NMR assignment and solution structure of the external DII domain of the yeast Rvb2 protein

Benoît Bragantini, Clément Rouillon, Bruno Charpentier, Xavier Manival, Marc Quinternet

NMR assignment and solution structure of the external DII domain of the yeast Rvb2 protein

Benoit Bragantini[1], Clément Rouillon[1], Bruno Charpentier[1], Xavier Manival[1] & Marc Quinternet[2*]

*Corresponding author: Marc Quinternet marc.quinternet@univ-lorraine.fr

1 Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA), UMR 7365 CNRS Université de Lorraine, Biopôle, Campus Biologie Santé, Vandœuvre-lès-Nancy, France

2 Ingénierie-Biologie-Santé Lorraine (IBSLor), UMS 2008, CNRS, INSERM, Université de Lorraine, Biopôle, Campus Biologie Santé, Vandœuvre-lès-Nancy, France

Abstract

We report the nearly complete ^{1}H, ^{15}N and ^{13}C resonance assignment and the solution structure of the external DII domain of the yeast Rvb2 protein, a member of the AAA+ ATPase superfamily.

Keywords : R2TP, Chaperone, snoRNP biogenesis, AAA+ ATPase, Rvb, RUVBL

Biological context

Rvb1 and Rvb2 proteins, also known as RUVBL1/RUVBL2, TIP49/TIP48 or pontin/reptin, are two highly conserved ADP/ATP-binding proteins belonging to the eukaryotic AAA+ (ATPases associated with diverse cellular activities) superfamily (Jha and Dutta 2009). They are essential for a large number of cellular processes and are involved in the assembly of an eclectic set of protein or ribonucleoprotein complexes, both in cytoplasm and nucleus (Mao and Houry 2017).

Rvb1 and Rvb2 are very similar in terms of amino-acid sequence and structural organization, each of them encompassing three distinct domains (DI, DII, and DIII). More in details, the domains I (Rossmann-like α/β/α-fold) and III (all α-helical) provide the signature of typical oligomeric rings of an AAA+ core, which is formed by a conserved Walker A motif (required for nucleotide binding), a conserved Walker B motif (required for ATP hydrolysis), sensor motifs 1 and 2 (required for the sensing of ADP or ATP), and an Arg-finger (Walker et al. 1982). Rvb1 and Rvb2 are organized in a hetero-hexameric ring made of 3 copies of each protein and several data in the literature suggest that a dodecameric double ring could be formed through direct contacts involv- ing either the AAA+ cores or the domains II of Rvbs (Ewens et al. 2016; Lakomek et al. 2015; Silva-Martin et al. 2016). Domain II of Rvbs contains two sub-domains: an oligonucleotide-binding fold (external DII or DII_{ext} or α/β Tip) and a base region closer to the AAA+ ring (internal DII). The DII domain was shown involved in several protein–protein interactions as observed in the INO80 complex (Tosi et al. 2013), in the yeast SWR1 complex (Nguyen et al. 2013) and in the R2TP complex (Rivera-Calzada et al. 2017). Moreo- ver, DII_{ext} domains seem to play a role in the Rvb1–Rvb2 dodecamerization interface (Ewens et al. 2016; Lakomek et al. 2015). Indeed, quite recently, gel filtration experiments performed using mutants of Rvb proteins from *Saccharomyces cerevisiae* showed that single amino-acid variants in DII_{ext} (R180E in Rvb1 or E166K in Rvb2) could disrupt the double ring in vitro, reinforcing the idea that Rvb1 DII_{ext} faces Rvb2 DII_{ext} to form an hetero-dodecamer (Ewens et al. 2016). More than helicases, Rvbs show a chaperone-like behavior necessitating energy coming from ATP hydrolysis (Tian et al. 2017; Zhou et al. 2017). Thus proteins, not DNA, seem to be the biologically appropriate substrate for Rvb1/2 and their DII_{ext} could therefore regulate their activity and oligomeric state of the Rvbs.

Only middle-resolution cryoEM data are available for Rvbs coming from budding yeast. In this study,

we deter- mined the $^{1}H–^{13}C–^{15}N$ chemical shift pattern and the 3D solution structure of the external DII domain of Rvb2 from *S. cerevisiae* with the aim to bring data making possible its study with putative partners at the atomic scale.

