that in addition to Dis3, Dis3l and Dis3l2 are also RNase II/R homologues in humans. Dis3l shares more similarities with Dis3 (which requires interaction with the exosome ring to become a functional unit) and contains a PIN domain at its N-terminus (Staals *et al.*, 2010; Tomecki *et al.*, 2010; Astuti *et al.*, 2012). Dis3l2 is composed of two cold-shock domains at its amino-terminus (CSD1 and CSD2) followed by a catalytic RNB domain and a carboxy-terminal S1 domain. Similar to the bacterial RNase II/R, Dis3l2 lacks the PIN domain that exists at the N-terminus of Rrp44 (Fig. 1*a*). Dis3l2 localizes in the cytoplasm and is involved in genetic interactions with components that function in the cytoplasmic mRNA-degradation pathway (Malecki *et al.*, 2013). For example, Dis3l2 can degrade uridylated pre-let-7 to block the expression of let-7 microRNAs in mouse embryonic stem cells (Chang *et al.*, 2013). Human Dis3l2 mutations can cause Perlman syndrome, which results in overgrowth and is related to Wilm's tumour. Moreover, the

variability in chromosome number and mitotic errors in HeLa cells are increased after *dis3l2* knockdown (Astuti *et al.*, 2012). Other studies have indicated that Dis3l2 preferentially degrades uridylated RNAs, while adenylated RNAs inhibit its catalytic activity (Malecki *et al.*, 2013); moreover, Dis3 can degrade both poly(A) and poly(U) RNA substrates with a similar preference (Lorentzen *et al.*, 2008). These significant functional alterations indicate that structural differences exist between Dis3 and Dis3l2. Recently, the structure of mouse Dis3l2 in complex with an oligoU RNA substrate has been solved (PDB entry 4pmw; Faehnle *et al.*, 2014) and was shown to adopt a conformation similar to those of the Rrp44–RNA complex (PDB entry 2vnu; Lorentzen *et al.*, 2008) and the RNase II–RNA complex (PDB entry 2ix1; Frazão *et al.*, 2006): the three OB-fold domains (CSD1, CSD2 and S1) lie on the top face of the catalytic RNB domain to form the entrance to the RNA path, and the conserved RNA-binding residues in the RNB domain contribute to the formation of the catalytic region. However, there are also differences in the RNA path among the three homologous protein–RNA complexes. For example, in the mouse Dis3l2–RNA complex the three RNA-binding domains adopt different orientations and form a funnel that defines a novel RNA-recognition path that is more open and straight (Faehnle *et al.*, 2014).

The product of the *Schizosaccharomyces pombe* gene SPAC2C4.07c has also been identified as a member of the conserved Dis3l2 family of exonucleases. Considering that budding yeast does not possess a Dis3l2 homologue, fission yeast seems to share more conservation in RNA metabolism with higher eukaryotes than with budding yeast (Malecki *et al.*, 2013). *S. pombe* Dis3l2 (*Sp*Dis3l2) has been shown to be unable to interact with the exosome complex using pulldown experiments. Genetic interactions between the *Sp*Dis3l2 gene and the components of the mRNA-degradation pathway have been observed using a double-deletion assay. Furthermore, deletion of the *Sp*Dis3l2 gene can result in the accumulation of related transcripts (especially uridylated RNAs) and a slower rate of degradation (Malecki *et al.*, 2013). The unique location of *Sp*Dis3l2 in combination with its physical and genetic interactions may shed light on its novel role as an essential factor in an alternative RNA-degradation pathway; this pathway exists in addition to the exosome and Xrn1 pathways observed in most other eukaryotes. To elucidate the RNA-recognition mechanism of *Sp*Dis3l2, we solved the structure of *Sp*Dis3l2 in its RNA-free form at 2.8 Å resolution. Instead of forming the entrance of the RNA channel with the S1 domain, as observed in the mouse Dis3l2–RNA complex, CSD2 is oriented on the side of the catalytic RNB domain. The results of a fluorescence polarization assay (FPA) indicate that the RNB and S1 domains contribute the major binding affinity for the RNA substrate; the CSD region also participates in the interaction process. Combining the structural comparisons with mutagenic experiments, we speculate that *Sp*Dis3l2 could adopt an RNA-recognition pattern that resembles that of its mouse homologue but not that of RNase II. Furthermore, the binding of an RNA substrate to *Sp*Dis3l2 could result in a drastic conformational change.
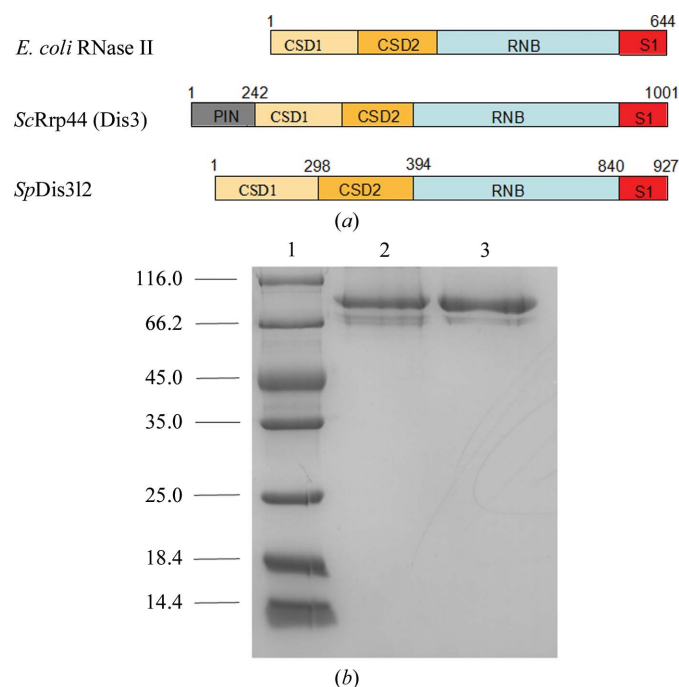

Figure 1
The domain organization and sequence conservation of Dis3l2s. (*a*) Schematic representation of the domain organization of *E. coli* RNase II, *Sc*Rrp44 and *Sp*Dis3l2. All three nucleases contain three OB-fold domains (CSD1, CSD2 and S1) and a catalytic RNB domain. Rrp44 has an N-terminal PIN domain that does not exist in RNase II/R and Dis3l2. (*b*) SDS–PAGE analysis of purified Dis3l (residues 170–927; lane 3) and the washed *Sp*Dis3l2 crystals (lane 2). Lane 1 contains molecular-weight marker (labelled in kDa).

