JMB

COMMUNICATION

The Three-dimensional Architecture of the EJC Core

M. Elizabeth Stroupe[1]†, Thomas Ø. Tange[2]†, Dennis R. Thomas[1,2] Melissa J. Moore[1,2] and Nikolaus Grigorieff*

[1]*Howard Hughes Medical Institute, Brandeis University 415 South Street, Waltham MA 02454, USA*

[2]*Department of Biochemistry Brandeis University 415 South Street, Waltham MA 02454, USA*

*Corresponding author

The exon junction complex (EJC) is a macromolecular complex deposited at splice junctions on mRNAs as a consequence of splicing. At the core of the EJC are four proteins: eIF4AIII, a member of the DExH/D-box family of NTP-dependent RNA binding proteins, Y14, Magoh, and MLN51. These proteins form a stable heterotetramer that remains bound to the mRNA throughout many different cellular environments. We have determined the three-dimensional (3D) structure of this EJC core using negative-stain random-conical tilt electron microscopy. This structure represents the first structure of a DExH/D-box protein in complex with its binding partners. The EJC core is a four-lobed complex with a central channel and dimensions consistent with its known RNA footprint of about ten nucleotides. Using known X-ray crystallographic structures and a model of three of the four components, we propose a model for complex assembly on RNA and explain how Y14:Magoh may influence eIF4AIII's RNA binding.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: spliceosome; exon junction complex; mRNA; electron microscopy; three-dimensional structure

Introduction

The exon junction complex (EJC) is a set of proteins deposited on mRNA by the process of pre-messenger RNA (pre-mRNA) splicing.[1] EJCs are deposited ~20 nucleotides upstream of exon–exon junctions and accompany spliced mRNAs to the cytoplasm where they participate in localization, translation, and decay of the bound mRNA.[2] This protein complex remains stably bound to the mRNA through different cellular environments. One factor that contributes to its tight association is eIF4AIII, a key component of the EJC core.[3–6] Three other factors are part of this stable platform, the Y14: Magoh heterodimer[7–10] and MLN51;[5,11] this αβγδ heterotetramer is a landing pad for many additional transiently interacting factors.[12]

When simultaneously overexpressed in mammalian cells, eIF4AIII, Y14:Magoh, and MLN51 form a stable complex of about 170 kDa amenable to purification and biochemical analysis.[12] Core formation can also be recapitulated *in vitro* using recombinant proteins individually expressed in and purified from *Escherichia coli. In vitro* core assembly requires both single-stranded RNA and ATP, resulting in a footprint of about ten nucleotides, indistinguishable from that observed for full EJCs deposited by pre-mRNA splicing.[1,13]

eIF4AIII is a member of the DExH/D-box protein family. These proteins constitute a large and important family of NTP-dependent nucleic acid binding proteins that have been implicated in almost every step of RNA metabolism.[14–16] The entire family is often referred to as the DExH/D-box family of RNA helicases because some family members exhibit RNA helicase activity *in vitro*. In a few cases, this RNA unwinding activity reflects protein translocation along single-stranded RNA. Additionally, some DExH/D-box proteins can disrupt RNA–protein interactions and/or catalyze RNA annealing[14–18] Nonetheless, their roles *in vivo* are poorly defined because they are found in the context of large macromolecular machines, such as the EJC, where their activities are modified by a host of interacting protein factors.[19] Our understanding of how these accessory proteins modulate their helicase partners

is limited at present by a lack of information about their quaternary structure.

In general, DExH/D-box family members contain two RecA-like motifs tethered by a flexible linker.[15] eIF4AIII, about 50 kDa in mass, contains all the canonical motifs (Q–VI) that define the DEAD-box subfamily. It also contains eight eIF4AIII-specific motifs (A–H) that map to variable surface loops in a structural model based on an X-ray crystallographic structure of another family member, eIF4AI.[20] Like other family members, eIF4AIII exhibits RNA-stimulated ATPase activity;[13,20,21] however, this activity appears to be dispensable for EJC assembly.[20] eIF4AIII thus represents a new functional class of DExH/D-box proteins that act as RNA placeholders or "clothes pins", rather than helicases or RNPases.[6]

