Title

Dimerization of human Rio2 kinase/ATPase locks its ATP-binding site in an apo state

Running title

Crystal structure of human Rio2 kinase

Frédérique Maurice, Natacha Pérébaskine and Sébastien Fribourg*

INSERM U1212, UMR CNRS 5320, Université de Bordeaux,

146, rue Léo Saignat, Bordeaux, 33076, France

*Correspondance e-mail: sebastien.fribourg@inserm.fr

ABSTRACT

Rio proteins form a conserved family of atypical protein kinases. Rio2 is a serine/threonine protein kinase/ATPase involved in pre-40S ribosomal maturation. Current crystal structures of archaeal and fungi Rio2 proteins report a monomeric form of the protein. Here, we describe three atomic structures of the human Rio2 kinase showing that it forms a homodimer. Upon self-association, the ATP-binding pocket is hidden from the solvent and the protein is locked in an apo state corresponding to an inactive form of the kinase. The homodimerization is mediated by key residues previously shown to be responsible for ATP binding and catalysis. This unusual protein kinase dimer reveals an intricate mechanism of mutually exclusive substrate binding and oligomeric state formation. We propose that this oligomeric state could serve a dual function in maintaining the protein in an inactive state and being a novel type of nuclear import signal.

Significance Statement:

Rio kinases form a family of atypical protein kinases that are believed to be ATPases rather than kinases. The three members of the Rio family are involved in ribosome biogenesis. We show here that contrarily to what was reported so far, Rio2 is able homodimerize in a conformation that locks it

in an apo state, preventing its (re)association to pre-mature ribosomes. This unconventional self-association is not seen in any other protein kinase. This mechanism is likely to be transient and could used to efficiently re-import the protein to the nucleus.

Keywords: Ribosome biogenesis, 40S, atypical Kinase, dimer.

Introduction

Ribosome biogenesis is a complex mechanism initiating in the nucleolus with the transcription of a long ribosomal RNA precursor (pre-rRNA) by RNA polymerase I and ending in the cytoplasm after a number of inter-twinned maturation events including methylations, nucleotide pseudouridylations, endo- and exonucleolytic cleavages and a mandatory transit through the nuclear pore (1, 2). Some of the 200 associated factors (AF) involved in ribosome biogenesis (3–8) display characteristic sequence motifs of enzymes such as Ser/Thr kinases, GTPases, RNA helicases, and are probably involved in the many checkpoint controls. The three proteins forming the RIO kinases family, namely Rio1, Rio2 and Rio3, have a distinct role in cytoplasmic maturation of the pre-40S pre-rRNA (9–13). They form a family of Ser/Thr kinases build around a central protein kinase domain but that do not exhibit further sequence similarity with other eukaryotic protein kinases (ePKs), and as such have been classified as atypical protein kinases (14, 15). Rio1, Rio2 and Rio3 central kinase domain is flanked by unrelated N-terminal and C-terminal domains and extensions. In yeast, RIO1 and RIO2 are essential genes and their product, Rio1p and Rio2p, associate to ribosomal pre-40S particles (11, 13, 16). Both factors are involved in D-site cleavage *in vitro* and *in vivo* (11, 13, 16, 17). Rio1 is additionally involved in cell cycle progression and rDNA stability (18–20). The individual depletion of Rio1p and Rio2p leads to an accumulation of the 20S pre-rRNA in yeast (16, 17, 21, 22) and of the 18S-E pre-rRNA in human cells (9). In human, hRio3 is exclusively cytoplasmic and is part of the pre-40S particles. It is associated to the 18S-E pre-rRNA. hRio3 deletion leads to 21S rRNA accumulation in human (10).

In human, the Rio2 protein (hereafter hRio2) presence, but not its activity, is necessary for the recycling of the hEnp1 factor from the cytoplasm to the nucleus, whereas hRio2 activity is required for the release of hDim2, hLtv1, hNob1 and hRrp12, from the 18S-E pre-rRNA (13, 23). Similarly, in yeast, Rio2p catalytic activity is not required for its association to the pre-40S particle but rather for its release (24, 25). Apart from a weak autophosphorylation activity, Rio2 is phosphorylating hDim1 in human cells (13, 24, 26). Altogether, this led to the proposal that Rio2 is rather an ATPase than a protein kinase. Finally, the position of Rio2p on the pre-40S particle has been established at the P-site close to domain IV of Tsr1p where both proteins block the A- and P-sites on the 40S subunit (27, 28).

The structure of the archaeal *Archeoglobus fulgidus* Rio2 (afRio2) and *Chaetomium thermophilum* Rio2 (ctRio2) have been solved in apo and ATP-bound states (14, 24, 29). These atomic models revealed the domain organization of Rio2 and gave insights into the enzymatic reaction mechanism. Each of these structures described Rio2 kinase as a monomer. Here we report that hRio2 can form a homodimer in solution. We solved the crystal structure of the homodimer of hRio2 and observe that several invariant or highly conserved residues are involved in the dimer interface. In the hRio2 homodimer, the ATP-binding site is locked in a conformation incompatible with ATP binding. Moreover, residues involved in the homodimer interaction are also required for ATP-binding and catalysis. The overlap of the enzymatic activity and dimer formation are mutually exclusive and incompatible with pre-40S particle association. We propose that this state might be either transient or represent a specific state that could be used *in vivo* for protein import to the nucleus for example.

MATERIAL AND METHODS

Constructs

All constructs were sequenced to ensure the absence of mutations. The human Rio2 kinase constructs, 1-321, 1-353 and 1-321 deleted from residues 131 to 150 were amplified from the DNA encoding the full-length hRio2 and subcloned into the *Nde*I and *BamH*I sites of a modified pET-15b

(Novagen) plasmid (30) to produce an N-terminal His-tag fused protein containing a TEV cleavage site and an ampicillin resistance marker. The *Chaetomium thermophilum* (ct) Rio2 (kind gift from N. Laronde-Leblanc) was subcloned into a pET-15b derived plasmid to generate an N-terminally His-tag protein.

Protein expression and purification

The various constructs were transformed in *E. coli* BL21(DE3) Gold cells. The cultures were grown at 37 ºC in 2xYT medium supplemented with 100 mg/L ampicillin and 30 mg/L kanamycin. Cells were induced at 20 ºC with 0.5 M IPTG for 14 hours and collected by centrifugation at 4,500 g and resuspended in buffer containing 50 mM Tris pH 8.0, 350 mM or 900 mM NaCl. Cell pellets were lysed with an Emulsi Flex-C3 (Avestin) and centrifuged at 50,000 g. The clarified cell lysate was combined with His-Select Co^{2+} resin (Sigma) for 1 hour at 4 ºC. The resin containing bound proteins was washed with 5 column volumes of resuspension buffer and the protein complexes were eluted with a 10 mM - 250 mM imidazole gradient. The hRio2 (1-321), hRio2 (1-321)Δ(131-146), hRio2 (1-353) Δ(131-146) were further purified with a HiQ-Sepharose (GE Healthcare) and size exclusion chromatography over a Superdex 200 (GE Healthcare). The Selenomethionine substituted hRio2 (1-321) Δ (131-146) proteins was purified as the native hRio2 (1-321).

Analytical Ultracentrifugation

Analytical ultracentrifugation experiments were performed in a Beckman Coulter ProteomeLab XL-I instrument (IGBMC Structural Biology Platform) at 4°C. ctRio2 and hRio2(1-321)Δ(131-46) proteins were independently purified and dialyzed against a buffer containing 25 mM Tris-HCl pH 7.4, 300 mM NaCl and 5% Glycerol (w/v). For analytical ultracentrifugation sedimentation velocity experiments at 50,000 RPM, 400 µl of a concentration series of each protein was prepared into the dialysis buffer at a final concentration of 60, 30, 15, 7.5 and 3.75 µM for ctRIO2 and 54, 27, 13.5, 6.75 and 3.375 µM for hRio2. Absorbance scans at 280 nm and interference scans were taken every 5 minutes. Sedimentation data were first analyzed using SEDFIT software (31) and the continuous

sedimentation coefficient distribution model c(s) were generated and then exported into SEDPHAT for fitting with the hybrid local continuous distribution model (32). Buffer density, buffer viscosity and protein partial specific volumes were calculated using SEDNTERP software. GUSSI was used to integrate the sedimentation peaks and to produce the graphs (33).