2. Materials and methods

2.1. Construction, expression and purification

The full length and several truncated fragments of the SpDis3l2 gene were amplified from fission yeast genomic DNA by PCR using PrimeSTAR HS DNA polymerase (Takara). The PCR fragments were inserted into the modified expression vector pET-28a(+) (Novagen) using the NdeI/XhoI restriction sites to generate recombinant plasmids containing an N-terminal His tag for expression in *Escherichia coli*. The plasmids were then transformed into *E. coli* BL21 (DE3) cells (Novagen). Expression of all of the recombinant proteins was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the cell growth reached an $OD_{600\ nm}$ of 0.8–1.0. After 20 h of incubation at 289 K, the cells were collected by centrifugation and lysed by sonication in lysis buffer consisting of 50 mM Tris–HCl pH 8.0, 500 mM NaCl, 5% glycerol; the same buffer was used in Ni^{2+}–nitrilotriacetate (Ni–NTA) affinity resin chromatography (Qiagen). The proteins were eluted with 500 mM imidazole followed by cation-exchange chromatography using a Q Sepharose Fast Flow column and size-exclusion chromatography using a Superdex 200 column (GE Healthcare). The final proteins were concentrated to 8 mg ml^{-1} in buffer consisting of 20 mM Tris–HCl pH 8.0, 50 mM NaCl, 5% glycerol for crystallization. Using the methionine-biosynthesis inhibition method, we purified a selenomethionine (SeMet) derivative of SpDis3l2 (residues 170–927). Mutants of SpDis3l2 were obtained using the Mutanbest kit (Takara). The SeMet-derivatized and mutated proteins were purified using the same protocol as described for native SpDis3l2.

2.2. Crystallization, data collection and structure determination

For crystallization, we used the hanging-drop vapour-diffusion method at 289 K. 1 µl protein solution and 1 µl reservoir solution were mixed together. Crystals of both native and SeMet-derivatized SpDis3l2 were grown for approximately 6–10 d under the same conditions consisting of 20% PEG 4000, 0.15 M ammonium sulfate, 0.1 M HEPES pH 7.0. For data collection, all crystals were soaked in a cryoprotectant solution consisting of the respective reservoir solution supplemented with 20%(v/v) glycerol and then flash-cooled in liquid nitrogen. All X-ray diffraction data were collected on beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF) at 100 K. Both the native and the SeMet-derivative data were processed and scaled with *HKL*-2000 and programs from the *CCP*4 package (Winn *et al.*, 2011). The structure of SeMet-derivatized SpDis3l2 was determined by the single-wavelength anomalous dispersion (SAD) phasing technique using the *AutoSol* program as implemented in *PHENIX* (Adams *et al.*, 2010). *AutoSol* found solutions for the RNB domain, and the initial model was built automatically using *AutoBuild* in *PHENIX*. CSD2 of SpDis3l2 was placed manually into the initial map and the *AutoBuild* routine was repeated. The structure was refined using automatically determined NCS restraints, occupancies and isotropic *B*-factor

Table 1
Data-collection and refinement statistics for SeMet-derivatized SpDis3l2.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 116.9$, $b = 54.2$, $c = 127.2$, $\alpha = 90$, $\beta = 107.1$, $\gamma = 90$
Wavelength (Å)	0.9791
Resolution limits (Å)	50.00–2.80 (2.90–2.80)
No. of unique reflections	37868
Completeness (%)	99.50 (97.37)
Multiplicity	3.8 (3.7)
R_{merge}† (%)	11.1 (44.8)
Mean $I/\sigma(I)$	20.3 (5.4)
Refinement statistics	
Resolution limits (Å)	40.53–2.80 (2.90–2.80)
R_{work}‡/R_{free}§ (%)	26.85/29.36
R.m.s.d., bond lengths (Å)	0.004
R.m.s.d., angles (°)	1.005
B factor (Å^2)	51.96
No. of non-H protein atoms	8291
Ramachandran plot (%)	
Most favoured regions	82.2
Additional allowed regions	17.6
Generously allowed regions	0
Disallowed regions	0.2
PDB code	4ro1

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all i observations of reflection hkl. ‡ $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. § R_{free} was calculated in the same way as R_{work} using a randomly selected 5% of the reflections that were omitted from refinement.

refinement in *PHENIX* (Adams *et al.*, 2010). The final model was validated with *MolProbity* (Chen *et al.*, 2010). All of the structures shown in the figures were prepared with *PyMOL*.

2.3. Fluorescence polarization assays

Fluorescence polarization assays were performed in buffer consisting of 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 5% glycerol at 298 K using a SpectraMax M5 microplate reader system. The wavelengths of fluorescence excitation and emission were 486 and 525 nm, respectively. Each well of a 384-well plate contained 100 nM of a fluorescently labelled (5′-FAM) RNA probe and different amounts of the SpDis3l2 mutant proteins with a final volume of 80 µl. DNA-free controls were included for each assay. The fluorescence polarization P (in mP units) was calculated as

$$P = (I_{|} - I_{\perp})/(I_{|} + I_{\perp}). \tag{1}$$

The fluorescence polarization change ΔP (in mP units) was fitted to

$$\Delta P = \Delta P_{max}[\text{protein}]/(K_d + [\text{protein}]). \tag{2}$$

The binding curves were fitted according to a one-site binding model using the *Origin* software.

3. Results and discussion

3.1. Overall structure of SpDis3l2

We expressed and purified full-length SpDis3l2, but obtained no crystals. Similar situations have previously beeen

reported for the crystallization of *Saccharomyces cerevisiae* Rrp44 (*Sc*Rrp44) and mouse Dis3l2 (Lorentzen *et al.*, 2008; Faehnle *et al.*, 2014). Finally, truncated *Sp*Dis3l2 including residues 170–927 was crystallized and the crystals diffracted to 2.8 Å resolution. The molecular-replacement method failed to solve this structure; therefore, the single-wavelength anomalous dispersion (SAD) technique was used to solve the structure of *Sp*Dis3l2. The crystallographic statistics are summarized in Table 1.

The *Sp*Dis3l2 crystals belonged to space group $P2_1$. There are two molecules of *Sp*Dis3l2 in each asymmetric unit. Superposition of the two molecules shows little structural variation [root-mean-square deviation (r.m.s.d.) of 0.41 Å; Fig. 2*a*]. Hence, we discuss only one of the two molecules in the asymmetric unit in the following.