Methods and experiments

The N- and C- terminal extremities of the external DII domain of the Rvb2 protein from S. cerevisiae (scRvb2DIIext) were designed on the basis of the available 3D structures of the full-length protein of *Chaetomium thermophilum* (ctRvb2, PDB codes : 4WVY, 4WW4, 5FM6 and 5FM7) (Lakomek et al. 2015; Silva-Martin et al. 2016). In the latter structures, the external DII domain is connected to the DI and internal DII domains via a β-sheet made of two antiparallel strands. In the budding yeast, these two strands should stretch from residues 125–135 and from residues 230–238. For our NMR study, we consequently chose fragment 132–234 of Rvb2 that should divide this connecting β-sheet in two parts. Then, the DNA sequence corresponding to this fragment was inserted into a pnEA-3CH expression vector. For protein overexpression, Ca2+-competent BL21 (DE3) *E. coli* cells were transformed and grown overnight at 37°C in 10 mL of selective LB medium. The latter culture was used to inoculate 1 L of a selective M9 medium supplemented with 0.5 g/L of 15NH4Cl and 2 g/L of 13C d6-glucose. Bacteria were put at 37 °C, under agitation. When OD600 reached ~ 0.6, overexpression of the 6xHis-tag protein was induced by 0.2 mM IPTG and bacteria were kept at 20 °C for 16 h, under agitation. Cells were then harvested by centrifugation and sonicated in lysis buffer [25 mM HEPES (pH 7.5), 300 mM NaCl, 0.5 mM TCEP, 10 mM Imidazole]. The overexpressed protein was purified from the soluble fraction of the cell sonicate using TALON beads previously equilibrated in the lysis buffer. Three fixation and washing steps were performed. PreScission protease was used to elute the protein in lysis buffer. Finally, size-exclusion chromatography using a Superdex 200 column was performed in the NMR buffer [10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP], enabling us to obtain around 10 mg of pure protein.

A 1 mM sample of 13C/15N-labelled protein in NMR buffer (supplemented with 5% D2O and 1 mM 2H–DTT) was analyzed in a 5 mm Shigemi tube at 298 K using a 600 MHz spectrometer equipped with a TCI cryoprobe (Bruker). An almost full assignment of informative 1H, 15N and 13C resonances was obtained using 1H–15N HSQC, 1H–13C HSQC, HNCA, CBCACONH, HNCACB, HNCO, HNCACO, HNHA, HBHACONH, HCCCONH, CCCONH, HCCH–TOCSY, 1H–15N NOESYHSQC and 1H–13C NOESYHSQC spectra. For backbone amide groups, the steady-state 1H–15N heteronuclear NOE was assessed through intensity ratios between interleaved reference and saturation experiments. Spectral data were processed with TopSpin 3.2 (Bruker) then handled with CARA (Keller 2004). All resonances were referenced to DSS (4,4-dime- thyl-4-silapentane-1-sulfonic acid) in regard of the IUPAC-IUB recommended chemical shift referencing ratios. In addition, secondary structure predictions were derived from backbone resonance assignment using TALOS-N (Shen and Bax 2013). Inter-proton distances were derived from 2D 1H–1H NOESY and 3D 1H–15N and 1H–13C

NOESYHSQC spectra all recorded with a mixing time of 120 ms. CYANA 3.97 was used to iteratively assign NOE cross-peaks (Guntert 2004). Hydrogen bond restraints derived from 1H/2H NMR exchange experiments and dihedral angle restraints originating from TALOS-N predictions were added to the calculation. The last CYANA cycle was carefully checked and the final set of distance and dihedral restraints was used to generate 100 water-refined structures thanks to RECOORD scripts (Nederveen et al. 2005). 1H–15N residual dipolar coupling (RDC) data were added in the RECOORD refinement steps. RDCs were measured at 298 K and 600 MHz using 1H–15N TROSY and 1H–15N semi-TROSY experiments (Lescop et al. 2007) recorded in the NMR buff supplemented or not with pf1 phages. The final concentration of pf1 was 13 mg/mL for 500 µM of 13C/15N scRvb2DIIext. Finally, the 20 structures with the lowest overall energies were selected as the most representative structures. The electrostatic potential of the best structure was determined using PDB2PQR and APBS webservers (Dolinsky et al. 2004).