Crystallization & Structure determination

Prior to crystallization, the native hRio2 (1-321), hRio2 (1-353)Δ(131-146) and the Selenomethionine substituted hRio2 (1-321)Δ(131-146) proteins were exchanged in buffer containing 25 mM Tris, pH 7.5, 900mM or 350 mM NaCl (respectively), 1.0 mM DTT and concentrated to 10 mg/ml. The complexes were crystallized at 20°C by the sitting drop method against a reservoir containing either condition D2 and G2 of Morpheus™ screen (Molecular Dimensions). Crystals were transferred to a cryo protectant with 60 % EDO-P8K, flash-frozen in liquid nitrogen, and maintained at 100 K in a nitrogen cryostream during data collection.

Crystals of hRio2 (1-321) and hRio2 (1-321)Δ(131-146) belong to the $P2_1$ space group with unit cell dimensions a=69.13 Å b=88.30 Å c=102.14 Å and contain four molecules per asymmetric unit. Crystals of hRio2 (1-353)Δ(131-146) belong to the $P6_5$ space group with unit cell dimensions a=67.55Å b=67.55 Å c=312.2 Å

The structure of the hRio2 (1-321)Δ(131-146) was solved by SAD using Selenomethionine substituted crystals. Datasets were reduced using XDS (34). 40 selenium sites were located with SHELXCD (35) and refined with Phaser (36). The initial model was build using Bucaneer (37). The final models were refined with BUSTER 2.10 (38) and have good stereochemistry (Table 1). The atomic model of hRio2 (1-321)Δ(131-146) comprises residues 6-312. The atomic model of hRio2 (1-353)Δ(131-146) comprises residues 9-314. The atomic model of hRio2 (1-321) comprises residues 12 to 291 with the exception of loop 66-70 and 129-149.

RESULTS

Solution analysis of human Rio2

Rio2 is a phylogenetically conserved protein whose size varies from approximately 30 kDa in the archaea *A. fulgidus* to about 60 kDa in human. Despite this large variation in size, the protein is organized around a N-terminal winged helix domain followed with a classical kinase domain architecture with a N- and C-terminal lobes. In higher eukaryotes, the C-terminus contains additional elements such as a Nuclear Export Signal (NES) (39, 40). To help characterizing hRio2, various constructs were generated as the full-length protein was found poorly expressed and extremely unstable. C-terminally truncated constructs were engineered in order to remove natively unfolded regions of the protein.

To evaluate the oligomeric state of hRio2 in solution, we used sedimentation velocity analytical ultracentrifugation (SV-AUC). To this end, we purified to homogeneity hRio2 construct encompassing residues 1-321 but deleted from the unfolded loop (residues 131-146) (24, 29). As a control, we also used the full-length ctRio2 (24). Both protein samples were used at protein concentrations varying from 3.75 μM to 30 μM for ctRio2 and from 3.37 μm to 27 μM for hRio2 in SV-AUC (Fig. 1A to 1D). For ctRio2 samples, the continuous c(s) distribution plots showed one major sedimentation species with a sedimentation coefficient of 3.16 S and an estimated molecular mass of 44.80 kDa, which is close to the theoretical value of 47.24 kDa for the monomer. These results indicate that ctRio2 behaves as a single monomeric species at all tested protein concentrations (Fig. 1A and 1C and Table 1). For hRio2(1-321)Δ(131-146), the c(s) analysis revealed a behavior depending on the protein concentration with a large majority (> 94 %) of the sample behaving as a monomer at lowest concentrations and an increasing proportion of dimer at higher concentrations (> 20% at 27 μM) (Fig. 1B and 1D and Table 1). In comparison to the reported atomic models of afRio2 and ctRio2, these sedimentation velocity experiments reveal an unforeseen behavior for hRio2 able to form homodimers.

Structure determination

In order to characterize this unexpected dimeric state of hRio2, we decided to use protein crystallization as it is usually performed at high concentrations, an experimental condition favourable for hRio2 dimerization. Diffracting crystals were obtained for the construct encompassing residues 1-321 and the best native diffraction datasets were collected up to 3.2 Å resolution. Due to crystallization reproducibility issues, limited diffraction arising from this construct and the failure to solve the structure by molecular replacement, novel constructs of hRio2 were designed. In particular, the residues 131 to 146, belonging to a known unfolded loop (29) were deleted and native or Se-Met derived protein crystals were grown and diffraction data were collected (Table 2). The structure of hRio2 (1-321)Δ(131-146) was determined by SAD phasing on Se-derived protein crystals. The refined structure was used as a model to solve the crystal structure of hRio2 (1-321) and of a longer construct, hRio2(1-353)Δ(131-146). All refinement statistic for the three atomic structures can be found in the Table 2.

Overall structure description

The hRio2 kinase adopts a similar fold as the archeon *A. fulgidus* afRio2 (1TQP) and the thermophilic fungi ctRio2 (4GYI) (24, 29).The N-terminus (residues 10 to 75) folds as a winged helix domain (Fig. 1E), adjacent to the two lobes of the RIO kinase domain (respectively residues 76-190 and 196-291). hRio2 structure contains the typical motifs of canonical Ser/Thr protein kinases with the C-Helix (light green), the F-Helix (magenta), the catalytic loop (in dark blue) and the glycine loop (red, also known as P-loop). The hinge linker (residues 191-195, black) involved in the recognition of the ATP adenine moiety in afRio2, connects the two lobes of the kinase domain (Fig. 1E). No electron density for ATP or any ligand is seen in the electron density map at or near the active site.

The hRio2(1-321) and hRio2(1-353)Δ(131-146) have 1.2 Å root-mean-square deviation (r.m.s.d.) over 249 C-α carbon and differ mainly by the relative position of the C-lobe (Fig. S1A). Despite their sequence identity and structure similarity with hRio2, cfRio2 and atRio2 display slightly different domain orientation as shown by their relative r.m.s.d. of 2.5 Å and 1.5 Å over 221 and 244 and C-α

carbon respectively (Fig. S1B and S1C). The N-terminal winged helix domain and the C-terminal lobe of the kinase domain are not in the exact same orientation in comparison to hRio2(1-321) (Fig. S1).

Structure of human Rio2 homodimer

In the asymmetric unit, we observe a tight homodimer formed by the association of two hRio2 molecules in a head to head orientation independently of the three space groups or of the construct size (Fig. 2A). The F- and the C-helices form the major part of the homodimer interface without significant impact on the N-terminal winged helix domains. The F-helix of one monomer interacts with the F-helix of the other monomer in an antiparallel mode (Fig. 2B).

The ATP binding regions, via the glycine loop (residues K105 and E106), the activation loop (D246 and Q249) and the catalytic loop (D228 and E231) are providing a significant portion of the interacting surface (Fig. 2C). Numerous polar interactions between D228 of one monomer are stabilized by the hydroxyl group of Y261 and R264 of the other monomer (Fig. 2C). A similar organization is seen for the other monomer. Each C-helix of a monomer (light green) is facing the C-helix of the opposite monomer in an antiparallel mode (Fig. 2D). The interaction between C-helices is mainly restricted to Van der Waals contacts provided by L153 and L156 of each monomer. Altogether the complete surface of interaction involves around 1,570 $Å^2$ per monomer. Bio-informatics analysis using the EPPIC webserver (41) to assess the biological significance of interacting surfaces found in protein-protein complexes, largely favours a biological role for the observed dimerization (score of 98 %) and thus would exclude a crystallographic artefact (score of 1%) (42, 43).

In the longest crystallized form of hRio2 (residues 1 to 353), some of the additional C-terminal residues form an α–helix packing at the interface between the two monomers (Fig. 2A and 2D). This helical extension bridges the C-helix stabilizing further the dimer (E299 and R129 in Fig. 2D). Importantly, all the solved structures of hRio2 shows a dimeric organisation in the asymmetric unit. Taken together with the results from the sedimentation experiments, the observed dimerization of hRio2 is likely to mimic a functional situation.