The solved *Sp*Dis3l2 structure contains the CSD2, RNB and S1 domains. Both the CSD2 and the S1 domains are located on the side face of the RNB domain (Fig. 2*b*). Because of the poor electron density, the CSD2 domain contains only four β-strands, which behave similarly to the antiparallel β-strands organized into a β-barrel motif displayed in characteristic OB-folds (residues 330–394; Figs. 3*b* and 4*c*). The RNB domain is the most conserved domain, and the characteristic RNase II-like catalytic core (residues 440–481) is surrounded by several α-helices (Fig. 2*b*). The residues at the active site that participate in binding to the magnesium ion resemble those reported for *Sc*Rrp44 and human Dis3l2 (ARDLDDA; residues 457–463 in *Sp*Dis3l2). The residues that are involved in binding to the RNA are also conserved (Asn658, Arg677, Arg733, His764, Arg770 and Arg771 in *Sp*Dis3l2; Fig. 3). The
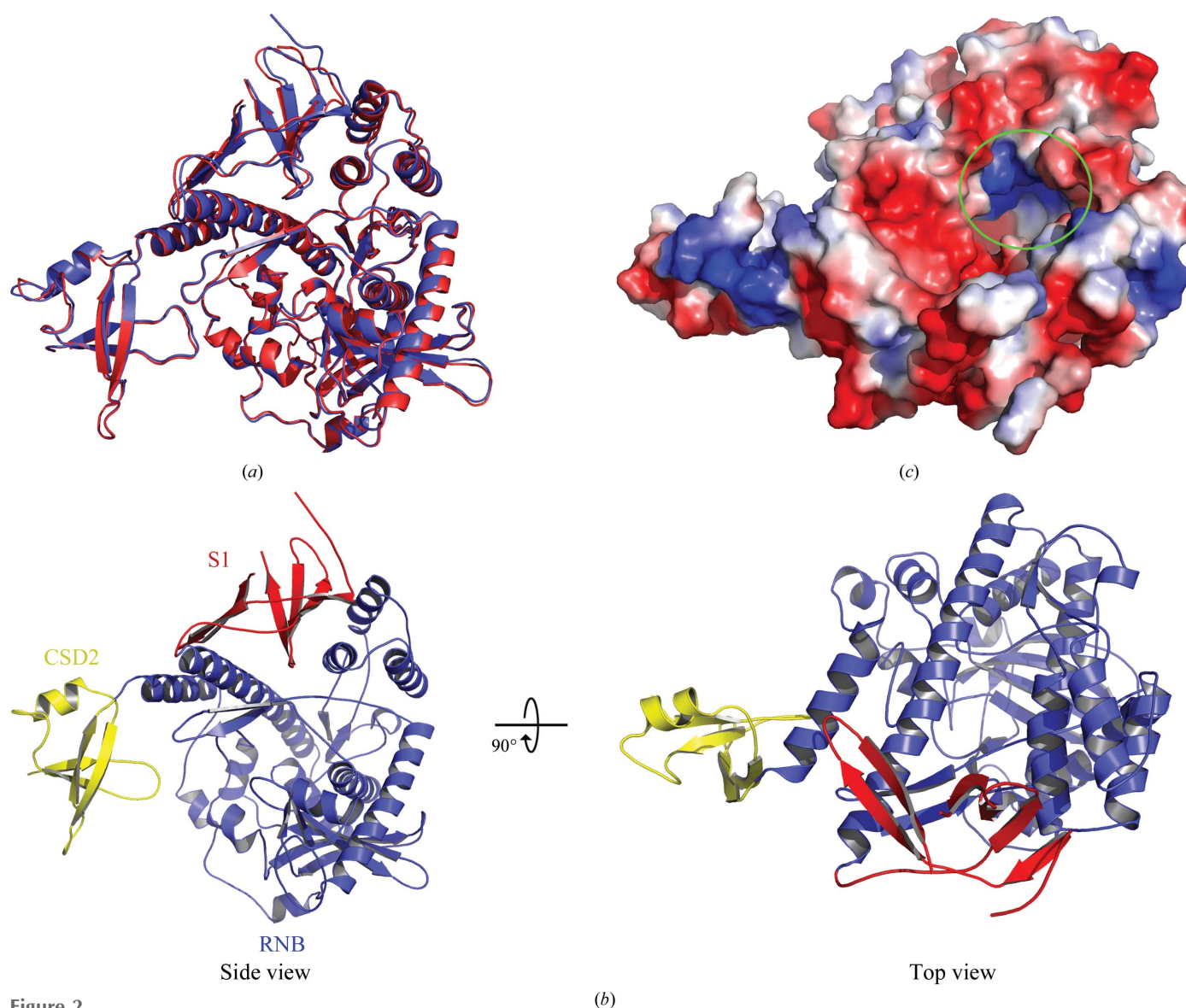

Figure 2
The overall structure of *Sp*Dis3l2. (*a*) Superposition of the two molecules of Dis3l2 in one asymmetric unit. (*b*) Overall structure of *Sp*Dis3l2. The structure is presented in two orientations related by a 90° rotation around the vertical axis. CSD2 flanks one side of the RNB domain. CSD2, RNB and S1 are coloured yellow, blue and red, respectively. Owing to the poor electron density, CSD1 and part of S1 were not modelled in the final structure. (*c*) Top view of the electrostatic surface potential of *Sp*Dis3l2. White, blue and red indicate neutral, positive and negative surfaces, respectively. The positively charged electrostatic surface within the RNB domain is circled in green. (*a*), (*b*) (left) and (*c*) are shown in the same orientation.
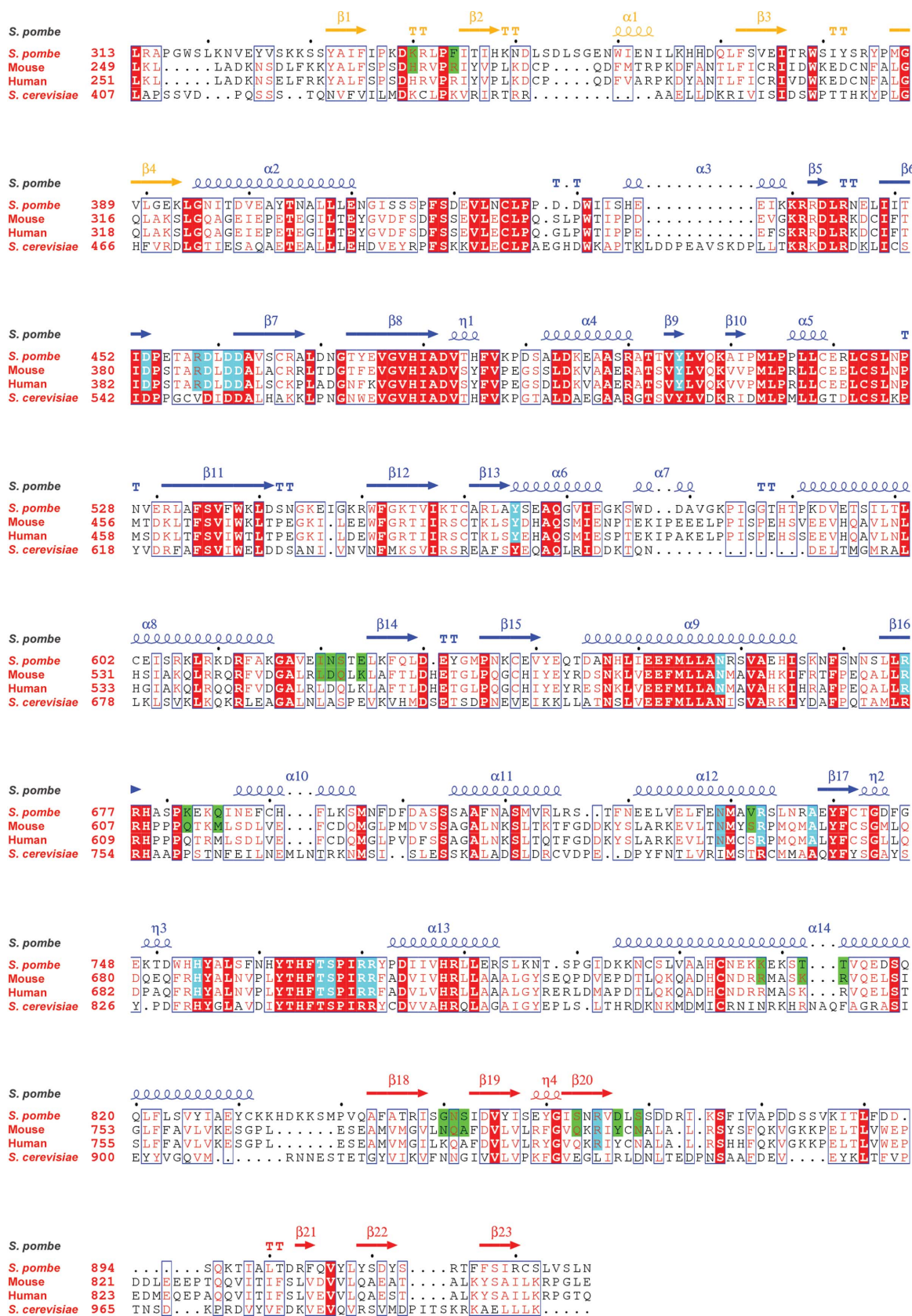
Figure 3
Sequence alignment of Dis3l2s and the secondary structure of *Sp*Dis3l2. A sequence alignment of the Dis3l2 homologues from fission yeast, mouse and human as well as budding yeast Rrp44 is shown. The numbered secondary-structural elements of *Sp*Dis3l2 are included on top of the corresponding sequences. CSD2, RNB and S1 are coloured yellow, blue and red, respectively. RNA-binding residues of mouse Dis3l2 and the corresponding residues of the Dis3l2 homologues are coloured green (conserved between fission yeast and mouse) and cyan (conserved in all three species).