Within the EJC core, eIF4AIII and MLN51 contribute to RNA binding, as both can be UV cross-linked to the RNA.[6,11,13] MLN51, an 80 kDa protein with no apparent sequence homology to other proteins, contains an N-terminal motif that exhibits non-specific RNA binding activity. This region, dubbed the SeLoR motif, also directs MLN51 variants to sub-nuclear speckle domains enriched in splicing factors. Within the EJC, MLN51 interacts directly with eIF4AIII; the nature of this interaction depends on the MLN51 construct, suggesting a redundant interface between the two proteins.[11]

Both *in vivo* and *in vitro*, Y14 and Magoh associate to form a tight heterodimer with a combined molecular mass of about 40 kDa. Although Y14 possesses a canonical RNA recognition motif (RRM), extensive cross-linking experiments failed to show any interaction between it and RNA.[22] This is consistent with Y14's putative RRM being buried completely at its dimer interface with Magoh.[23–25] Rather than binding RNA, the main role of the Y14: Magoh heterodimer in the EJC core is to inhibit ATP hydrolysis, effectively locking eIF4AIII into a conformation that cannot release bound RNA.[13]

To better understand how all four EJC core components work together to effect tight, sequence independent anchoring onto RNA, we determined the three-dimensional (3D) structure of the native EJC core purified from human cells. To our knowledge, this is the first example of a DExH/D-box protein in complex with its binding partners. Analysis of the negatively stained EJC core imaged in the electron microscope allowed us to identify heterogeneity that correlates to degradation of MLN51 in a subset of the particles. The identification of this alteration was confirmed by gel filtration analysis and independent reconstruction of EJC cores containing C-terminally truncated MLN51, allowing us to identify this portion of MLN51. Together with extensive mutagenesis of eIF4AIII, we
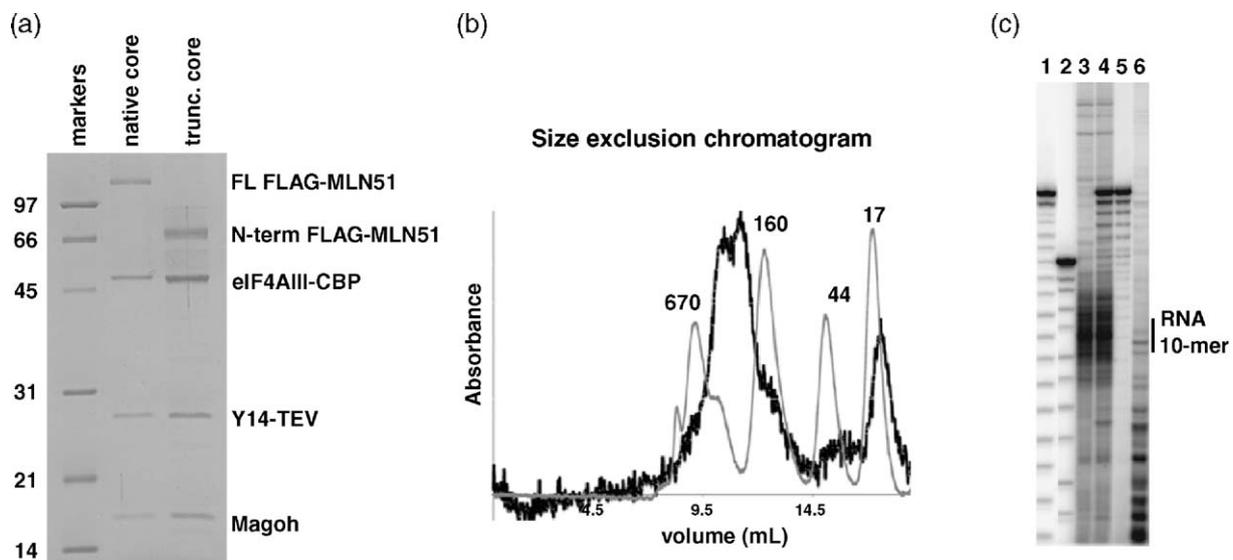
Figure 1. EJC core purification and analysis. (a) SDS-13% (w/v)PAGE analysis of 10 μl native EJC cores or those formed with intentionally C-terminally truncated MLN51. Staining with colloidal Coomasie G250 (Sigma) shows that the four EJC core proteins, Y14:Magoh, eIF4AIII, and MLN51 were purified with approximate stoichiometry *via* a split TAP tag where the protein A moiety was fused to the C terminus of Y14 and the CBP moiety was fused to the C terminus of eIF4AIII.[12] (b) Gel filtration analysis of the EJC core. Native EJC core was passed over a Superdex 200 10/300 GL (Amersham) size exclusion column (black). A double peak eluted between the 670 kDa and 160 kDa markers (gray) (BioRad 151–191). According to Western blot analysis, EJC core components Magoh, Y14, eIF4AIII, and FLAG-MLN51 only elute in this major peak (Supplementary Data, Figure S1). Western blotting revealed that the doublet is due to truncation of MLN51 (Supplementary Data, Figure S1). The molecular mass of each standard is given in kDa. (c) Determination of RNA content of the EJC core. Lane 1, 21-mer ssRNA hydrolysis ladder; 2, 15-mer ssRNA hydrolysis ladder; 3, (γ-³²P)-labeled EJC core RNA; 4, (γ-³²P)-labeled EJC core RNA + ssRNA; 5, (γ-³²P)-labeled ssRNA alone; 6, (γ-³²P)-labeled EJC core RNA + labeled ssRNA + RNaseA. The EJC core was purified as described[12] but the lysate was treated with micrococcal nuclease in place of RNase A before being heated to release proteins, cooled, and treated with PNK and [γ-³²P]ATP. Parallel end-labeling reactions of the EJC core were done with and without a known amount of ssRNA. One reaction was treated with RNase A after labeling. Samples were phenol/chloroform extracted, precipitated, resuspended in formamide buffer and separated on a 27%(w/v) denaturing gel alongside RNA hydrolysis markers.