The αI of ctRio2 adopts a different conformation in hRio2.

ctRio2 and hRio2 superimpose over 244 Cα residues with a r.m.s.d. of 2.5 Å (Fig. 3A). One most significant difference between hRio2 and ctRio2 models apart from the oligomeric state, is the position of the eukaryotic specific α-helix (αI, orange), which packs in the close vicinity to the ATP-binding site in ctRio2 (24)(Fig. 3A and 3B). In ctRio2, the C-terminal residues (325-343) form a long α-helix (αI) that wraps around the protein and contributes to the ATP binding site environment (Fig. 3B). αI interacts with the N- and C-kinase lobes of the protein and it is linked to the remainder of the protein through a short α-helix (residues 317-322, pink) embedded within a loop (Fig. 3B). In hRio2, most of the equivalent residues form a continuous α-helix that packs against the C-helix, away from the ATP binding site (Fig. 3C). It appears that the conserved triad residues G-F/Y-T adopts a different secondary structure (Fig. 3D). Due to hRio2 homodimerization, αI cannot adopt the same orientation as observed in ctRio2. These data reveal the existence of two possible organizations for the residues 297-312 forming the αI helix in which the G-F/Y-T triad acts as a switch.

Rio2 homodimerization surface is conserved in eukaryotes

To the best of our knowledge, dimerization of afRio2 or ctRio2 has not been observed in the published crystal structures (24, 29). In an attempt to understand this discrepancy with hRio2, we looked at sequence alignments of Rio2 orthologues to identify hot spot(s) of residue conservation (Fig. S2). We then plotted the residue conservation across species on the surface of hRio2 atomic model using the webserver ConSurf (44)(Fig. 4). When all species are considered, sequence conservation is limited to the ATP-binding pocket (Fig. 4A and 4B). However, when sequence conservation is calculated using only eukaryotic sequences or multicellular eukaryotic sequences, invariant residues cluster not only at the ATP-binding pocket but also along the dimer interface (Fig. 4C to 4E). Of note, the same surface area is used for the association of Rio2p with 20S pre-rRNA (helices 1, 18, 28, 34, and 44), Tsr1, Ltv1, uS12, us7, uS13 and uS19 (27, 28). Altogether, this suggests

an overlapping role for a subset of residues that are involved in pre-40S particle association and dimer formation. The superimposition of ctRio2 and hRio2 suggests that αI prevents ctRio2 dimer formation. This observation is in agreement with the absence of dimer for ctRio2 in AUC experiments. It is also interesting to notice that this eukaryotic-specific α-helix is disordered in the recent structure of the pre-40S particle (27, 28), suggesting that the αI is a mobile secondary structure that may not be involved in ATP catalysis in pre-40S maturation.

Homodimerization locks hRio2 ATP-binding pocket in an apo form

The afRio2 and ctRio2 structures have been solved in complex with various ligands including ADP, ATP and toyocamycin (14, 24). In order to analyse the local rearrangements of the Cα backbone around the ATP binding site, the ADP-bound structure of ctRio2 and the hRio2 structure were superimposed and the overall structure around the ADP is shown (Fig. 5). In comparison to the ADP-bound structure of ctRio2, the ligand-binding pocket of hRio2 has undergone substantial local rearrangements. The main difference occurs at the glycine loop (red) that adopts a closed conformation in hRio2. In that orientation, the glycine loop prevents ligand binding (Fig. 5A and 5B). The glycine loop of one monomer is stabilized by a series of interactions with the other monomer (Fig. 5B, grey shading). These interactions depend on residues D228, R264 and Y261 of the adjacent monomer. The glycine loop is also stabilized through interactions between G104 and H255, E106 and Q249. Remarkably every residue involved in the glycine loop stabilization is highly conserved across species (Fig. S2). Despite many attempts to soaking in the crystals with various ATP analogues and to incubate the protein sample prior crystallization with ATP analogues, we never managed to obtain a liganded form of hRio2, nor managed to crystallize a monomeric from. Altogether, it suggests a concerted molecular mechanism by which hRio2 homodimerization favours a non-liganded form of the protein.

DISCUSSION

In this article, we report the crystal structure of three different constructs of human Rio2 protein

kinase. The structures extend from the N-terminus up to the residues 290 and 312 respectively. Homodimeric forms of hRio2 where no ligand was observed in the ATP-binding pocket were obtained. Dimer formation relies on a network of interactions resulting in the inward movement of the glycine loop and the occlusion of the ATP-binding pocket. There is an intricate molecular mechanism linking homodimer formation and ATP pocket closure. This observation is reminiscent of the function of the glycine (P-loop) described in pre-40S binding and ATP hydrolysis stimulation (25).Residues involved in dimer formation are highly conserved in higher eukaryotes suggesting a functional relevance for dimer formation in these species.

In other protein kinases, dimerization is a common feature driving their activation (45, 46). This is true for EGFR, RAF and CHK2 kinases for instance (47). In the case of B-RAF, dimerization occurs at the opposite face of the ATP-binding pocket and leaves the protein able to bind ATP (48). Activation is thought to occur via an allosteric mechanism in which one of the monomer acts as a scaffolding protein (the activating kinase) to stabilize the active conformation of the other monomer (the receiving kinase). CHK2 protein kinase dimerization is promoted by the phosphorylation of T68 residue by ATM kinase. This phosphorylation leads to a transient dimerization and CHK2 activation through the activation-loop auto-phosphorylation (47). Once phosphorylated the dimer rapidly dissociates into monomers. Homodimer formation is also observed in the *Legionella pneumophila* LegK4 kinase (49). In this structure, the protein dimerizes through αF, αG and the αG-αI loop of the C-lobe in both the apo and AMP-PNP bound states, but this is different from the dimer formation observed for hRio2. Here we have a dimeric form of hRio2 that locks its ATP-binding pocket in a conformation where ATP cannot be accommodated, in an opposite mechanism to CHK2. Finally, In the cryo-EM reconstructions of yeast and human pre-40S particles, only a monomer of Rio2 could be fitted in the attributed density for Rio2 (24, 27, 28, 50). Altogether, these data strongly suggest that the dimeric form of Rio2 is an inactive form.

The C-terminal residues described in ctRio2 and hRio2 structures display different conformations. In the dimeric form of hRio2, they provide additional interactions to strengthen the dimer. In the

monomeric form of ctRio2, the C-terminal helix participates in interactions around the catalytic centre but may play an auto-inhibitory action on ATP catalysis (24). Finally, in the pre-40S particle cryo-EM reconstruction, this eukaryotic specific α-helix is not observed (27, 28). Its position in ctRio2 structure is also incompatible with the dimer formation. Together, this suggests that this eukaryotic specific substructure may have a role as a switch between a dimeric inactive and an monomeric active form of Rio2.

Several point mutations impair hRio2 kinase activity. This is the case for the K123A/D246A double mutation in hRio2 (13) and the D229A substitution in Rio2p (22). While the K123A/D246A double mutation will clearly affect the ATP recognition based on ctRio2/ADP complex structures (Fig. 5A & 5B), the D228A substitution may have a double effect on both the catalytic activity and the dimer formation. Indeed, the D228 side chain (D229 in ctRio2) points towards the neighbouring hRio2 monomer and not towards the ATP binding pocket. This suggests that an identical set of residues is used for ATP recognition/catalysis or peptide substrate activation and dimer formation in a mutually exclusive mechanism. This might be an elegant molecular mechanism for segregating active versus inactive forms of the enzyme.