C-terminal S1 domain is composed of antiparallel β-strands that show OB-fold characteristics. The three β-strands (β18, β19 and β20) align well with their counterparts in mouse Dis3l2, and the end of Dis3l2 features two antiparallel β-strands (β21 and β22; Figs. 2b and 4d).

SDS–PAGE analysis of the washed crystals and the purified $SpDis3l2^{170–927}$ revealed that no proteolysis occurred during the crystallization process (Fig. 1b). Therefore, the electron density of the missing N-terminal region (residues 170–329) and the partial S1 region (residues 867–894 and 924–927) in the SpDis3l2 structure is most likely to be owing to structural flexibility issues rather than peptide degradation. The arrangement of the secondary-structure elements is shown in Fig. 3.

3.2. The different domain orientations of RNA-free SpDis3l2 and the mouse Dis3l2–RNA complex

Three domains (RNB, CSD2 and S1) were detected in the RNA-free SpDis3l2 structure. Sructural comparison revealed that the above three domains resembled the corresponding domains in the mouse Dis3l2–RNA complex, with r.m.s.d.s of 0.7 Å for RNB (residues 396–835), 0.97 Å for CSD2 (residues 330–395) and 0.75 Å for S1 (residues 836–927). However, some discrete secondary-structural element movements were still obvious between them (Fig. 4a). In the most conserved domain (RNB), helices α10 and α12 are located in closer proximity to helix α14, resulting in a narrower RNA channel in SpDis3l2 than in the mouse Dis3l2–RNA complex (Fig. 4b). In the most variant domain (CSD2), the β-strands (β1–β4) match well in the two Dis3l2s, while a long helix α1 in SpDis3l2 is substituted by two small and discrete helices (residues 285–295) in the mouse Dis3l2–RNA complex (Fig. 4c). In the S1 domain, the β22 strand is four residues longer compared with the mouse Dis3l2–RNA complex. Moreover, a corresponding antiparallel β-sheet region (residues 867–894) cannot be observed in the SpDis3l2 structure owing to the poor electron density.
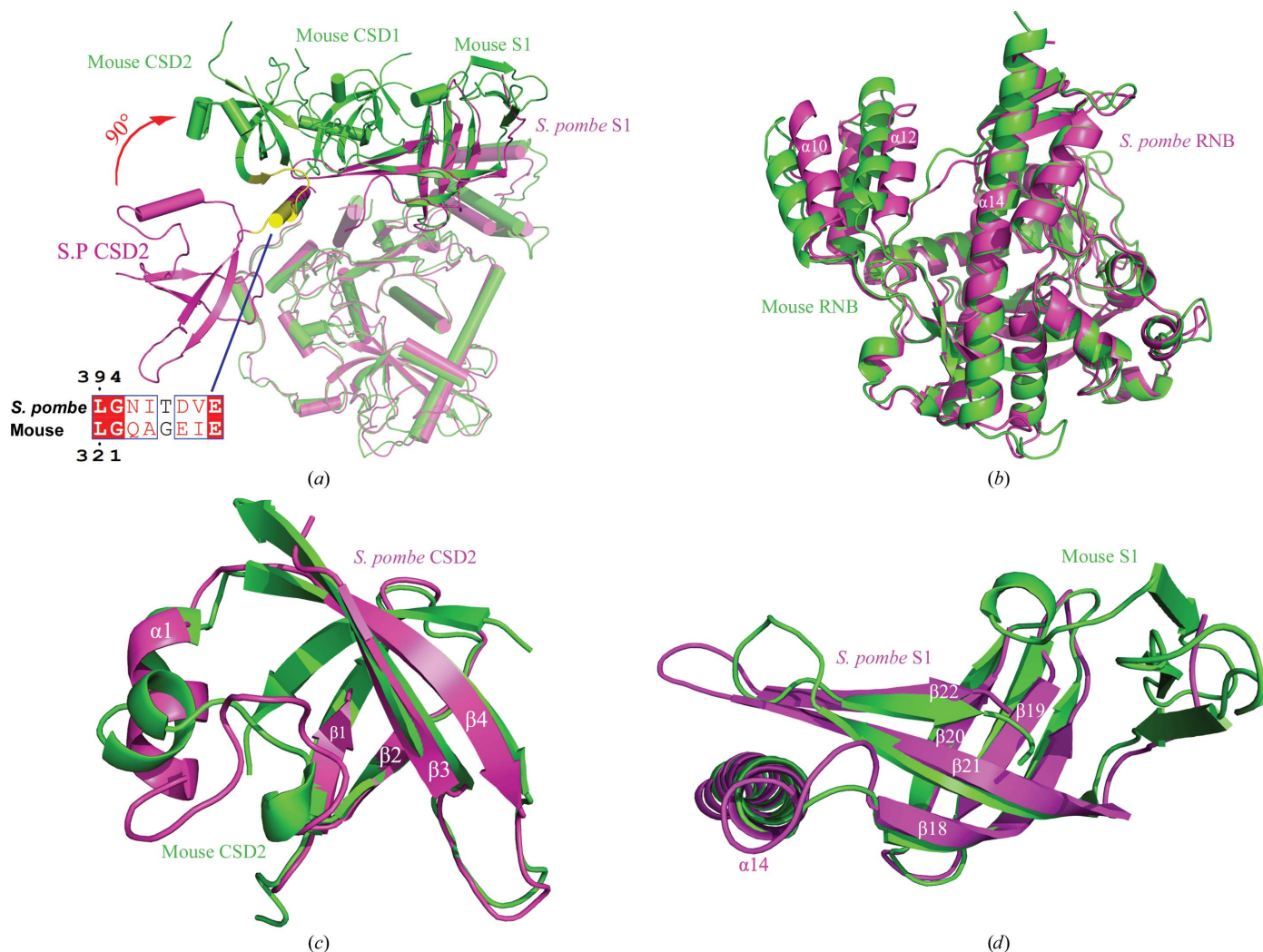

Figure 4
Structural comparison of SpDis3l2 and mouse Dis3l2 (RNA-bound form). (a) Integral contrast of the two structures (fission yeast in purple and mouse in green). The conserved residues coloured in yellow are located in the interdomain region that displays different secondary structures in the two species. (b) Superposition of the RNB domains of the two structures. (c) Superposition of the CSD2 domains of the two structures. (d) Superposition of the S1 domains and the RNB α14 helix of the two structures.