propose a model for how eIF4AIII, Y14:Magoh, and MLN51 assemble on RNA to form the EJC core.

Expression, Purification, and Analysis of EJC Cores

EJC core recombinantly expressed from HEK293T cells was purified through two affinity steps using a TAP tag[26] whose protein A moiety/TEV cleavage site was fused to the C terminus of Y14 and whose calmodulin binding peptide (CBP) moiety was fused to the C terminus of eIF4AIII[12] (Figure 1(a)). Gel filtration analysis partitioned the native sample in a broad peak that eluted between the 670 kDa and 160 kDa size markers (Figure 1(b)). Western blotting with antibodies against eIF4AIII, Y14, Magoh, or the N-terminal FLAG tag on MLN51 detected all expected proteins in this peak (Supplementary Data, Figure S1). Nonetheless, species with electrophoretic mobility between that of full-length MLN51 and eIF4AIII that also reacted with the α-FLAG antibody were present in the latter half of the peak, suggesting that some of the particles contained C-terminally truncated MLN51 (Supplementary Data, Figure S1). No EJC core components were detected in the minor peak that elutes after the low-molecular mass marker (data not shown).

Previous analyses of both fully-assembled EJCs purified from HeLa cell nuclear extract[1] and the tetrameric EJC core assembled from recombinant proteins *in vitro*[13] revealed that the EJC has a footprint of about ten nucleotides. Stable assembly of the core *in vitro* is absolutely dependent on incorporation of RNA.[13] To determine whether our *in vivo*-derived EJC core also contained RNA, EJC cores purified from micrococcal nuclease-treated lysate were heat denatured and then incubated with T4 polynucleotide kinase and [γ-^{32}P]ATP. Denaturing gel electrophoresis revealed RNA fragments of predominantly 9–12 nucleotides (Figure 1(c)), indicating that this recombinantly expressed EJC core bound RNA in the cell and carried that RNA through purification. The RNA content in the purified sample appeared approximately stoichiometric with the protein subunits, based on protein concentration estimates from colloidal Coomassie stained SDS-PAGE gels and RNA concentrations calculated from similar-length synthetic RNA molecules combined with the EJC core labeling reaction (Figure 1(a) and (c)).
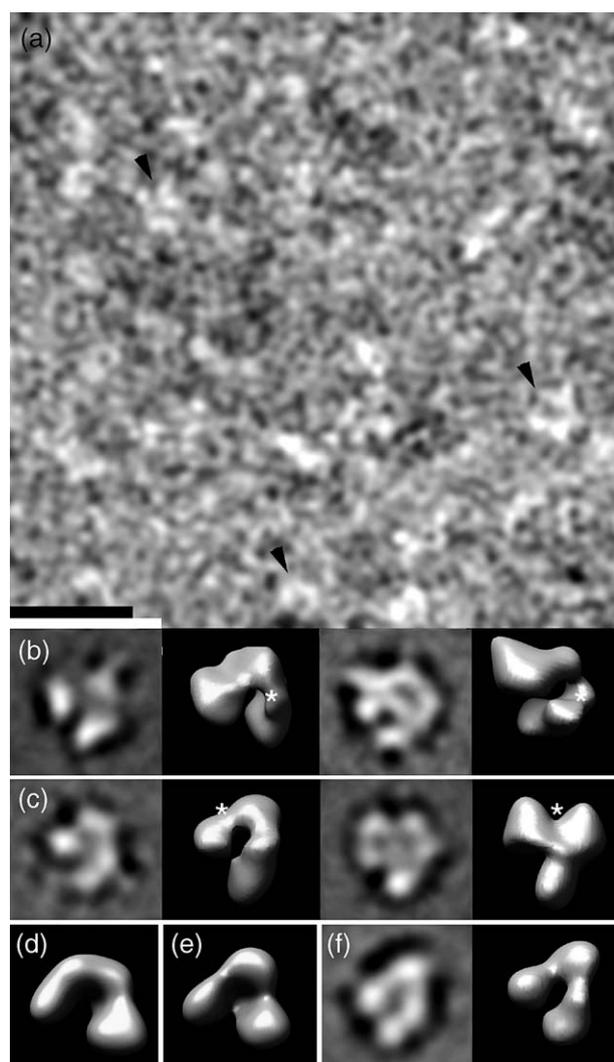