Assignment and data deposition

1H–15N HSQC spectrum of scRvb2DIIext is shown in Fig. 1. All observable correlations in this spectrum have been assigned to a backbone NH group or to an Asn/Gln NH2 group. Very few signal overlaps were observed, suggesting first that each of the 103 native residues in scRvb2DIIext possesses a distinguishable chemical environment. The finger- print left by the 1H–15N HSQC spectrum was in agreement with the presence of secondary structure elements.

89.1% of the 1H, 13C and 15N chemical shifts were assigned in scRvb2DIIext. More in details, 98.5% of the backbone resonances were assigned (i.e. HN, N, CO, Cα and Hα nuclei) and, if we focus on 1H attached to 13C and 15N and on heteroatoms attached to 1H, 92.5% of the 1H shifts, 98.6% of the 13C shifts and 76.2% of the 15N shifts were assigned. In the very large majority, missing nuclei came from positively- charged groups in Arg and Lys residues. Chemical shifts have been deposited in the BioMagResBank under Accession Number 34173 (http://www.bmrb.wisc.edu). Extraction of secondary structures using TALOS-N permitted to mainly predict 6 β-strands and one α-helix in the N-terminal part of scRvb2DIIext (Fig. 2A). As for it, the C-terminal part did not exhibit any propensity for structuration.