The most remarkable difference between the overall structures of these two homologues is that CSD2 adopts a different position relative to the RNB domain in SpDis3l2. Mouse Dis3l2 and ScRrp44 adopt a similar conformation when binding to the RNA substrate, with the CDSs and the S1 domain positioned on the top surface of the RNB domain. The three RNA-binding domains together form a positively charged entrance to the compact and rigid RNA channel (Faehnle et al., 2014). In the RNA-free SpDis3l2 structure, CSD2 is linked to the α2 helix located to one side of the RNB domain instead of lying on top of the RNB domain (Fig. 4a). This change in domain position induces a variation into the interdomain region between CSD2 and RNB (coloured yellow in Fig. 4a). In the mouse Dis3l2–RNA complex, the inter-domain loop (residues 323–327) is oriented vertically to the α2 helix and is followed by the contiguous β-strand of CSD2; in SpDis3l2 the α2 helix is seven residues longer than in the mouse Dis3l2–RNA complex, resulting in the placement of the CSD2 domain distant from the RNB domain, with a 90°

rotation compared with the orientation of the CSD2 domain in the mouse Dis3l2–RNA complex (Fig. 4a).

There are several probable explanations as to why the CSD2 domain of SpDis3l2 is oriented differently relative to the catalytic RNB domain compared with the structure of the mouse homologue. Firstly, the protein sequences of the two homologues are different, which may lead to a change of orientation of the CSD2 of SpDis3l2. Secondly, the crystal packing may have differently influenced the conformations of the Dis3l2s. Thirdly, RNA binding may result in the adoption of a different conformation by the Dis3l2s. RNA is bound to the protein in the mouse Dis3l2 structure, but this is not the case for the yeast structure; therefore, these differences may lead to changes in the three-dimensional structure.

Structural superposition of the RNA-free SpDis3l2 and the mouse Dis3l2–RNA complex shows that the orientation change of CSD2 starts at the interdomain region (coloured yellow in Fig. 4a). The residues in this region are highly conserved between the two species. Comparison of the
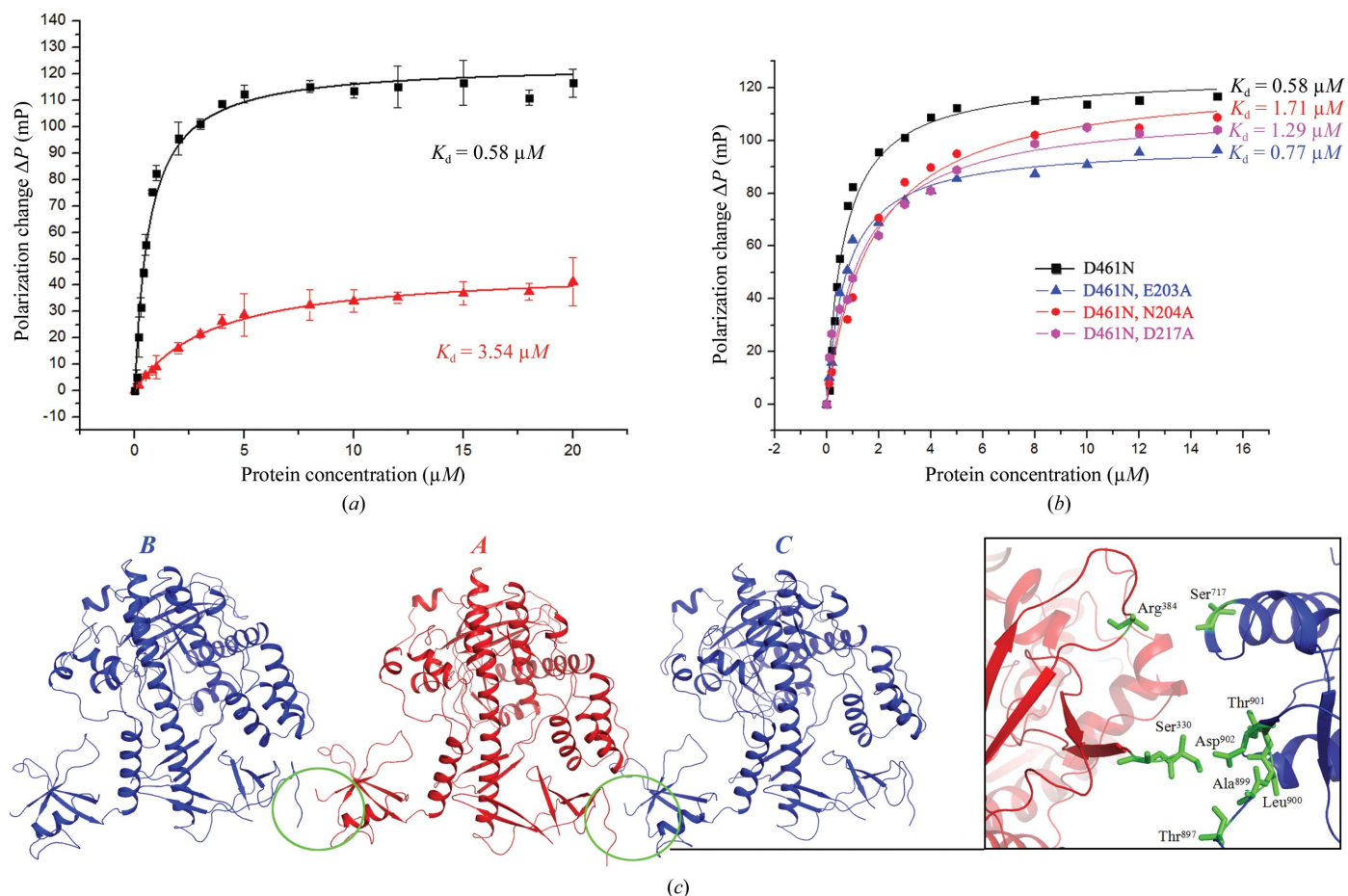

Figure 5
RNA-binding and crystal-packing analysis of SpDis3l2. (a) RNA-binding curves for full-length and CSDs-truncated forms of SpDis3l2. A catalytically inactive D461N mutant was used instead of the wild-type SpDis3l2 protein, and this catalytically inactive mutant was also included in all of the truncation/point mutants used in the RNA-binding analysis. Full-length SpDis3l2 (with D461N, black) and a fragment without the CSDs (with D461N, red) bind to 14-base poly(U)$_{14}$ RNA with K_d values of 0.584 and 3.54 μM, respectively. The binding abilities of the CSDs were too weak to be detected using the FPA method. (b) The RNA-binding curves of the SpDis3l2 mutants. The binding curves of full-length SpDis3l2 (with D461N, black) and the E203A (with D461N, blue), N204A (with D461N, red) and D217A (with D461N, purple) mutants are shown, with K_d values of 0.584, 0.744, 1.71 and 1.29 μM, respectively. (c) Two adjacent molecules (B and C, coloured blue) interact with the selected monomer of SpDis3l2 (A, coloured red) by crystal packing. The residues participating in intermolecular interactions are highlighted in green.