Figure 2. Images and random-conical tilt structures of the EJC core with full-length MLN51 and truncated MLN51. (a) A field of native EJC cores negatively stained with 1%(w/v) uranyl formate over a glow discharged thin carbon film was imaged at about 1 μm under focus using a Philips CM12 electron microscope at 120 keV. Images were recorded at 60,000 times magnification and the micrographs were scanned at a resolution of 7 μm per pixel on a Zeiss scanner to give a resolution of 1.17 Å per pixel on the specimen. Particles were picked using the pattern-matching algorithm in Signature[32] based on rotationally-averaged class averages from manually picked particles aligned and classified in IMAGIC.[33] In this way, particles were selected based on their diameter. The computationally identified particles were then manually screened in WEB[34] to eliminate any spurious particles. The length of the black scale bar represents 100 nm and its width 10 nm for the view of the particle field. Particles are marked with arrowheads. To enhance contrast for display purposes, the image has been low-passed filtered at 50 Å resolution in SPIDER.[34] (b) Two representative class averages yielding four-lobed native structures calculated using K-means classification in SPIDER[34] from the native data alongside their corresponding structures shown in a similar orientation. The structures were calculated using the random-conical tilt method[27] from pairs of particles imaged at 40° and 0° tilt. The length of the white scale bar represents 15 nm for the class averages and structures. The asterisk denotes the fourth lobe unique to this subset of structures. (c) Two representative class averages yielding three-lobed native structures calculated as in (b) alongside their corresponding structures shown in a similar orientation. The asterisk denotes the missing fourth lobe of density that identifies this subset of structures. (d) Structure determined by aligning and averaging 20 structures similar to those in (b). (e) Structure determined by aligning and averaging seven structures similar to those in (c). (f) A representative class average calculated as for (b) from particles harboring deliberately truncated MLN51 and its corresponding structure calculated from the tilted particles.