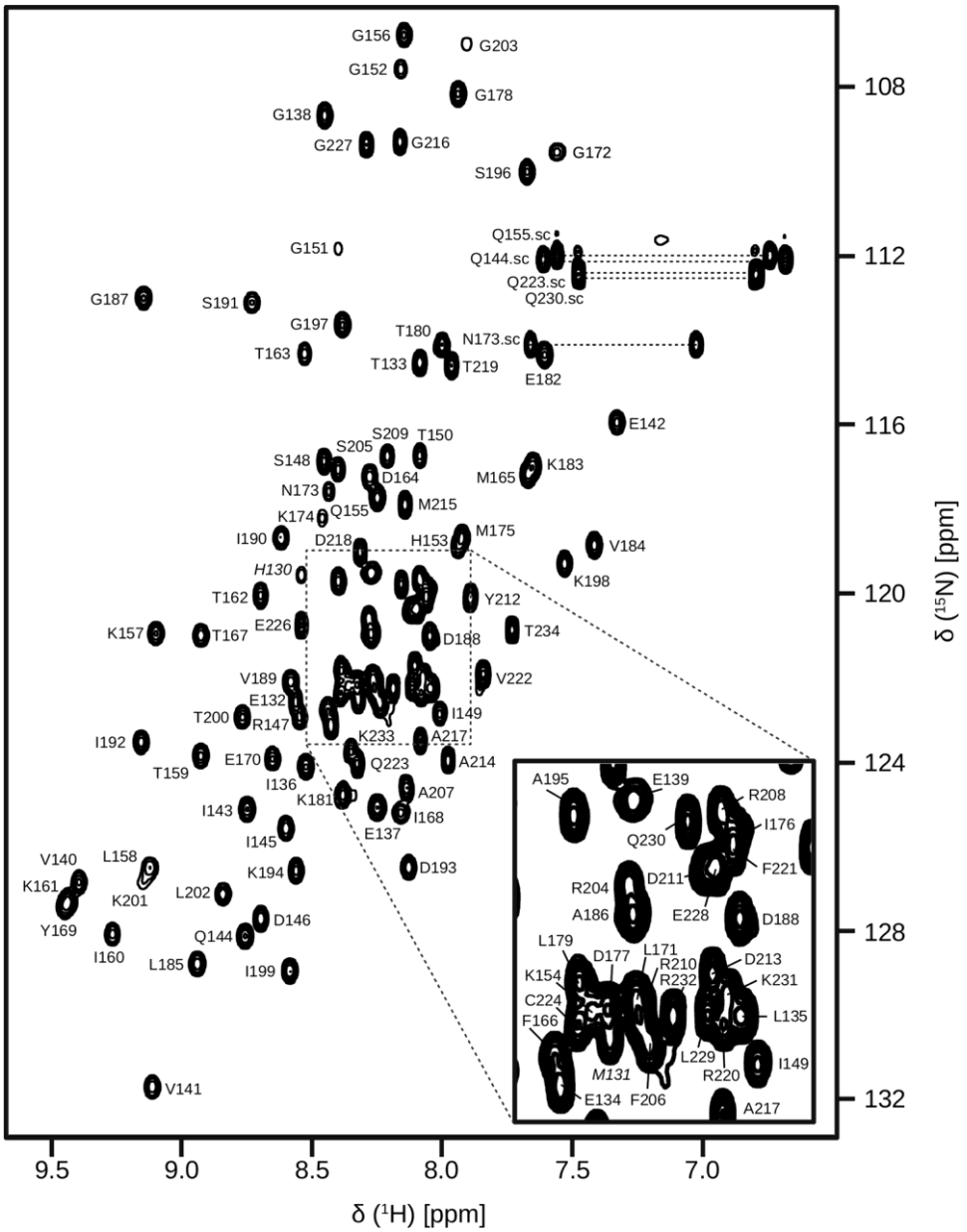

Fig. 1 $^{1}H–^{15}N$ HSQC spectrum of $^{15}N/^{13}C$ *sc*Rvb2DII$_{ext}$ recorded at 600 MHz, 298 K in 10 mM NaPi (pH 6.4), 150 mM NaCl, 0.5 mM TCEP and 1 mM $^{2}H–$ DTT. Assignments of side- chain –NH2 groups in Gln and Asn residues are labelled with identifier ".sc". The assignments of residues belonging to the expression tag are labelled in italic. A zoom on the central region of the spectrum is also displayed

We went further and obtained for scRvb2DIIext a well- resolved ensemble of final 3D structures with respective RMSD values to the best structure of 0.36 ± 0.05 and 0.98 ± 0.15 Å for backbone and heavy atoms included in secondary structures (Fig. 2B). 3D coordinates and NMR restraints were deposited in the Protein Data Bank under Accession Code 5OUN (http://www.rcsb.org/pdb/home/home.do). Calculation statistics are summarized in Supplementary Table 1. The pattern of secondary structures forms a core domain consisting in five β-strands and one α-helix, in agreement with predictions based on chemical shifts (Fig. 2A, C). More precisely, the long strand β1 bridges two β-sheets, both composed of three strands (β-sheet 1: β1–β2–β3 and β-sheet 2: β1–β5–β6) (Fig. 2B, C). Helix α1 contains hydrophobic residues that interact with the internal faces of β-sheets 1 and 2 to induce the curvature of strand β1. As a result, helix α1 looks as if it was trapped inside a jaw formed by the two β-sheets (Fig. 2D). This 3D organization makes it possible to protect hydrophobic residues and

to expose several charged residues to the solvent. Analysis of the electrostatic potential of the core domain scRvb2DIIext revealed a large negative region that could represent a putative interaction surface (Fig. 2E). We also observed that scRvb2DIIext has a highly disordered C-terminal tail comprising residues 203–234 (Fig. 2B). The 1H–15N heteronuclear NOE ratios measured in the backbone amide groups of scRvb2DIIext were in good agreement with the lack of structural organization in this tail (with a ratio < 0.6) (Fig. 2F). Similarly, three other regions corresponding to the N-terminal residues and to the inner loops β1–β2 (147–155) and β2–β3 (163–165) can be described as flexible (Fig. 2B, F). The NMR structure determined for scRvb2DIIext is in good agreement with the X-ray structure of ctRvb2 (Cα RMSD on secondary structures of 0.8 Å; PDB code: 5FM7) and with the NMR structure of Rvb2DIIext from Homo sapiens (Cα RMSD on secondary structures of 0.7 Å; PDB code: 2CQA), for which no chemical shifts were deposited (Fig. 2A).