sequences between fission yeast and mouse shows that Leu, Gly and Glu are completely conserved in the two Dis3l2s, while Asn/Gln, Ile/Ala, Asp/Glu and Val/Ile are similar in terms of polarity. Moreover, we searched for the secondary structure of the corresponding sequence LGNITDVE from SpDis3l2 in the PDB. The same protein sequence formed an α-helix in a neutral peroxidase from *Arabidopsis thaliana* (PDB entry 1qgj; Mirza *et al.*, 2000), a β-sheet in the BT0572 protein from *Bacteroides thetaiotaomicron* (PDB entry 2f06; Midwest Center for Structural Genomics, unpublished work) and may form a flexible loop in the K7 protein from *Vaccinia virus* (PDB entry 2k36; Kalverda *et al.*, 2009). These results indicate that the conserved sequence of the interdomain region of SpDis3l2s is elastic and is capable of forming different secondary structures (α-helix, β-sheet or loop) depending on the environmental conditions and the energy state of each protein. However, the residues on the potential binding surface between the RNB and CSD2 domains are conserved across species. This finding indicated that the capability for interaction between the RNB and CSD2 domains of SpDis3l2 was retained in terms of protein sequence conservation. Thus, the difference in the protein sequence between SpDis3l2 and mouse Dis3l2 cannot be responsible for the drastic conformational change between the RNA-free SpDis3l2 and the mouse Dis3l2–RNA complex.

To investigate the possible effect of crystal packing on the SpDis3l2 structure, all of the adjacent molecules were analysed using *PISA* (Krissinel & Henrick, 2007). Two adjacent molecules were found to interact with SpDis3l2 in a repeating manner. As shown in Fig. 5(*c*), two molecules (*B* and *C*) interacted with the selected Dis3l2 molecule (*A*) *via* two hydrogen bonds and one salt bridge formed between Ser330 and Asp902; no disulfide and covalent bonds were found in the interface. Additionally, Ser331 and Arg384 of one molecule and Ser717, Thr897, Ala899, Leu900 and Thr901 of another molecule participated in the interaction through van der Waals

forces. These tiny interactions and the limited interface area (approximately 61.4 Å) between SpDis3l2 molecules would make the induction of such a drastic conformation change from a compact conformation similar to the mouse Dis3l2–RNA complex to the noncompact conformation of SpDis3l2 difficult.

Eliminating the possibilities of sequence differences and crystal packing, RNA binding could be the driving force behind the observed conformational heterogeneity between SpDis3l2 and the mouse Dis3l2–RNA complex. When the RNA substrate is bound to mouse Dis3l2, the CSDs (CDS1 and CDS2) are in close proximity to the RNB domain and form a pincer-like funnel surrounding the RNA substrate and the S1 domain. The interacting area between the CSDs and the RNB domain is 634.4 Å^2 in the mouse Dis3l2–RNA complex and seven nucleotides of the RNA substrate are recognized by the RNB and S1 domains. Without the CSDs–RNB and Dis3l2–RNA contacts, SpDis3l2 could display a less compact conformation in which the CSD2 domain stretches out and adopts a flanked position compared with CSD2 in the mouse Dis3l2–RNA complex.

3.3. RNA-binding characteristics of SpDis3l2

The structural variation between RNA-free SpDis3l2 and the mouse Dis3l2–RNA complex suggest that an investigation into the RNA-binding ability of SpDis3l2 is needed. To avoid RNA-substrate degradation during the RNA-binding measurements, a catalytically inactive D461N mutant was used instead of the wild-type SpDis3l2 protein; this catalytically inactive mutant was also included in all of the truncation/point mutants used in the RNA-binding analysis. Fluorescence polarization assays revealed that the full-length SpDis3l2[1–927] bound to the 14-base poly(U)$_{14}$ RNA substrate with a K_d value of 0.584 μM. In contrast, the K_d value of the CSDs-truncated mutant RNB/S1[398–927] decreased drastically to 3.54 μM (Fig. 5*a*). Moreover, the binding ability of the CSDs region (residues 170–397) was too weak to be detected by FPA (data not shown).

The above results indicated that full-length SpDis3l2 contains RNA-binding ability by itself, that the RNB and S1 domains contribute the major binding affinity to the RNA substrate and that the CSDs region also participates in the interaction process.

Because the S1 domain displayed a conserved OB-fold conformation, its RNA-binding ability would be similar to that of the CSD1 or CSD2 domains. The individual OB-fold only contains a small positively charged surface with very limited RNA-binding ability. Therefore, the dominant RNA-binding ability of the RNB and S1 domains would be contributed by the RNB domain. A
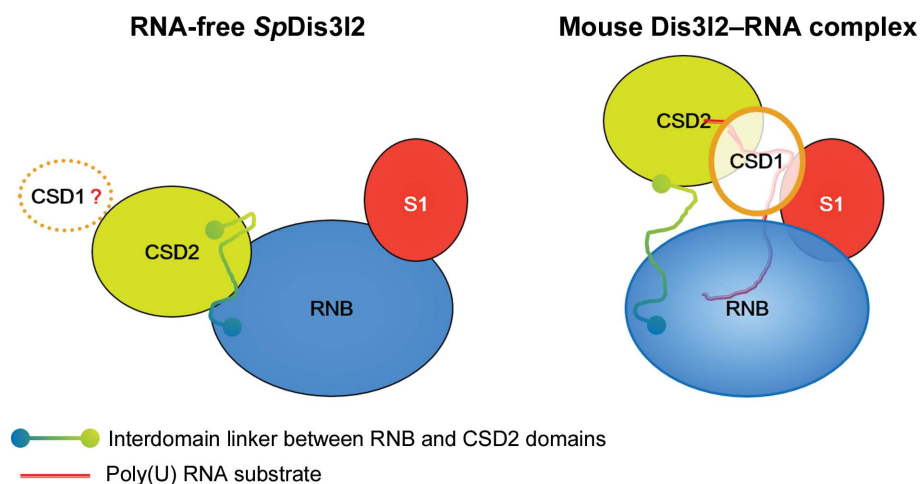

Figure 6
Schematic representation of the domain orientations in SpDis3l2 and the mouse Dis3l2–RNA complex. CSD2, RNB and S1 in both complexes are coloured yellow, blue and red, respectively. The CSD1 and RNA substrates in the mouse Dis3l2–RNA complex are coloured white and brick red, respectively. CSD1 of SpDis3l2 is shown as a dotted circle owing to its positional uncertainty.

structural comparison between *Sp*Dis3l2 and the mouse Dis3l2–RNA complex also supported this hypothesis. The residues involved in the RNA-binding surface of the mouse RNB domain (Asp381, Arg386, Asp387, Asp389, Asp340, Tyr433, Asn588, Arg606, Asn661, Arg665, Ala670, His686, Thr698, Ser699, Arg702 and Arg703; Fig. 3) are rigidly conserved in *Sp*Dis3l2 and form a positively charged channel with a strong RNA-binding ability (Fig. 2c).