Random Conical Tilt Negative-Stain Electron Microscopy

EJC core negatively stained on a thin carbon film and imaged in the electron microscope appear monodisperse and uniformly sized, with a characteristic U or V shape (Figure 2(a)). Initial structures calculated from particle image classes determined in the random-conical tilt method[27] revealed that there were two predominant species, one with four lobes (Figure 2(b)) and one with three (Figure 2(c)). Both structures were about 90 Å in diameter with a central hole of about 10 Å diameter. Initial reconstructions were calculated separately for the four- (Figure 2(d)) and three-lobed (Figure 2(e)) structures by aligning and combining those that were similar, which accounted for only a subset of the particles. The untilted images were aligned to a single reference dataset generated from each of the two structures to determine the structure to which they belonged. In the end, 9655 particles aligned to the four-lobed structure and 3892 particles to the three-lobed structure. The tilted and untilted particles were then split into two datasets based on this alignment and the rotational and translational parameters for each particle were iteratively aligned to reference projections[28] calculated from the corresponding starting model to generate final structures (Figure 3(a) and (b)). Two independently determined reconstructions from two different native datasets resulted in identical structures,
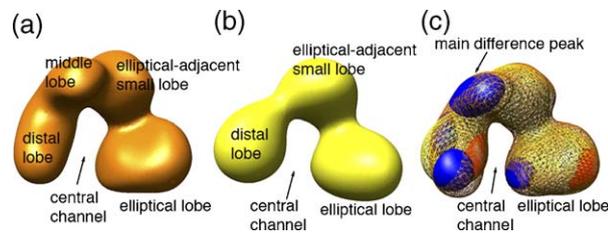
Figure 3. The structure of the native EJC core and mapping of MLN51's C terminus. (a) The EJC core at 28 Å resolution is a four-lobed structure with a central 10 Å wide channel. Three of the four lobes are about the same size (30 Å across) and the fourth is a long, elliptical domain whose long axis runs parallel to the page in the orientation shown. (b) The EJC core formed with C-terminally truncated MLN51 (Figure 1(a)) shares many features with the native EJC core, including the 10 Å wide central channel and the elongated elliptical domain. (c) Difference maps calculated between the aligned and scaled native and mutated EJC cores show that the most significant absence occurs in the middle, small lobe; however, small changes occur in the shape of the distal small lobe and along the interior side of the elliptical domain. The structures were aligned by centering each, based on its center of mass, and then the three common lobes were rotationally aligned. The middle lobe was masked out from the four-lobed structure so the Fourier amplitudes calculated from the two structures of different mass could be scaled accurately within each resolution zone by a least-squares algorithm. Once scaled, the structures were subtracted from one another and negative (blue) and positive (red) density is highlighted in the difference map.

affirming our four-lobed native structure. The resolution for both structures was estimated to be about 28 Å using the Fourier shell correlation criterion at a value of 0.5 and 23 Å at a value of 0.143 (Supplementary Data, Figure S2).

With these structures in hand, and with the aim of dissecting the location of each protein component within the core, an EJC core lacking the dispensable 400 amino acid residue C-terminal domain of MLN51 was expressed and purified in the same manner as the parent complex (Figure 1(a)). Random conical tilt data were collected to produce an independent reconstruction of this truncated core. Class averages were similar to those of the original particle, but the resulting random-conical tilt reconstruction revealed only structures similar to the three-lobed complex (Figure 2(f)). The rotational and translational parameters for all of the untilted and tilted particles were then iteratively refined against projections calculated from the 3D reconstruction to generate a final structure of the core containing C-terminally truncated MLN51 from tilt pairs of 16,416 particles (Figure 3(b)). The resolution for this structure was estimated to be about 20 Å using the Fourier shell correlation criterion at a value of 0.5 and 15 Å at a value of 0.143 (Supplementary Data, Figure S2). The increased resolution of this structure compared with the native structure confirms the reduced heterogeneity of this particle population.

Organization of the EJC Core

The RNA-bound EJC core provides a protein platform to which a multitude of transiently interacting factors bind. The constellation of these transient EJC factors is subject to dynamic exchange as the mRNA moves to different subcellular locations.[2,12] In general, a protein platform that contacts many other factors would be expected to have a large surface area. Indeed, the EJC core containing full-length MLN51 forms an extended, asymmetric U-shaped complex with a central channel of about 10 Å diameter, providing a large surface area for protein–protein or protein–nucleic acid interactions (Figure 3(a)).

At 28 Å resolution, one can discern four discrete domains in the major structural species (Figure 3(a)). Three are of similar size and shape, spheres of about 30 Å diameter. The fourth domain is an ellipsoid whose short axis is also about 30 Å wide and whose long axis is about 80 Å in length. Assuming a protein density of 0.735 Da/Å[3,29] these four lobes together account for about 190 kDa mass, which corresponds well to the calculated mass of 170 kDa for the protein components of the complex. The short RNA molecule associated with the structure contributes only about 3 kDa to this mass. The minor structural species in the native preparations that has only three lobes is an asymmetric V-shaped molecule (Figure 2(e)). The largest domain and two of the three smaller domains are similar in shape and size to those in the four-lobed structure. The central small domain is smaller and

protrudes in the opposite direction from the elliptical domain (Figure 2(c)). EJC cores harboring intentionally C-terminally truncated MLN51 are virtually identical to the minor species of native EJC cores (Figure 2(f) and Figure 3(b)). In this smaller structure, the large ellipsoidal domain is also slightly smaller than in the native complex.