At this stage, the biological role of a dimer of hRio2 remains elusive but hypotheses can be formulated towards a putative regulatory mechanism. A lot of efforts have been made in order to understand protein and RNA trafficking across the nuclear pore complex (51–53), however there is a limited knowledge concerning the many AFs required for ribosome biogenesis that have to be imported to the nucleolus (52, 54). For the yeast Rio2p, it has been shown that it possesses a C-terminal NES required for the Crm1p-driven pre-40S particle export (40). In human, hCrm1 binds residues 391 to 403 of hRio2 as a NES (39). An efficient re-import mechanism of AFs has to be set in order to ensure the next round of pre-mature ribosomal particles export. This is especially true if one considers that some of the trans-acting factors might be in limiting quantities. This is likely to be true for Rio2 as previous experiments showed that if you prevent Rio2p release from the yeast cytoplasmic pre-40S particles, you rapidly deplete the pool of Rio2p present in the nucleus and no

free Rio2p is observed (24). This implies that the whole pool of Rio2p is immobilized on pre-40S particles. Thus, for an efficient maturation process starting in the nucleus, nuclear (re)import of Rio2 must be as efficient as pre-40S particle export in order to recycle Rio2. Having Rio2 as a dimer, thus in its apo form, may combine several advantages. It would i) prevent re-association of Rio2 to pre-mature or mature 40S particles in the cytoplasm after its release, ii) prevent any ATPase activity to occur if not bound to the pre-40S particle, iii) dimerization could serve as a recognition signal or be required for its import as no classical NLS can be identified from the primary sequence, no importin in charge of Rio2 has been reported so far. Alternatively, the use of any importin in an opportunistic and serendipity manner could be envisaged. Dimer formation for nuclear import has been reported in the case of uS3 (RpS3). In this mechanism, the nuclear import of uS3 is chaperoned by Yar1 and the ternary complex is imported to the nucleus by Kap60/Kap95 (52, 55). Dimer dissociation would have to be promoted by an external and yet unidentified factor. It has been shown, in the case afRio1 and hRio1 that the oligomeric state of the protein impacts on its kinase activity (56) and that auto-phosphorylation plays an important role in such a mechanism. The switch from a cytoplasmic inactive and dimeric state to a nuclear active and monomeric state remains elusive so far and awaits further analysis. However, it provides an exciting regulatory mechanism for specific AF re-import to the nucleus after their cytoplasmic release.

ACKNOWLEDGEMENTS

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and we would like to thank staff members for assistance in using beamline ID14-1 and ID29. We also acknowledge the Synchrotron SOLEIL and Drs Pierre Legrand and Andrew Thompson for providing access to Proxima-1 beamline. We acknowledge N. Laronde-Leblanc for providing the ctRio2 plasmid. We thank S. Thore for critical reading of the manuscript.

FUNDING

This work was supported by an ANR Blanc 2010 grant RIBOPRE40S, University of Bordeaux, CNRS and INSERM. The authors acknowledge the support and the use of resources of the French Infrastructure for Integrated Structural Biology FRISBI ANR-10-INBS-05 and of Instruct-ERIC.

Authors contributions

FM and NP expressed, purified and crystallized human Rio2. SF solved and refined the structure. SF analyzed the data, wrote the paper and supervised the work.

The authors declare that they have no competing financial interests.

Figure Legends

Fig. 1: Analytical UltraCentrifugation (AUC) studies and structure of human Rio2 kinase.

AUC experiments were performed at 4 °C and 50,000 rpm. For clarity one over three acquisitions is shown corresponding to about 15 min time intervals. Sedimentation coefficient distributions *c*(s) determined from Sedfit/Sedphat (31, 33) analysis of absorbance data at 280 nm.

A) Superposition of experimental and fitted sedimentation velocity profiles (top), and their differences (bottom) for ctRio2 at 15 µM and **B)** hRio2 at 13.5 µM.

C) ctRio2 at 30 µM (orange line), 15 µM (black line), 7.5 µM (dashed line) and 3.75 µM (dotted line) appeared at all concentrations as a monomer. Sedimentation velocity data were plotted using the software GUSSI.

D) hRio2 samples at 27 µM (orange line), 13.5 µM (black line), 6.75 µM (dashed line) and 3.375 µM (dotted line) were analyzed. hRio2 appears as a dimeric upon increasing concentrations.

E) Structure of the hRio2 encompassing residues 1 to 353. The winged helix is located at the N-terminus, the two kinases lobes are labelled and connected by a hinge linker loop (in black). The so-called C-helix (light green), F-helix (magenta), glycine loop (red), catalytic loop (blue) and activation loop (yellow) are also shown and labelled accordingly to the canonical representation (15). C-terminal residues to the F-Helix are coloured in pink and orange. This and all subsequent structural Fig.s were generated with PyMOL (57).

Fig. 2: Human Rio2 forms homodimers

A) Human Rio2 interacts with itself to form an inter-twinned homodimer with a head to head orientation. A general view of hRio2 homodimer is shown under the same orientation as in Fig. 1A with monomer A in the exact same orientation.

B) Close-up view of the homodimer interface around F-Helix. One monomer is shown with a grey surface. Salt bridges between residues D271 and K267, and residues K267 and D271 are maintaining the two F-helices in close interaction. R264 of each monomer is stacking together through Van der Waals interactions, while Y261 is involved in a H-bond network with K106 and D228 of the other monomer.

C) Close-up view of the homodimer interface around the ATP-binding region. One monomer is shown with a grey surface. The catalytic residue D228 is involved in H-bonding with Y261 and K106 while K105 backbone interacts with H255 and Q249 stabilizes K105 side chain. E231 is interacting with W260. The catalytic residues K123 and D246 are shown.

D) Close-up view of the homodimer interface around C-Helix. The C-terminal α-helix is coloured in pink and orange. Aliphatic residues such as L153 and L156 are providing Van Der Waals interaction force that add to the many other interactions between the two monomers.

Fig. 3: αI can adopt two conformational orientations

A) Superimposition of hRio2(1-353)Δ(131-146) and ctRio2 has been performed in Coot (58). ctRio2 is displayed in grey.

B) Focus on the αI of ctRIo2. The residues 311 to 324 (pink and orange) of ctRio2 form a short α-helix embedded in a linker region that packs against the C-helix (light green). The GF/YT switch region is labelled.

C) The similar region of hRio2 as residues 311 to 324 of ctRio2 are displayed under the same orientation as in B. This region adopts an extended α-helix conformation projecting the C-terminal end of the downstream α-helix residues out of the ATP-binding pocket vicinity.

D) Sequence alignment of Rio2 orthologues in eukaryotes around the 296 to 308 residues region of hRio2. The secondary structures observed in ctRio2 and hRio2 are represented below the alignment. The GF/YT switch residues are boxed in red.

Fig. 4: Rio2 homodimer interface is conserved in eukaryotes.

The sequence conservation has been calculated using Consurf server (http://consurf.tau.ac.il/) with the sequence alignment displayed in Fig. S2. The conservation above 70 %, 80 %, 90 % and equal to 100 % are shown in a gradient from white to magenta.

A) Structure of hRio2(1-353)Δ(131-146) monomer. The N- and C-termini are labelled as the ATP-binding pocket.

B) Sequence conservation in all Rio2 orthologues is displayed at the surface of hRio2 and cluster at the ATP-binding pocket.

C) Sequence conservation in eukaryotic Rio2 sequences is displayed at the surface of hRio2. The fingerprint extends around the ATP-binding pocket.

D) and E) Sequence conservation in higher eukaryotes is displayed at the surface of hRio2 structure. The fingerprint covers the entire homodimer interface.

The bar indicates the residue conservation from white (non-conserved) to magenta (identical residues).

Fig. 5: Homodimer formation locks the ATP-binding pocket in an apo form.

ctRio2 and hRio2 structures have been superimposed and are displayed separately.

A) The ATP-binding pocket of ctRio2 in complex with ADP (PDB code 4GYI) (24). Residues involved in the ADP binding and the phospho-aspartate residue (p-D257) are indicated. A Mn^{2+} ion is also shown in beige.

B) Human Rio2 ATP-binding pocket is shown under same orientation as ctRio2. The equivalent residues involved in ADP binding in ctRio2 are displayed. Residues belonging to the adjacent hRio2 monomer are shaded in grey.

References

1. Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A, Gleizes PE (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA* 6(2):225–242.

2. de la Cruz J, Karbstein K, Woolford Jr. JL (2015) Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annu Rev Biochem* 84:93–129.

3. Bassler J, et al. (2001) Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol Cell* 8(3):517–529.

4. Dragon F, et al. (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* 417(6892):967–970.

5. Harnpicharnchai P, et al. (2001) Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol Cell* 8(3):505–515.

6. Milkereit P, Kuhn H, Gas N, Tschochner H (2003) The pre-ribosomal network. *Nucleic Acids Res* 31(3):799–804.

7. Nissan TA, Bassler J, Petfalski E, Tollervey D, Hurt E (2002) 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *Embo J* 21(20):5539–5547.

8. Schafer T, Strauss D, Petfalski E, Tollervey D, Hurt E (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *Embo J* 22(6):1370–1380.

9. Rouquette J, Choesmel V, Gleizes PE (2005) Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. *Embo J* 24(16):2862–2872.

10. Baumas K, et al. (2012) Human RioK3 is a novel component of cytoplasmic pre-40S pre-ribosomal particles. *RNA Biol* 9(2):162–174.

11. Widmann B, et al. (2012) The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. *Mol Biol Cell* 23(1):22–35.

12. Ferreira-Cerca S, Kiburu I, Thomson E, Laronde N, Hurt E (2014) Dominant Rio1 kinase/ATPase catalytic mutant induces trapping of late pre-40S biogenesis factors in 80S-like ribosomes. *Nucleic Acids Res* 42(13):8635–8647.

13. Zemp I, et al. (2009) Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *J Cell Biol* 185(7):1167–1180.

14. LaRonde-LeBlanc N, Guszczynski T, Copeland T, Wlodawer A (2005) Autophosphorylation of Archaeoglobus fulgidus Rio2 and crystal structures of its nucleotide-metal ion complexes. *Febs J* 272(11):2800–2810.

15. Endicott JA, Noble ME, Johnson LN (2012) The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem* 81:587–613.

16. Vanrobays E, et al. (2001) Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. *Embo J* 20(15):4204–4213.

17. Vanrobays E, Gelugne JP, Gleizes PE, Caizergues-Ferrer M (2003) Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae. *Mol Cell Biol* 23(6):2083–2095.

18. Iacovella MG, et al. (2015) Rio1 promotes rDNA stability and downregulates RNA polymerase I to ensure rDNA segregation. *Nat Commun* 6:6643.

19. Angermayr M, Roidl A, Bandlow W (2002) Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression. *Mol Microbiol* 44(2):309–324.

20. Angermayr M, Hochleitner E, Lottspeich F, Bandlow W (2007) Protein kinase CK2 activates the atypical Rio1p kinase and promotes its cell-cycle phase-dependent degradation in yeast. *Febs J* 274(17):4654–4667.

21. Leger-Silvestre I, et al. (2004) The ribosomal protein Rps15p is required for nuclear exit of the 40S subunit precursors in yeast. *Embo J* 23(12):2336–2347.

22. Geerlings TH, Faber AW, Bister MD, Vos JC, Raue HA (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae. *J Biol Chem* 278(25):22537–22545.

23. Wyler E, et al. (2011) Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *Rna* 17(1):189–200.

24. Ferreira-Cerca S, et al. (2012) ATPase-dependent role of the atypical kinase Rio2 on the evolving pre-40S ribosomal subunit. *Nat Struct Mol Biol* 19(12):1316–1323.

25. Knüppel R, et al. (2018) Insights into the evolutionary conserved regulation of Rio ATPase activity. *Nucleic Acids Res* 46(3):1441–1456.

26. Sloan KE, Knox A, Wells G, Schneider C, Watkins NJ (2019) Interactions and activities of factors involved in the late stages of human 18S rRNA maturation. *RNA Biol*:1–15.

27. Heuer A, et al. (2017) Cryo-EM structure of a late pre-40S ribosomal subunit from Saccharomyces cerevisiae. *Elife* 6. doi:10.7554/eLife.30189.

28. Scaiola A, et al. (2018) Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit. *EMBO J* 37(7):e98499.

29. LaRonde-LeBlanc N, Wlodawer A (2004) Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases. *Structure* 12(9):1585–1594.

30. Romier C, et al. (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr* 62(Pt 10):1232–1242.

31. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78(3):1606–1619.

32. Vistica J, et al. (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem* 326(2):234–256.

33. Brautigam CA (2015) Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzym* 562:109–133.

34. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133–144.

35. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64(Pt 1):112–122.

36. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.

37. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62(Pt 9):1002–1011.

38. Bricogne G, et al. (2009) BUSTER, version 2.8.0. *Cambridge, United Kingdom Glob Phasing Ltd*.

39. Fung HY, Fu SC, Brautigam CA, Chook YM (2015) Structural determinants of nuclear export signal orientation in binding to exportin CRM1. *Elife* 4. doi:10.7554/eLife.10034.

40. Fischer U, et al. (2015) A non-canonical mechanism for Crm1-export cargo complex assembly. *Elife* 4. doi:10.7554/eLife.05745.

41. Bliven S, Lafita A, Parker A, Capitani G, Duarte JM (2018) Automated evaluation of quaternary structures from protein crystals. *PLoS Comput Biol* 14(4). doi:10.1371/journal.pcbi.1006104.

42. Duarte JM, Srebniak A, Scharer MA, Capitani G (2012) Protein interface classification by evolutionary analysis. *BMC Bioinformatics* 13:334.

43. Scharer MA, Grutter MG, Capitani G (2010) CRK: an evolutionary approach for distinguishing biologically relevant interfaces from crystal contacts. *Proteins* 78(12):2707–2713.

44. Ashkenazy H, et al. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44(W1):W344-50.

45. Jura N, et al. (2011) Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Mol Cell* 42(1):9–22.

46. Lavoie H, Li JJ, Thevakumaran N, Therrien M, Sicheri F (2014) Dimerization-induced allostery in protein kinase regulation. *Trends Biochem Sci* 39(10):475–486.

47. Cai Z, Chehab NH, Pavletich NP (2009) Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. *Mol Cell* 35(6):818–829.

48. Wan PT, et al. (2004) Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 116(6):855–867.

49. Flayhan A, et al. (2015) The structure of Legionella pneumophila LegK4 type four secretion system (T4SS) effector reveals a novel dimeric eukaryotic-like kinase. *Sci Rep* 5:14602.

50. Larburu N, et al. (2016) Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. *Nucleic Acids Res* 44(17):8465–8478.

51. Sloan KE, Gleizes PE, Bohnsack MT (2016) Nucleocytoplasmic Transport of RNAs and RNA-Protein Complexes. *J Mol Biol* 428(10 Pt A):2040–2059.

52. Mitterer V, et al. (2016) Nuclear import of dimerized ribosomal protein Rps3 in complex with its chaperone Yar1. *Sci Rep* 6:36714.

53. Christie M, et al. (2016) Structural Biology and Regulation of Protein Import into the Nucleus. *J Mol Biol* 428(10 Pt A):2060–2090.

54. Huber FM, Hoelz A (2017) Molecular basis for protection of ribosomal protein L4 from cellular degradation. *Nat Commun* 8:14354.

55. Mitterer V, et al. (2016) Sequential domain assembly of ribosomal protein S3 drives 40S subunit maturation. *Nat Commun* 7:10336.

56. Kiburu IN, LaRonde-LeBlanc N (2012) Interaction of Rio1 kinase with toyocamycin reveals a conformational switch that controls oligomeric state and catalytic activity. *PLoS One* 7(5):e37371.

57. DeLano WL (2014) The PyMOL Molecular Graphics System, Version 1.8. *Schrödinger LLC*:http://www.pymol.org.

58. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta*

Crystallogr D Biol Crystallogr 66(Pt 4):486–501.

Figure 1

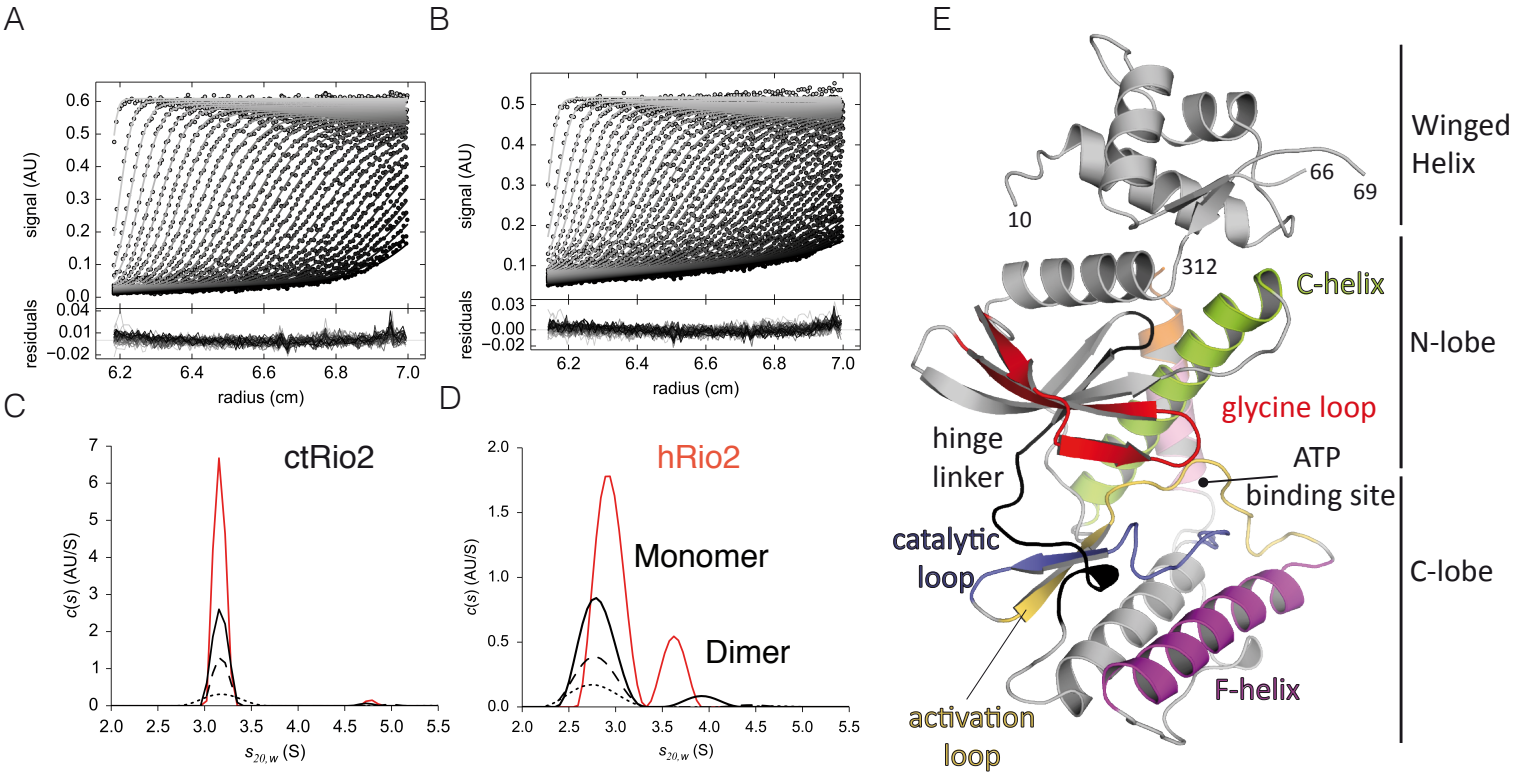

Figure 2

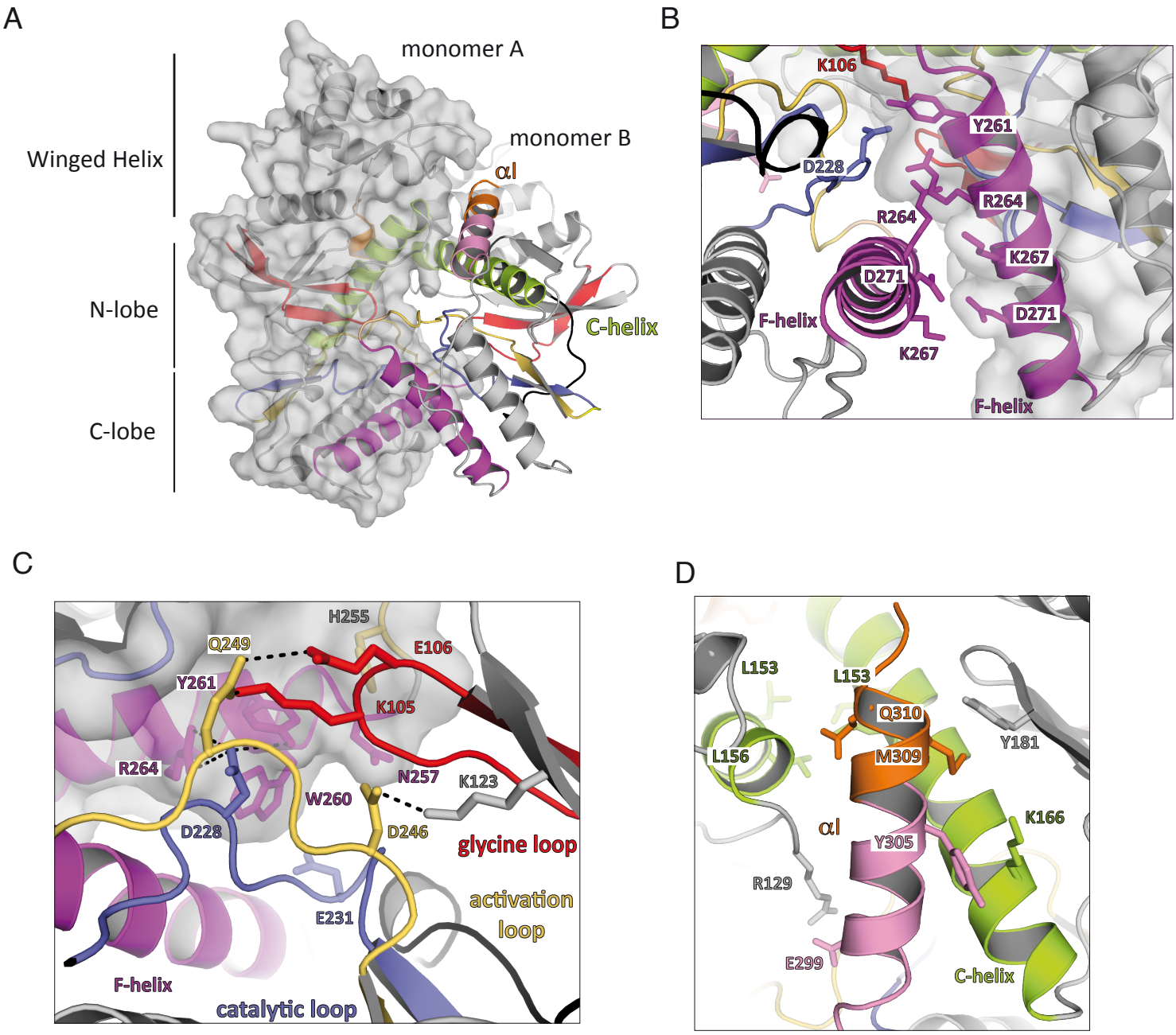

Figure 3

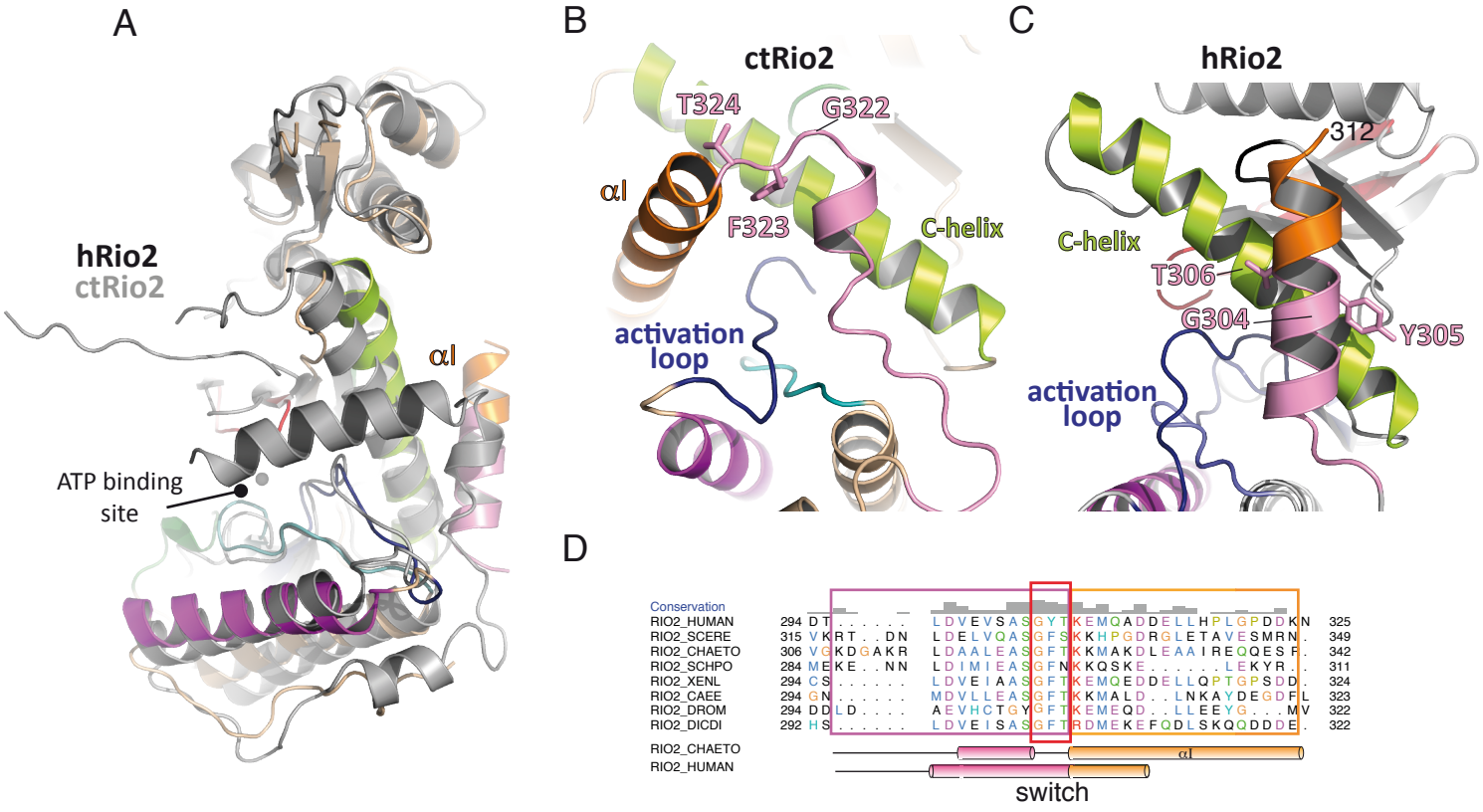

Figure 4

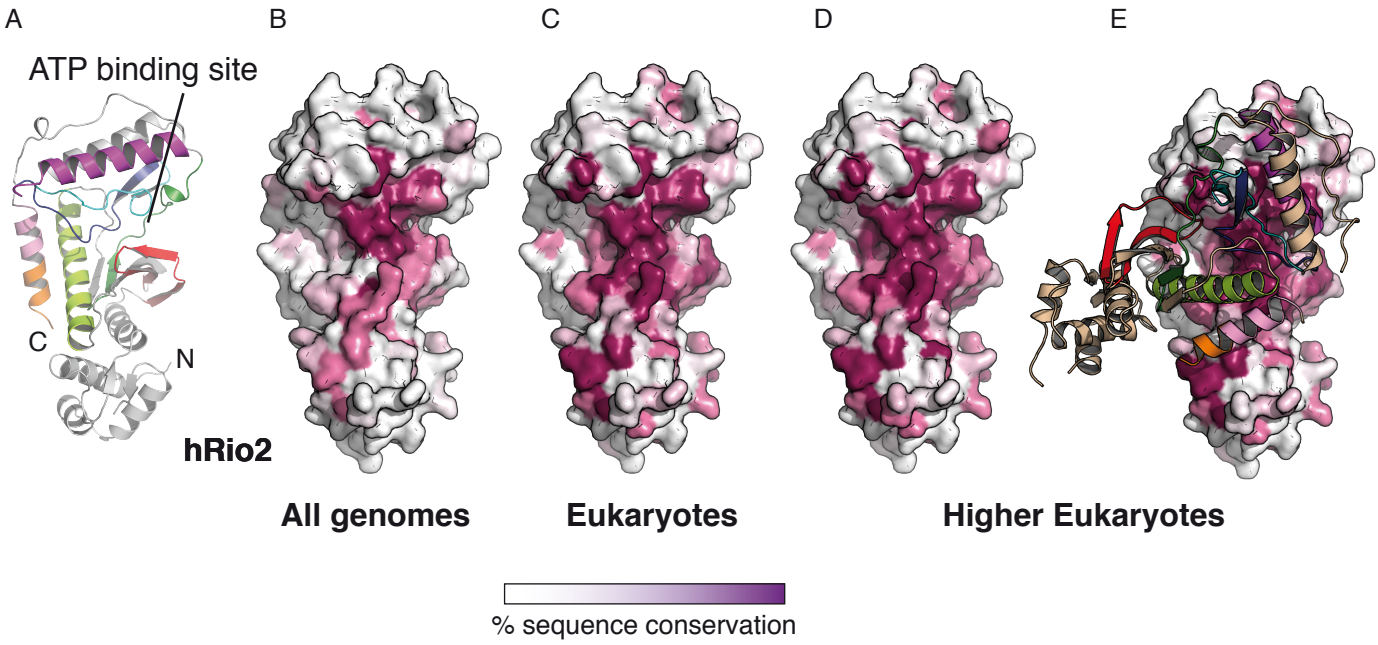

Figure 5

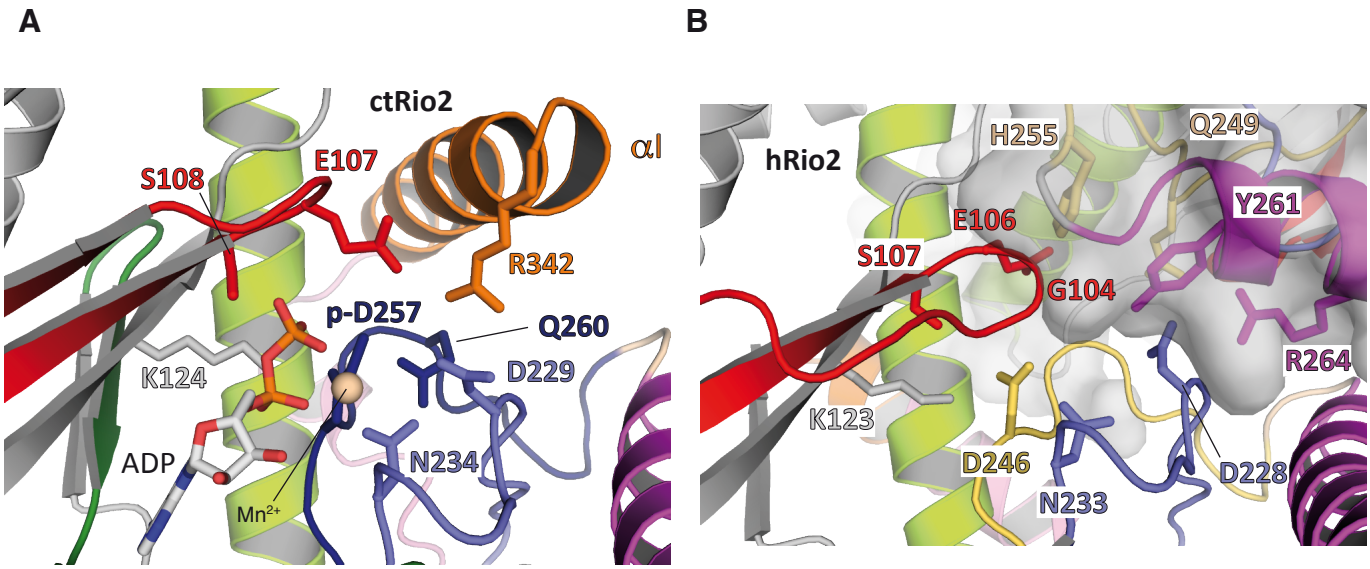

Figure S1

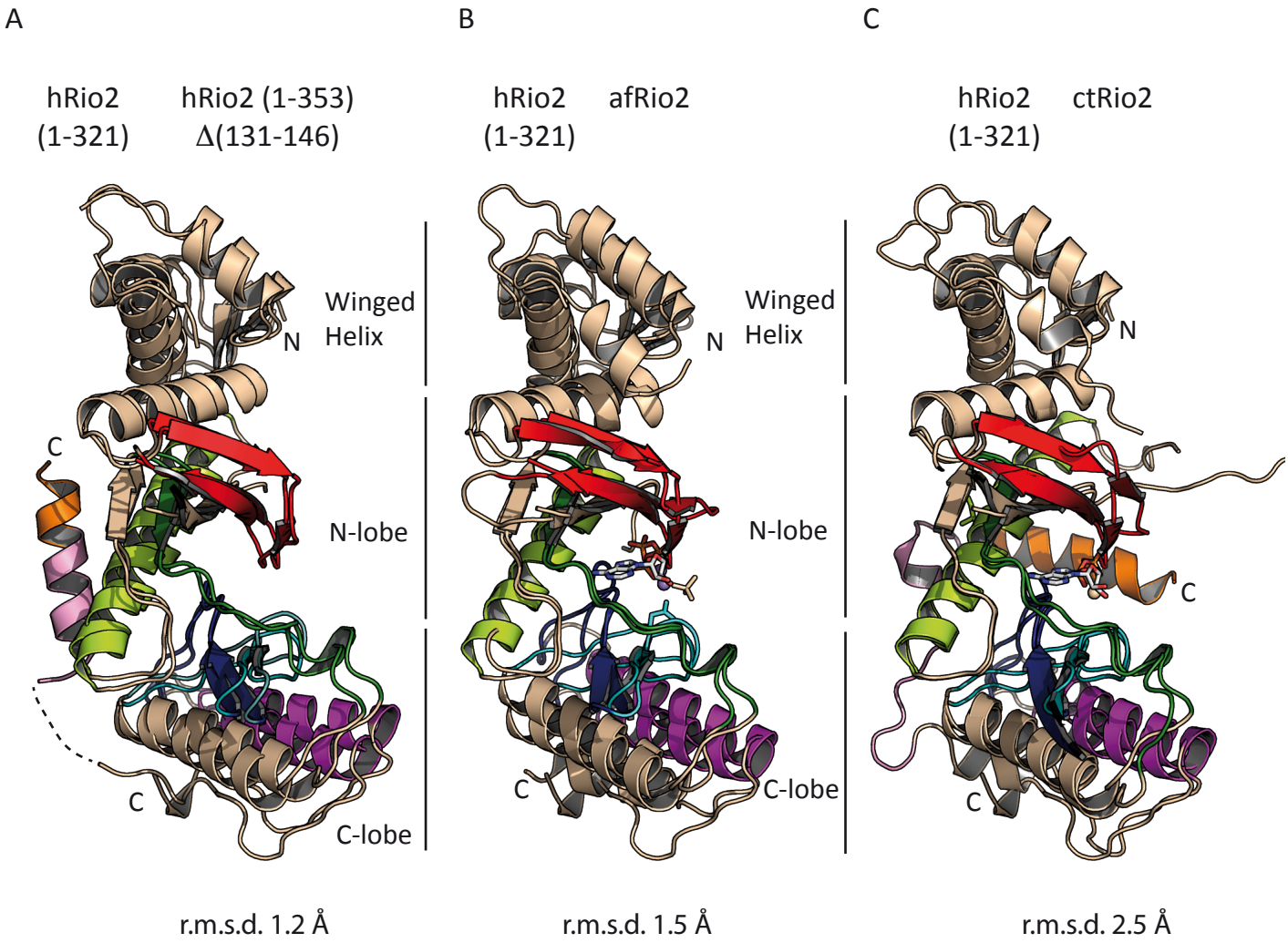

Figure S2

Table 1 Analytical ultracentrifugation studies of human Rio2 and ctRio2.

[hRio2]	**% monomer**	**$s_{20,w}$ monomer**	**% dimer**	**$s_{20,w}$ dimer**	**% oligomer**
3.37 µM	94.7	2.76 S	0.7	4.28 S	4.6
6.75 µM	95.7	2.78 S	1.7	4.35 S	2.6
13.5 µM	91.2	2.80 S	8.5	3.92 S	0.3
27 µM	79.8	2.93 S	20.2	3.62 S	0
[ctRio2]	**% monomer**	**$S_{20,w}$ monomer**	**% dimer**	**$S_{20,w}$ dimer**	**% oligomer**
3.75 µM	96.6	3.18 S	0	-	3.4
7.5 µM	97.3	3.17 S	2.0	4.98	0.7
15 µM	97.6	3.16 S	2.0	4.73	0.4
30 µM	96.6	3.16 S	1.9	4.78	1.5

Percentage of species and $s_{20,w}$ values determined by integration of the peaks observed in the c(s) distributions shown in Figure 1. For hRio2, the integration range was 2.22 to 3.31 S for monomers, 3.32 to 4.49 S for dimers and 4.5 to 20 S for oligomers. For ctRio2, the integration range was 2.72 to 3.69 S for monomers, 4.5 to 5.0 S for dimers and 5.01 to 20 S for oligomers

Table 2 Crystallographic data and refinement statistics.

	hRio2(1-321)Δ(131-146) Se-Met	**hRio2 (1-321)Δ(131-146)**	**hRio2 (1-321)**	**hRio2 (1-353)Δ(131-146)**
Beamline	PX1	PX1	PX1	PX1
Space group	$P2_1$	$P2_1$	$P2_12_12_1$	$P6_5$
Unit cell parameters				
a, b, c (Å)	67.15 99.08 117.48	66.92 98.327 117.069	58.59 86.40 121.54	67.55 67.55 312.19
α, β, γ (°)	90.00 95.05 90.00	90.00 95.22 90.00	90.00 90.00 90.00	90 90 120
Wavelength (Å)	0.9797	0.9801	0.9788	0.9789
Resolution range (Å)	46.10 - 2.37 (2.43-2.37)	41.67 - 2.1 (2.17- 2.10)	70.4- 2.9 (3.07- 2.90)	42.69 - 2.6 (2.69 - 2.6)
Total reflections	232,087 (24,051)	331,294 (31,840)	47,340 (4,662)	279,056 (31,169)
Unique reflections	61,746 (6,173)	87,771 (8,466)	13,696 (1,224)	24,610 (2,417)
Rpim (%)	4.09 (35.8)	2.16 (33.59)	4.06 (35.45)	5.78 (73.21)
I/σI	12.24 (2.26)	20.59 (2.32)	13.81 (2.62)	8.73 (1.49)
Completeness (%)	99.40 (99.37)	98.49 (96.68)	92.17 (88.49)	99.98 (100.00)
Redundancy	3.8 (3.9)	3.8 (3.7)	1.93 (1.90)	11.3 (12.9)
Se sites	40	-	-	-
Wilson Plot B-factor (Å^2)	46.89	42.09	72.28	67.2
CC 1/2	0.997 (0.845)	0.999 (0.879)	0.997 (0.744)	0.991 (0.493)
FOM acentric (before /after solvent flattening)	0.394 / 0.715	-	-	-
Refinement statistics				
Rwork (%) Rfree (%)	19.70 22.89	20.05 23.52	22.58 26.67	21.44 25.00
No. of non-H atoms	9,713	9,792	3,924	4,636
Protein	1,144	1,133	506	565
Water	413	596	-	61
Average B factor (Å^2)	60.13	56.77	77.97	81.0
Ramachandran (%)				

Preferred/allowed/outliers (%)	97.05 / 1.79/ 1.16	99.46 / 0.54/ 0.0	93.99/ 3.86 / 2.15	92.38 / 5.81 / 1.81
Rmsd bond/angle (°)	0.022 / 1.87	0.020 / 1.76	0.008/ 1.03	0.015/ 1.87
Clash score	7.25	5.70	5.28	4.95
PDB	6FDV	6FDM	6FDN	6FDO

Numbers under brackets refer to the highest resolution shell