The difference in RNA-binding affinity between full-length *Sp*Dis3l2 and the CSDs-truncated form indicates an indispensable role for the CSDs in stimulating the RNA-binding affinity of *Sp*Dis3l2. The RNA-binding ability of the individual OB-folds is limited; however, when the three OB-folds form a pincer-like funnel to fix the RNA substrate, the RNA-binding affinity of the mouse Dis3l2–RNA complex is highly stimulated. Similar to these results, the conserved CSD1, CSD2 and S1 domains of *Sp*Dis3l2 are likely to form the same pincer-like funnel to fix the RNA substrate and increase the RNA-binding affinity of *Sp*Dis3l2. This hypothesis explains why the addition of the CSDs region with little RNA-binding ability can increase the RNA-binding affinity of *Sp*Dis3l2 by approximately an order of magnitude. However, CSD2 and S1 are located distantly from one another in the *Sp*Dis3l2 structure and cannot form a pincer-like funnel in their present spatial positions (Fig. 6). A drastic domain allostery will be required to allow the OB-folds to form a pincer-like funnel surrounding the RNA substrate.

3.4. The possible RNA-binding pattern of *Sp*Dis3l2

Mouse Dis3l2 and *E. coli* RNase II act as general exonucleases that degrade a variety of RNA sequences independently of their association with the exosome but contain different RNA-binding patterns (Fig. 7). In the *E. coli* RNase II–RNA complex, CSD1 does not directly interact with the RNA substrate, and CSD2 forms interactions with the RNA substrate only through a flexible loop (residues 102–105); moreover, the cleft between CSDs and S1 is much wider and open (Frazão *et al.*, 2006). In the mouse Dis3l2–RNA complex, the CSDs play a key role in binding RNA *via* a series of residues (especially in CSD1, including Arg74, Pro77, Lys79, Phe80 and Asp93), which are distinct from those in RNase II. This strong interaction shortens the distance between the CSDs and the top surface of the RNB in the complex with RNA substrate (Faehnle *et al.*, 2014).

Sequence alignment of the CSDs regions of *Sp*Dis3l2, mouse Dis3l2 and *E. coli* RNase II revealed that the *Sp*Dis3l2 CSD1 domain possesses a similar sequence identity to those of mouse Dis3l2 and *E. coli* RNase II (12.1 and 11.39%, respectively); however, the sequence identity between the
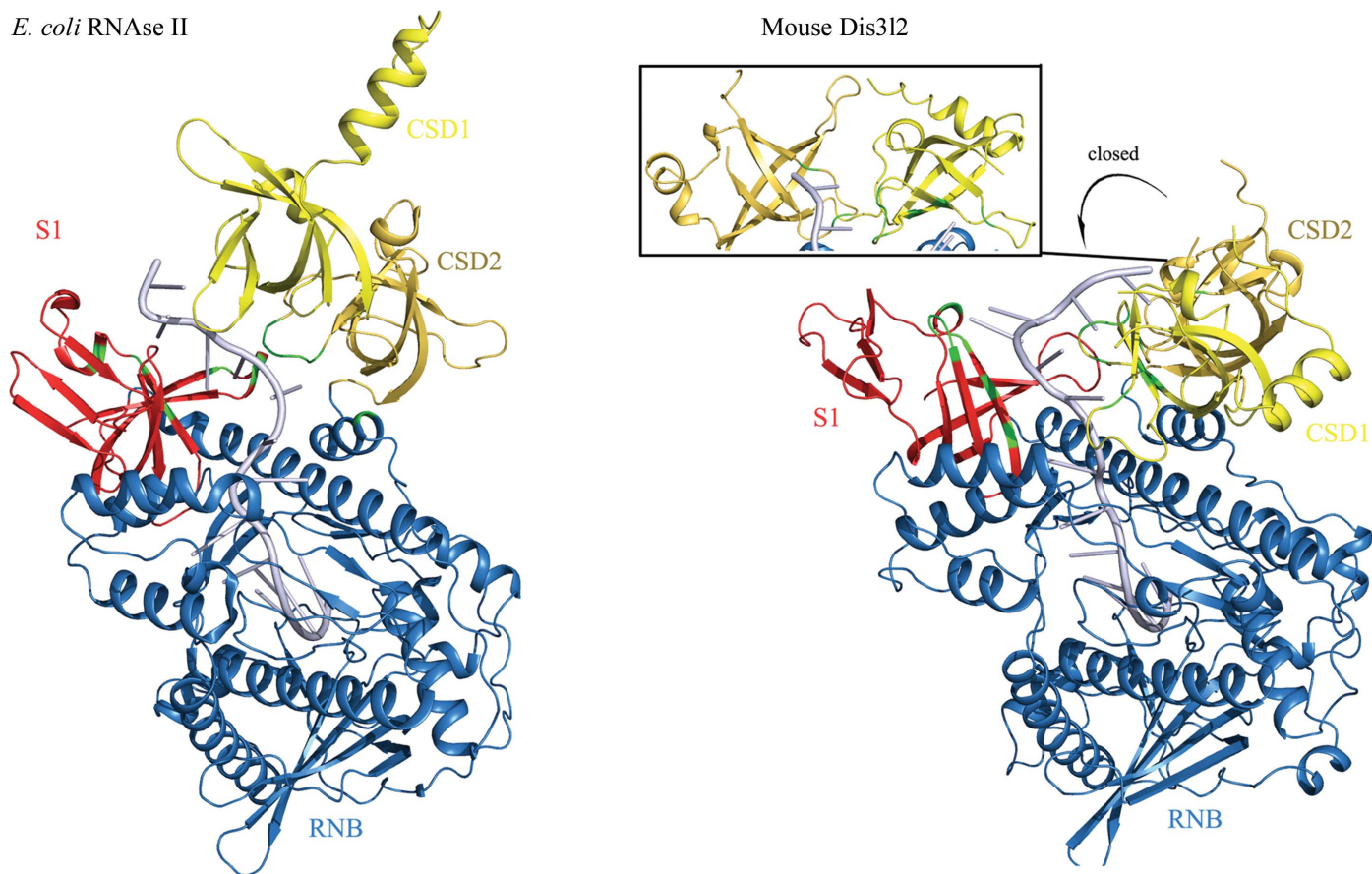

Figure 7
Discriminating interactions between ssRNA bound to either *E. coli* RNase II or mouse Dis3l2. The two structures are shown in the same orientation. The residues participating in RNA binding in the two complexes are highlighted in green.

CSD2 domains of *Sp*Dis3l2 and mouse Dis3l2 (19.32%) was much higher than that between those of *Sp*Dis3l2 and *E. coli* RNase II (8.33%). Based on the sequence similarity between the CSD2 domains of *Sp*Dis3l2 and mouse Dis3l2, RNA binding to *Sp*Dis3l2 would result in the CSD2 domain leading the CSD1 domain to adopt an RNA-binding pattern that resembles that of its mouse orthologue.