Direct comparison of the EJC core harboring truncated MLN51 to the four-lobed native structure by difference density mapping highlights three areas of difference (Figure 3(c)). First, and most striking, is the missing middle small lobe. This is a real absence in the smaller structure and the negative peak is not paired with positive difference density elsewhere in the map that might indicate a conformational change. Further, the structure of the EJC core with truncated MLN51 accounts for about 130 kDa; the 60 kDa mass difference between the wild-type and truncated structures corresponds well with the expected missing 40 kDa of truncated MLN51. In contrast, a pair of negative and positive peaks at opposite ends of the elliptical domain and another pair in the distal small domain suggest that in removing the middle lobe, other small conformational changes occur in the structure. Based on the significance of these peaks, we propose that MLN51 could link disparate parts of the EJC core, and that its C terminus resides predominantly in the middle lobe of the three smallest domains.

Placement of the EJC Core Factors into the Em Density

MLN51 and eIF4AIII interact *via* their respective N-terminal domains. The N-terminal 399 amino acid residues of Barentsz (*Drosophilia* MLN51) were shown to interact with *Drosophilia* eIF4AIII in a two-hybrid assay[5]; a major site of this interaction has been mapped to the SeLoR motif (amino acid residues 137–283).[13] Sequence swapping experiments further revealed that replacement of a section of the N-terminal domain of eIF4AIII with the analogous region of eIF4AI disrupts this eIF4AIII–SeLoR interaction.[13] Within this region, mutation of eIF4AIII-specific motif C disrupts an *in vitro* interaction between eIF4AIII and the N-terminal 400 amino acid residues of MLN51.[20] However, the same motif C mutations failed to disrupt full EJC formation either *in vivo* or *in vitro*,[20] indicating additional stabilizing contacts in the full-length proteins.

X-ray crystal structures for the Y14:Magoh heterodimer[23–25] and a reasonable atomic model for eIF4AIII[20] facilitated their docking into our structure. Y14 and Magoh form a tight pyramidal complex with a diameter of about 40 Å. Given the large buried surface that forms their interface, the Y14:Magoh heterodimer likely remains intact within the EJC core and we modeled it into our EM density map as such. In contrast, DEAD-box family proteins' characteristic ATP and RNA-binding motifs are distributed between linked N and C-terminal domains that can adopt open or closed conformations. Upon binding

either ligand, the domains coalesce to form the ATP and RNA binding sites. eIF4AIII binds both unhydrolyzed ATP and RNA within the EJC core.[13] Therefore, we modeled the N and C-terminal domains of eIF4AIII in a closed conformation based on that observed in the X-ray crystallographic structure of a DEAD-box protein homolog from *Methanococcus jannaschii* (*mj*DEAD).[30] *mj*DEAD's topology is similar to eIF4AIII, with two globular RecA-like domains and no intervening regions.[30] As crystallized, *mj*DEAD is bound to neither ATP nor RNA but adopts a compact conformation similar to that seen in the X-ray crystallographic structures of the DNA-bound DEAD-box NS3 helicase from the hepatitis C virus[31] and the DNA-bound DNA helicase PcrA.[32]

Additional constraints on fitting eIF4AIII and Y14:Magoh into the structure of the EJC core were provided by previous protein interaction data and the structural complementarity of its components. Pull-down analyses implicated a eIF4AIII-specific motif H, which lies in the C-terminal domain just next to the active site cleft and near the linker that joins the two domains, as part of the interaction site for the Y14:Magoh heterodimer.[20] The Y14:Magoh heterodimer has only three major surfaces through which it might interact with this region of eIF4AIII (Figure 4(a)). One is the convex surface of Y14, which is unlikely to fit onto the equally convex surface of eIF4AIII's C-terminal domain. The second
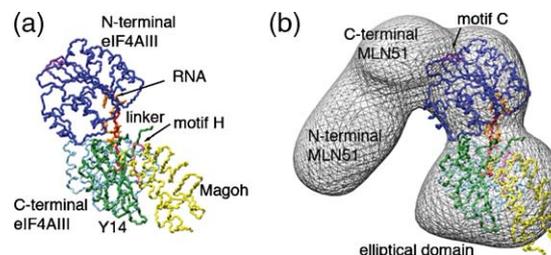
Figure 4. Docking of eIF4AIII and Y14:Magoh into the native EJC structure. (a) eIF4AIII N and C-terminal domains (dark blue and light blue, respectively) modeled from X-ray crystallographic structures of yeast eIF4AI and *mj*DEAD[20,30,36] docked onto the X-ray crystallographic structure of the Y14:Magoh heterodimer[23–25] (green and yellow, respectively). Docking was performed manually using the Xfit module from the XtalView suite[35] and UCSF Chimera[37] and based on mutational studies[13,20], as well as surface complementarity between the C-terminal domain of eIF4AIII and Y14:Magoh, which are known to interact at motif H (pink).[13,20] The flexible linker connecting eIF4AIII's N and C-terminal domains (red) is near Y14:Magoh. eIF4AIII's RNA binding site points slightly into the plane of the page and touches both its N and C-terminal domains, demarked by superposition of an oligonucleotide (orange) from the X-ray crystallographic structure of PcrA.[33] (b) As modeled, the C-terminal domain of eIF4AIII and Y14:Magoh (colored as for (a)) fit well into the largest, elliptical domain of the 3D structure of the EJC core (gray mesh). In this orientation, the N terminus of eIF4AIII projects into the adjacent small domain, presenting motif C, which is important for MLN51 binding,[13,20] towards the MLN51 binding site.

is formed by Magoh's large β-sheet and is also an unlikely binding partner for eIF4AIII because they show little surface complementarity. Opposite Magoh's β-sheet, however, is a large, concave surface that fits well onto the face of eIF4AIII containing motif H. By placing Y14:Magoh in this orientation, the heterodimer falls near eIF4AIII's linker; this may explain how binding of Y14:Magoh inhibits eIF4AIII ATPase activity by locking it in the clamped conformation.

When docked together in this way, eIF4AIII's C-terminal domain and Y14:Magoh fit into the elliptical domain of the EJC core structure (Figure 4(b)). Some of the elliptical domain remains empty, regardless of how Y14:Magoh are placed on eIF4AIII's C terminus. In the difference maps between the native and the truncated core (Figure 3(c)), there is a peak in this same region that could be filled by parts of MLN51. The N terminus of eIF4AIII docks nicely into the adjacent small lobe, with motif C projecting towards the middle small domain. This arrangement of subunits would place the RNA binding groove of eIF4AIII on the opposite side from the EJC core's largest domain, running perpendicular to the long axis of the ellipse (Figure 4(b)). We propose that the remaining density corresponds to MLN51, whose C terminus would sit in the middle domain and whose N terminus would occupy the small lobe distal to the largest domain (Figure 4(b)). This density includes a mass of 90 kDa, which corresponds well with the known mass of full-length MLN51. MLN51 would have access to the RNA where MLN51 and eIF4AIII come into contact at the vertex of the three-lobed complex.

Conclusions

To better understand how eIF4AIII, Y14:Magoh, and MLN51 form a complex that serves as a platform for the binding of other transiently interacting factors and is locked onto spliced mRNAs in a sequence independent manner, we determined the 3D structure of the EJC core (Figure 3(a)). By truncating MLN51, we were able to localize the C-terminal region of MLN51 and to confirm the identification of a low level of contaminant present in our native preparation (Figures 2 and 3). By combining difference mapping of these structures with previously acquired muta-genesis data[13,20], we propose a means by which eIF4AIII interacts with Y14:Magoh to enforce a compact arrangement of eIF4AIII's two domains and allow MLN51 to interact with both eIF4AIII and the RNA (Figure 4).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.05.049

References

1. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860–6869.
2. Tange, T. O., Nott, A. & Moore, M. J. (2004). The ever-increasing complexities of the exon junction complex. *Curr. Opin. Cell Biol.* **16**, 279–284.
3. Chan, C. C., Dostie, J., Diem, M. D., Feng, W., Mann, M., Rappsilber, J. & Dreyfuss, G. (2004). eIF4A3 is a novel component of the exon junction complex. *RNA*, **10**, 200–209.
4. Ferraiuolo, M. A., Lee, C. S., Ler, L. W., Hsu, J. L., Costa-Mattioli, M., Luo, M. J. *et al.* (2004). A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proc. Natl Acad. Sci. U. S. A.* **101**, 4118–4123.
5. Palacios, I. M., Gatfield, D., St Johnston, D. & Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature*, **427**, 753–757.
6. Shibuya, T., Tange, T. O., Sonenberg, N. & Moore, M. J. (2004). eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nature Struct. Mol. