THE
EMBO
JOURNAL

Manuscript EMBO-2011-79534

Crystal Structure of Cwc2 Reveals a Novel Architecture of a Multipartite RNA-binding Protein

Jana Schmitzova, Nicolas Rasche, Olex Dybkov, Katharina Kramer, Patrizia Fabrizio, Henning Urlaub, Reinhard Luehrmann and Vladimir Pena

Corresponding authors: Vladimir Pena and Reinhard Luehrmann, MPI for Biophysical Chemistry

Review timeline:

Submission date:	18 September 2011
Editorial Decision:	19 October 2011
Revision received:	24 January 2012
Editorial Decision:	14 February 2012
Response (author):	14 February 2012
Accepted:	14 February 2012

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 19 October 2011

Thank you for submitting your manuscript for consideration by The EMBO Journal. I have now received the final report from the three referees who evaluated your study and I enclose their comments below. As you will see two of the referees raise the concern of the recent publication of the same structure of Cwc2, however, it is the policy of The EMBO Journal that papers published after the date of submission are not taken into account when making the final decision. Implementing this policy means that we will continue to consider the study but we require additional experimentation to make it suitable for the journal. Therefore as all the referees point out we would require additional evidence and a model on how cwc2 binds to the U6 RNA for it to be further considered. Should you be able to provide the requested information I would like to invite you to submit a revised version of the manuscript.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process initiative, please visit our website: http://www.nature.com/emboj/about/process.html

Thank you for the opportunity to consider your work for publication. I look forward to your

© European Molecular Biology Organization

revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

The paper Schmitzova et al presents the structure of the splicing factor Cwc2 in its apo form. The structure reveals a very compact structure involving two classical RNA binding domains (a zinc finger and an RRM) which are "glued" together by a so-called Torus domain that wrapped around the zinc-finger. This structure is very compact and represents one of the few RNA binding proteins structure composed of a zinc-finger and an RRM. The authors also show that the N-terminal part of the protein which was crystallized here is sufficient for splicing. The study also revealed that RNA binding (to U6 snRNA) involved not only the RRM but also the zinc-finger and the interdomain linker, forming a clear RNA binding surface on the protein. Single amino-acid mutant in all regions identified lead to decrease in splicing.
Altogether, the paper presents a very nice and original structure which is supported by functional data. Two aspects of the paper diminish the impact of the findings.

1_ A paper describing the exact same structure has just been published in the biochemical journal(Lu et al, Bioch. J. 2011). The structure presented this other paper is from the same organism and with a higher resolution 1.8A compare to 2.4A here .

2_Considering that five cross-linked could be obtained for cwc2 in complex with RNA, couldn't the authors identify on which part of the RNA the different part of cwc2 can crosslink to? Having evidence on how cwc2 is bound to U6 would clearly lift the impact of the present work.

Minor comments:
1_Figure 2A is not shown in stereo
2_part of the discussion on how cwc2 bind RNA could be shortened since this maybe too speculative without evidence on where cwc2 binds the RNA.

Referee #2

Schmitzowa et al. present the crystal structure of a functionally active N-terminal fragment of the yeast Cwc2 protein, a member of the nineteen complex (NTC) that is implicated in organizing the catalytic site of the active spliceosome. The structure reveals a zinc finger domain that is encircled by secondary structure elements (a torus domain) and an RRM domain. The authors show that the deletion variant studied is able to support splicing when added back to Cwc2 depleted extracts. UV-induced protein-RNA cross-links of Cwc2 and U6 snRNA (in lieu of using purified spliceosomes) show RNA contacts of different regions of the protein. The authors show that full-length and the N-terminal fragment bind U6 snRNA in vitro, while the RRM alone has strongly reduced binding affinity. The authors propose that the Cwc2 protein presents a multipartite RNA binding protein that helps to organize the catalytic centre in the active spliceosome.
The manuscript provides novel insight for an important protein in the heart of the spliceosome. The structural data are sound and the functional relevance of the protein studied is provided by splicing assays. A limited set of mutations are presented to support the structural analysis, although a clear picture regarding the proposed RNA-binding interface is not obtained. The topic is highly important and the manuscript suitable for publication. However, the authors should provide additional evidence and mutational analysis to support their conclusions regarding the proposed RNA binding activity of the protein.

© European Molecular Biology Organization

Specific points:

- The authors should provide additional mutational analysis to support their conclusions. The proposal that Cwc2 is a multipartite RNA binding protein should be substantiated by such data. One important variant to study is a deletion of the connector linker, which is proposed to be important for RNA binding and thus function, but can presumably deleted without disrupting the fold of the domain.
- The authors should provide EMSA assays to probe the RNA binding of wild type and mutant proteins. This is mentioned as data not shown but normally is a more direct assay to monitor and semi-quantitate RNA binding, as it is not affected by cross-linking efficiencies and accessibility.
- At several places the authors refer to "data not shown", or manuscript "in preparation", i.e. contacts of Cwc2 to the RNA (p.18), "earlier structure-function investigations" (p.19). This is not acceptable as a number of these data are important and would further strengthen the manuscript.
- Fig. 1 should include a schematic picture of the full-length Cwc2 protein indicating the different domains and the construct crystallized.
- The discussion is too long and in part very speculative. It should be more concise.

Referee #3

Crystal Structure of Cwc2 Reveals a Novel Architecture of a Multipartite RNA-binding Protein
Schmitzova et al.

Cwc2, a component of the multi-protein "nineteen complex" in yeast, is believed to interact with U6 snRNA, and has been shown to be an essential splicing factor; the current manuscript along with previous work suggests that this factor is involved in the catalytic activation of the spliceosome.

The submitted paper reports an X-ray structural analysis of a proteolytically stable and central fragment of Cwc2. The crystal structure was solved by single anomalous dispersion made possible by the presence of a protein-bound zinc. Notable features include a predicted RRM, the metal binding Zinc finger of the CCCH class, a massive feature, referred to by the authors as a "torus", that surrounds the Zn finger, and a protruding loop. In addition to their structural analysis, the authors performed add-back experiments to Cwc2-depleted yeast extracts to analyze the roles of both domains and individual amino acid residues in splicing assays. They also examined the binding of Cwc2 to U6 snRNA and mapped RNA binding sites on the protein by a combination of crosslinking and mass spectrometric analysis.

Similar work to that reported here was just published by Yin and coworkers in the Biochemical Journal. I would note that the functional analysis in the current manuscript is more comprehensive including the crosslinking data.

This paper presents a novel structure of a key splicing component of great interest. The authors' analysis of the structure is elegant and insightful. Although the biochemical and mapping experiments are fine for what they are, the authors over-interpret their data as described below (given issues of specificity). There needs to be considerable revision in terms of the interpretation of the non-structure related results. Also, I believe this paper would be strengthened by including some of the unpublished data referred to within the submission.

Major points
1. Although others (including Yin et al) have published data showing interaction of isolated U6 snRNA with Cwc2, there is nothing in this (or other papers) to indicate the specificity of this interaction. This is a major weakness in the presentation of that data and should be addressed. I have little difficulty in accepting that the crosslinking/MS data reveals RNA binding surfaces but the relevance to specific interaction with U6 snRNA is unclear. Where are the control crosslinking/mapping experiments (or gel shift - see below)

2. Related to the above, no mention is made of the multiple complexes observed in the EMSA experiments. Do these represent multiple binding to U6 snRNA (probably and hence non-specific interaction) or some rearrangement of U6 structure? I also believe that the crosslinking experiments

© European Molecular Biology Organization

were performed at ~6x the concentration required to see the highest shift.

3. This paper contains an extensive (and somewhat confusing - see below) list of references to a manuscript in preparation - nine occurrences by my count. This includes extensive references to data not shown. I think it imperative to include at least some of this data in the current paper. For example:

"It was recently shown by crosslinking and chemical structure-probing, that in isolated catalytically active spliceosomes Cwc2 simultaneously contacts several elements of their catalytic RNA network, including the U6 ISL, the intron upstream of the base-paired U6-ACAGAGA/5'SS element, and also part of U2/U6 helix Ia (Rasche et al, in preparation)."

Other references to data from a paper in preparation should be removed or modified.

Minor points

1. Various aspects of references to the paper in preparation are unclear:

"The structure of Cwc2 thus provides strong support for the hypothesis, based upon our earlier structure-function investigations (Rasche et al, in preparation)..."

Earlier? From a paper in preparation? Again, this needs to be fixed or deleted at several points in the current manuscript.

1st Revision - authors' response 24 January 2012

Point-by-point Response

Referee 1:

1. A paper describing the exact same structure has just been published in the biochemical journal (Lu et al, Bioch. J. 2011). The structure presented in this other paper is from the same organism and with a higher resolution 1.9A compare to 2.4A here.

Lu et al. have used for their crystallographic analyses a crystal form that differs from the one reported by us. The two crystal forms grow in different conditions and the cell parameters are also different {Lu, 2012 #239}. Nevertheless, the difference in resolution between our structure and the one reported by Lu et al. does not account for a distinct interpretation of the biological relevance of the structure. Because we succeeded to solve the structure at 2.4 Å resolution, we did not screen for other crystal forms in order to gain resolution. All structural features that we present and discuss in our manuscript are well-defined in the electron density map at a resolution of 2.4 Å. Our crystal form has the advantage that the connector element is stabilized in crystal contacts, allowing us to trace it entirely in the electron density map. The connector element could be only partially revealed in the crystal form reported by Lu et al. 2012.

In the Lu et al. paper, insight into RNA recognition by Cwc2 is achieved by performing rational mutagenesis exclusively based on their structural analysis, while in our work we provide an

experimental means (i.e., RNA crosslinking followed by mass spectrometry) that allows us to precisely identify the crosslinked amino acids in Cwc2 and thus those sites that contact RNA. The gel shift assays presented in the two papers are consistent with one another, but mostly complementary rather than overlapping. In addition, after the mass spectrometric identification of the residues crosslinked to RNA, we perform mutagenic analyses followed by splicing complementation assays, which further allows us to assess the relevance of the structure-function analysis in the context of intact spliceosomes. Indeed Referee 3 also points out: "Similar work to that reported here was just published by Yin and coworkers in the Biochemical Journal. I would note that the functional analysis in the current manuscript is more comprehensive including the crosslinking data." In short, our manuscript clearly provides additional, highly significant information that is not found in Lu et al.

2. Considering that five cross-links could be obtained for cwc2 in complex with RNA, couldn't the authors identify on which part of the RNA the different part of cwc2 can crosslink to? Having evidence on how cwc2 is bound to U6 would clearly lift the impact of the present work.

Currently, mass spectrometry methods do not allow us to map the position of crosslinks on a complex RNA species in an RNA-protein crosslink as not enough RNA sequence information can be obtained. Reverse transcription analyses could shed some light on crosslinked nucleotides in U6 but would not provide information regarding what region of Cwc2 is crosslinked. To gain insight into which part of Cwc2 binds which region of the U6 snRNA, we now performed cross-linking with a binary complex consisting of the U6 intramolecular-stem-loop (ISL) and Cwc2. Interestingly, U6-ISL was crosslinked frequently to both RNP2 and RNP1 of Cwc2's RRM, and to a lesser extent (compared to U6) to its zinc finger and connector element, while no crosslinks were observed for the Torus domain. This result suggests that the RRM domain has a binding preference towards the U6-ISL, while the ZnF, Torus domain and connector element preferentially interact with other U6 snRNA nucleotides at least in a binary system.
Our experiments performed with Cwc2 and free U6 snRNA allow the identification of the Cwc2 sites that bind RNA in isolation. The biological relevance of these sites was deduced indirectly based on mutagenesis and splicing assays. Previous studies examining binary interactions of Cwc2 with RNA indicate that it binds all spliceosomal snRNA and the pre-mRNA with similar affinity, indicating that Cwc2 does not exhibit sequence specific binding per se (McGrail et al., 2009). However, in the presence of intact spliceosomes, Cwc2 contacts exclusively U6 snRNA (McGrail et al., 2009). Recent crosslinking experiments coupled with primer extension analyses performed in our lab with purified spliceosomes (Rasche et al., 2012), allowed the identification of U6 snRNA bases that are specifically contacted by Cwc2 within the spliceosome. The recognized bases are located in a region upstream of the U6 ACAGAGA box and in the U6 intramolecular-stem-loop (ISL). Unfortunately, despite our extensive efforts, it has not been possible to identify whether the same regions of Cwc2 contact the U6 snRNA in the spliceosome as those observed in vitro with

© European Molecular Biology Organization

purified Cwc2 and U6 snRNA, due to the low amounts of purified yeast spliceosomes that could be obtained coupled with low crosslinking efficiencies.

Minor comments:
1. Figure 2A is not shown in stereo

We have downsized this figure and it is now shown in stereo.

2. Part of the discussion on how cwc2 binds RNA could be shortened since this maybe too speculative without evidence on where cwc2 binds the RNA.

We have shortened the discussion from 19453 characters in the first version to 14624 characters in the revised version.

Referee 2:
Specific points:
- The authors should provide additional mutational analysis to support their conclusions. The proposal that Cwc2 is a multipartite RNA binding protein should be substantiated by such data. One important variant to study is a deletion of the connector linker, which is proposed to be important for RNA binding and thus function, but can presumably be deleted without disrupting the fold of the domain.

To substantiate that Cwc2 acts as a multipartite RNA binding protein, we have performed additional mutational analyses (as requested by the referee) targeted to the RNA crosslinked sites within Cwc2 (except for C87 whose structural role makes it unsuited for mutagenesis studies) followed by electrophoretic mobility shift assays (EMSA). These new results are shown in Figure 5C. Although single point mutations had little or no affect on the affinity of Cwc2 for U6 (Supplemental Figure S6), two of the newly tested double mutants had reduced binding affinity. Specifically, the double mutants C181A/Y138A and Y120/Y138A exhibited reduced U6 RNA binding affinity, demonstrating that Y120 from the connector element and C181 from RNP2 are important for RNA binding. Taken together, the data demonstrate that in addition to the RRM, the connector element also contributes to RNA binding, supporting the idea that Cwc2 functions as a multipartite RNA-binding unit. We would also note that Lu et al. performed detailed mutational analyses that demonstrates the crucial role of the connector for RNA binding in vitro {Lu, 2011 #210}. Specifically, they showed that, although single point mutations do not have any effect, a quadruple mutant in the connector (R131A, K132A, K133A, K135A) nearly abolished RNA binding as assayed by EMSA. In addition, a nick in the connector linker leads to a dramatic decrease in RNA binding, demonstrating that the integrity of the connector is essential for RNA binding by Cwc2.

- The authors should provide EMSA assays to probe the RNA binding of wild type and mutant proteins. This is mentioned as data not shown but normally is a more direct assay to monitor and semi-quantitate RNA binding, as it is not affected by cross-linking efficiencies and accessibility.

The EMSA assays mentioned previously as data not shown indicated that single point mutations in Cwc2 do not have a significant effect on RNA binding. In the current version, we have included these data in a new figure in the supplementary material (Figure S6).

- At several places the authors refer to "data not shown", or manuscript "in preparation", i.e. contacts of Cwc2 to the RNA (p.18), "earlier structure-function investigations" (p.19). This is not acceptable as a number of these data are important and would further strengthen the manuscript.

This manuscript is currently accepted for publication and now available online. {Rasche, 2012 #238}.

- Fig. 1 should include a schematic picture of the full-length Cwc2 protein indicating the different domains and the construct crystallized.

We now include a schematic diagram in Figure 3A that shows Cwc2 domains and the crystallized region.

- The discussion is too long and in part very speculative. It should be more concise.

We have shortened the discussion from 19453 characters in the first version to 14624 characters in the revised manuscript.

Referee 3:

Major points

1. Although others (including Yin et al) have published data showing interaction of isolated U6 snRNA with Cwc2, there is nothing in this (or other papers) to indicate the specificity of this interaction. This is a major weakness in the presentation of that data and should be addressed.

Our binary interaction studies with purified Cwc2 and U6 snRNA were carried out primarily to identify potential RNA binding surfaces in Cwc2. Indeed, it was shown previously by EMSA that Cwc2 binds several snRNAs and pre-mRNA, indicating that Cwc2 does not possess sequence specificity per se (McGrail et al., 2009). However, in the presence of intact spliceosomes, Cwc2 contacts solely the U6 snRNA (McGrail et al., 2009), in particular several discrete bases in the

intramolecular-stem-loop (ISL) and upstream of the U6 ACAGAGA box, plus one base from the pre-mRNA {Rasche, 2012 #238}. This indicates that Cwc2 binds with high specificity when U6 snRNA is correctly integrated within the RNA-RNA network of the spliceosome. Our interpretation is that the specificity of the Cwc2 - U6 snRNA interaction is achieved from the accurate positioning of Cwc2 within the spliceosome (possibly mediated by other splicing factors) and probably from the existence of U6 snRNA in a pre-folded state (under the constraints exerted by other spliceosomal components). The work of Rasche et al. 2012, which was not published when this manuscript was first submitted, demonstrates the specificity of Cwc2 – U6 snRNA interaction in the context of purified spliceosomes. The work of Rasche et al. is now accepted and available online.
In order to directly address whether Cwc2 exhibits some kind of RNA sequence preference in isolation, we have now performed comparative crosslinking experiments with the binary complexes Cwc2 - U6 snRNA, Cwc2-U6-ISL and Cwc2 - U4 snRNA. These experiments, which are described on page 13 of the revised manuscript and summarized in Supplementary Table 2, show that the peptides from Cwc2's RRM domain are crosslinked with similar intensities to U6, U6-ISL and U4 snRNA, while the peptides from the ZnF and connector element show strong crosslinking to U6 snRNA and very weak or no crosslinking to U4 snRNA or the U6-ISL (Supplementary Table 2). These results provide a first indication that, at least in a binary system, the RRM does not differentiate between U6 and U4 snRNA, while the connector element and the ZnF could potentially exhibit a binding preference towards regions of the U6 snRNA other than its ISL.

2. I have little difficulty in accepting that the crosslinking/MS data reveals RNA binding surfaces but the relevance to specific interaction with U6 snRNA is unclear.
Where are the control crosslinking/mapping experiments (or gel shift - see below)?

See our response to point 1 of referee 3 above.

3. Related to the above, no mention is made of the multiple complexes observed in the EMSA experiments. Do these represent multiple binding to U6 snRNA (probably and hence non-specific interaction) or some rearrangement of U6 structure?

The large excess of Cwc2 used in our EMSAs might trigger the formation of oligomers, although we cannot exclude re-arrangements of the U6 structure. A similar decrease in electrophoretic mobility with increasing Cwc2 concentration was also reported by McGrail et al. 2009 and by Lu et al. 2012. We now include the following sentence on page 6 of the revised manuscript: " Note that the decrease in electrophoretic mobility of the Cwc2-U6 RNP complex with increasing Cwc2 concentration may be due to oligomerization or changes in the structure and/or surface exposed charge of the complex, and was also observed in previous studies {McGrail, 2009 #71}{Lu, 2012 #239}”. To investigate, whether multiple complexes might form in solution, we have mixed Cwc2 and U6 snRNA in a molar ratio of 10:1 and subjected the mixture to size exclusion chromatography. The profile shows the formation of a complex that runs as a single peak and whose apparent

molecular weight corresponds to about 260 kDa (Figure S5). However, in the EMSA experiments Cwc2 is added at a 100-fold higher excess than used for size exclusion chromatography (Cwc2:U6snRNA = 1000:1) and thus we cannot rule out that larger complexes containing more than one copy of Cwc2 might form. As our crosslinking experiments were performed under conditions (Cwc2: U6snRNA = 30:1) similar to those used for size exclusion chromatography, we would stress that the Cwc2 crosslinking sites identified were most likely obtained from stoichiometric Cwc2-U6 complexes.

I also believe that the crosslinking experiments were performed at ~6x the concentration required to see the highest shift.

Crosslinking experiments were performed at a Cwc2 concentration of 25 µM, which means ~3x the concentration required to see the highest shift at EMSA (8 µM). However, the molar excess of Cwc2 versus RNA is ~330x higher in EMSA (1000:1) as compared to the crosslinking experiments (30:1). The ratio between Cwc2 and RNA in the crosslinking experiments (30:1) is much closer to the one used for size exclusion chromatography (10:1), where the complex runs as a single and monodisperse peak.

3. This paper contains an extensive (and somewhat confusing - see below) list of references to a manuscript in preparation - nine occurrences by my count. This includes extensive references to data not shown. I think it imperative to include at least some of this data in the current paper. For example:" It was recently shown by crosslinking and chemical structure-probing, that in isolated catalytically active spliceosomes Cwc2 simultaneously contacts several elements of theircatalytic RNA network, including the U6 ISL, the intron upstream of the base-paired U6-ACAGAGA/5'SS element, and also part of U2/U6 helix Ia (Rasche et al, in preparation)."Other references to data from a paper in preparation should be removed or modified.

This manuscript is currently accepted for publication and is now available online. {Rasche, 2012 #238}.

Minor points

1. Various aspects of references to the paper in preparation are unclear:

"The structure of Cwc2 thus provides strong support for the hypothesis, based upon our earlier structure-function investigations (Rasche et al, in preparation)..."Earlier? From a paper in preparation? Again, this needs to be fixed or deleted at several points in the current manuscript.

All references to the work of Rasche et al. have been changed accordingly.

2nd Editorial Decision 14 February 2012

Thank you for sending us your revised manuscript.

Our original referee 1 has now seen it again, and you will be pleased to learn that in his/her view you have addressed all criticisms in a satisfactory manner and that the paper will now be publishable in The EMBO Journal.

There is one remaining issue raised by the referee with respect to the cross-eye stereo view of figure 2A. Could you send us a new amended version of the figure via e-mail, please? We will then upload the file for you.

Thank you very much again for considering our journal for publication of your work.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

Although the revised version did not provide information on the precise protein-RNA contact of Cwc2 with U6 RNA, the revised version is acceptable in my view.

Just one minor comment:
the figure 2A is still not visible as a cross-eye stereo view.

Response (author) 14 February 2012

Please find attached the amended version of figure 2. I decreased the distance between the left and right sided images to 4.3 cm, so that anyone can see it as stereo.

© European Molecular Biology Organization