Mutagenic experiments were performed to further investigate whether the CSD1 domain takes part in the interaction with RNA in *Sp*Dis3l2. Combinations of the D461N mutation with three mutations in the CSD1 domain (E203A[D461N], N204A[D461N] and D217A[D461N]) were constructed and their binding affinity to the 14-base poly(U)$_{14}$ RNA was tested. The results revealed that the RNA-binding affinities of the three mutations were decreased compared with the full-length *Sp*Dis3l2$^{1-927}$, with K_d values of 0.74 μM (E203A[D461N]), 1.71 μM (N204A[D461N]) and 1.29 μM (D217A[D461N]). This result indicates that the CSD1 domain of *Sp*Dis3l2 participates in the interaction with the RNA substrate and that *Sp*Dis3l2 may adopt an RNA-binding pattern that resembles that of its mouse orthologue.

Combining the sequence-alignment analysis with the experimental mutagesis results, we are inclined to propose that *Sp*Dis3l2 may adopt an RNA-binding pattern that resembles that of the mouse Dis3l2–RNA complex following RNA substrate binding, which is characterized by the movement of the three OB-fold-containing domains (CSD1, CSD2 and S1) to form a pincer-like funnel surrounding the RNA substrate.

3.5. A drastic conformational change could occur during the RNA-binding process of *Sp*Dis3l2

Taking the RNA-free structure and the deduced RNA-binding pattern of *Sp*Dis3l2 into consideration, a drastic conformational change would need to occur during the RNA-binding process. During the RNA-capturing process, CSD2 together with CSD1 would need to rotate by approximately 90° to anchor the RNA chain and to significantly enhance the affinity between *Sp*Dis3l2 and the RNA substrate. Similar conformation changes caused by RNA binding have been found in human HuR (Wang *et al.*, 2013). The RNA-free structure of human HuR reveals an open conformation with no interdomain contacts between the two RRM domains; following binding of the primary RNA, the RRM1 domain captures the RNA substrate with five nucleotides (U5–U8). The released binding energy induces the subsequent conformational changes of the interdomain linker and RRM2, forming a closed conformation that improves the RNA-binding affinity of HuR (Wang *et al.*, 2013). During the RNA substrate-binding process, the CSDs (CDS1 and CDS2) of *Sp*Dis3l2 could move close to the RNB domain and form a pincer-like funnel surrounding the RNA substrate with the S1 domain based on the sequence conservation between *Sp*Dis3l2 and mouse Dis3l2. The potential interaction area between *Sp*CSD2 and the RNB domain could be approximately 600 Å^2 (the area of the interface between the CSDs

and the RNB domain is 634.4 Å^2 in the mouse Dis3l2–RNA complex), and the RNB and S1 domains of *Sp*Dis3l2 could recognize approximately seven nucleotides of the RNA substrate. Furthermore, the free energy released from the above interactions of both CSDs–RNB and Dis3l2–RNA could support the energy consumption during the entire allosteric process.

4. Conclusion

To elucidate the RNA-recognition mechanism of *Sp*Dis3l2, we solved its crystal structure at 2.8 Å resolution and investigated its RNA-binding properties. The *Sp*Dis3l2 structure displays an open conformation that is distinct from its mouse homologue in the RNA-bound form. Significantly, CSD2 accompanied by CSD1 in RNA-free *Sp*Dis3l2 adopt a nearly vertical orientation compared with the counterparts in the mouse Dis3l2–RNA complex. The FPA results indicated that the RNB and S1 domains contribute the major binding affinity for the RNA substrate and that the CSDs region also participates in the interaction process. Combining the structural comparisons with mutagenic experiments, we speculate that *Sp*Dis3l2 could adopt an RNA-recognition pattern that resembles that of its mouse homologue but not that of *E. coli* RNase II. When the RNA substrate binds to *Sp*Dis3l2, a drastic conformational change could occur. However, the detailed allosteric mechanism and protein–RNA contact information for *Sp*Dis3l2 remain unclear. Further studies of its protein–RNA complex structure and its biochemical and molecular-biology characteristics are urgently needed.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Astuti, D. *et al.* (2012). *Nature Genet.* **44**, 277–284.
Büttner, K., Wenig, K. & Hopfner, K. P. (2005). *Mol. Cell*, **20**, 461–471.
Chang, H.-M., Triboulet, R., Thornton, J. E. & Gregory, R. I. (2013). *Nature (London)*, **497**, 244–248.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Dziembowski, A., Lorentzen, E., Conti, E. & Seraphin, B. (2007). *Nature Struct. Mol. Biol.* **14**, 15–22.
Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. (2014). *Nature (London)*, **514**, 252–256.

Frazão, C., McVey, C. E., Amblar, M., Barbas, A., Vonrhein, C., Arraiano, C. M. & Carrondo, M. A. (2006). *Nature (London)*, **443**, 110–114.

Hernández, H., Dziembowski, A., Taverner, T., Séraphin, B. & Robinson, C. V. (2006). *EMBO Rep.* **7**, 605–610.

Houseley, J. & Tollervey, D. (2009). *Cell*, **136**, 763–776.

Ibrahim, H., Wilusz, J. & Wilusz, C. J. (2008). *Biochim. Biophys. Acta*, **1779**, 256–265.

Kalverda, A. P., Thompson, G. S., Vogel, A., Schroder, M., Bowie, A. G., Khan, A. R. & Homans, S. W. (2009). *J. Mol. Biol.* **385**, 843–853.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Liu, Q., Greimann, J. C. & Lima, C. D. (2006). *Cell*, **127**, 1223–1237.

Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A. & Conti, E. (2008). *Mol. Cell*, **29**, 717–728.

Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G. & Conti, E. (2005). *Nature Struct. Mol. Biol.* **12**, 575–581.

Malecki, M., Viegas, S. C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M. G. & Arraiano, C. M. (2013). *EMBO J.* **32**, 1842–1854.

Mirza, O., Henriksen, A., Østergaard, L., Welinder, K. G. & Gajhede, M. (2000). *Acta Cryst.* D**56**, 372–375.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. (1997). *Cell*, **91**, 457–466.

Schilders, G., van Dijk, E., Raijmakers, R. & Pruijn, G. J. (2006). *Int. Rev. Cytol.* **251**, 159–208.

Schoenberg, D. R. & Maquat, L. E. (2012). *Nature Rev. Genet.* **13**, 246–259.

Staals, R. H., Bronkhorst, A. W., Schilders, G., Slomovic, S., Schuster, G., Heck, A. J., Raijmakers, R. & Pruijn, G. J. (2010). *EMBO J.* **29**, 2358–2367.

Tomecki, R., Kristiansen, M. S., Lykke-Andersen, S., Chlebowski, A., Larsen, K. M., Szczesny, R. J., Drazkowska, K., Pastula, A., Andersen, J. S., Stepien, P. P., Dziembowski, A. & Jensen, T. H. (2010). *EMBO J.* **29**, 2342–2357.

Vanacova, S. & Stefl, R. (2007). *EMBO Rep.* **8**, 651–657.

Wang, H., Zeng, F., Liu, Q., Liu, H., Liu, Z., Niu, L., Teng, M. & Li, X. (2013). *Acta Cryst.* D**69**, 373–380.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.