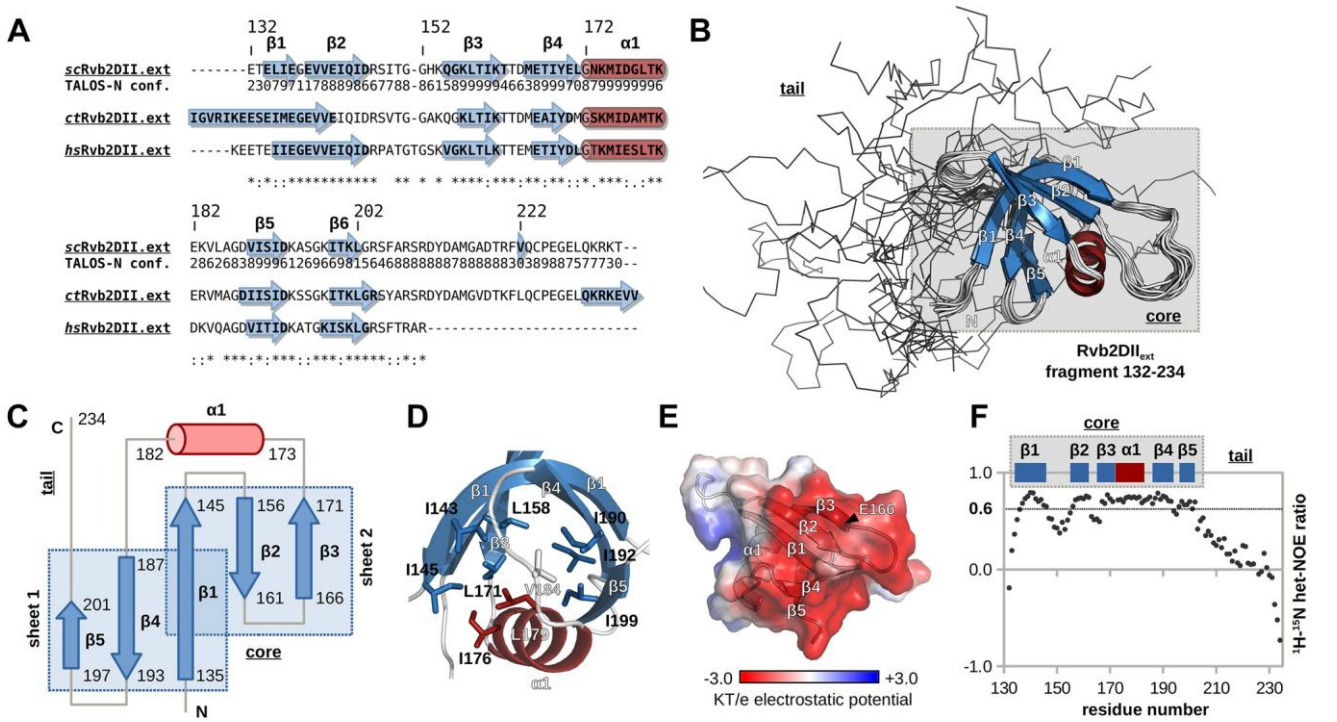

Fig. 2 **A** Sequence alignment of external DII domains from *Sac- charomyces cerevisiae* (*sc*Rvb2DII$_{ext}$), *Chaetomium thermophilum* (*ct*Rvb2DII$_{ext}$) and *Homo sapiens* (*hs*Rvb2DII$_{ext}$). TALOS-N results obtained for *sc*Rvb2DII$_{ext}$ are indicated onto the amino-acid sequence and confidence score of the prediction is indicated for each residue. Secondary structures derived from the 3D structures of *ct*Rvb2DI- I$_{ext}$ (PDB code: 5FM6) and *hs*Rvb2DII$_{ext}$ (PDB code: 2CQA) are also indicated on the amino-acid sequence. β-strands and α-helices are represented using arrows and cylinders respectively. Fully con- served, strongly similar and weakly similar amino-acids are respec- tively indicated with an asterisk (*), a colon (:) and a period (.). Alignment was done using Clustal Omega (Sievers et al. 2011). Theresidue numbering corresponds to *sc*Rvb2DII$_{ext}$. **B** Cartoon repre- sentation of the NMR structures of *sc*Rvb2DII$_{ext}$. The core region is in the grey box. **C** 2D topology of *sc*Rvb2DII$_{ext}$. The limits of the secondary structures are indicated. **D** Zoom on the

hydrophobic core of *sc*Rvb2DII$_{ext}$ and on the curvature of strand β1. Side-chains and labels of hydrophobic residues are given. For ease of interpretation, hydrogens were not represented. **E** Electrostatic potential calculated on the core region of *sc*Rvb2DII$_{ext}$. Negative and positive charges are represented on a "red to blue" scale and shown on the molecular sur- face. **F** Steady-state $^{1}H–^{15}N$ heteronuclear NOE values of backbone amide groups plotted against the residue number. The dotted line on the graph represents a value of 0.6.

In conclusion, our construct of scRvb2DIIext is not able to achieve the formation of a β-sheet involving its N- and C- terminal parts (as thought for the DIIext in the full-length Rvb2) for the benefit of a disordered tail. However, it forms an expected folded core that encompasses residues important for the dodecamerization of the Rvb1/Rvb2 proteins (E166, Fig. 2E). From recent studies showing the biological relevance of the hetero-hexameric form of Rvbs in presence of protein partner, we can assume that the DIIext core is a target of choice for Rvbs substrates. Our study provides a toolbox to follow in solution the self-association of Rvbs or interactions involving the DIIext core of Rvb2 at the atomic scale.

Acknowledgements
We thank the platform of biophysics and structural biology of UMS 2008 IBSLor (CNRS-INSERM-UL) for NMR facilities. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the pôle Biologie, Médecine, Santé (BMS) of the University of Lorraine (UL) and the Agence Nationale de la Recherche [ANR-11-BSV8-01503; ANR-16-CE11-0032-02].

References

Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB- 2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res 32:W665–W667. https://doi.org/10.1093/nar/gkh381

Ewens CA, Su M, Zhao L, Nano N, Houry WA, Southworth DR (2016) Architecture and nucleotide-dependent conformational change of the Rvb1-Rvb2 AAA+ complex revealed by cryoelectron microscopy. Structure 24:657–666. https://doi.org/10.1016/j. str.2016.03.018

Guntert P (2004) Automated NMR structure calculation with CYANA. Methods Mol Biol 278:353–378. https://doi.org/10.1385/1-59259-809-9:353

Jha S, Dutta A (2009) RVB1/RVB2: running rings around molecular biology. Mol Cell 34:521–533. https://doi.org/10.1016/j.molce l.2009.05.016

Keller RLJ (2004) Computer-aided resonance assignment tutorial, 1st edn. CANTINA, Zürich

Lakomek K, Stoehr G, Tosi A, Schmailzl M, Hopfner KP (2015) Struc- tural basis for dodecameric assembly states and conformational plasticity of the full-length AAA+ATPases Rvb1. Rvb2. Structure 23:483–495. https://doi.org/10.1016/j.str.2014.12.015

Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-reso- nance experiments for time-optimized protein resonance assign- ment. J Magn Reson 187:163–169. https://doi.org/10.1016/j.jmr.2007.04.002

Mao YQ, Houry WA (2017) The role of pontin and reptin in cellular physiology and cancer etiology. Front Mol Biosci 4:58. https:// doi.org/10.3389/fmolb.2017.00058

Nederveen AJ et al (2005) RECOORD: a recalculated coordinate data- base of 500 + proteins from the PDB using restraints from the BioMagResBank. Proteins 59:662–672. https://doi.org/10.1002/prot.20408

Nguyen VQ, Ranjan A, Stengel F, Wei D, Aebersold R, Wu C, Leschziner AE (2013) Molecular architecture of the ATP-depend- ent chromatin-remodeling complex SWR1. Cell 154:1220–1231. https://doi.org/10.1016/j.cell.2013.08.018

Rivera-Calzada A et al (2017) The structure of the R2TP complex defi a platform for recruiting diverse client proteins to the HSP90 molecular chaperone system. Structure 25:1145–1152. https://doi.org/10.1016/j.str.2017.05.016

Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural net- works. J Biomol NMR 56:227–241. https://doi.org/10.1007/s1085 8-013-9741-y

Sievers F et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75

Silva-Martin N et al (2016) The combination of X-ray crystallography and cryo-electron microscopy provides insight into the overall architecture of the dodecameric Rvb1/Rvb2 complex. PLoS ONE 11:e0146457. https://doi.org/10.1371/journal.pone.0146457

Tian S et al (2017) Pih1p-Tah1p puts a lid on hexameric AAA+ATPases Rvb1/2p. Structure 25:1519–1529. https://doi.org/10.1016/j. str.2017.08.002

Tosi A et al (2013) Structure and subunit topology of the INO80 chro- matin remodeler and its nucleosome complex. Cell 154:1207– 1219. https://doi.org/10.1016/j.cell.2013.08.016

Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myo- sin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1:945–951

Zhou CY et al (2017) Regulation of Rvb1/Rvb2 by a domain within the INO80 chromatin remodeling complex implicates the yeast Rvbs as protein assembly chaperones. Cell Rep 19:2033–2044. https://doi.org/10.1016/j.celrep.2017.05.029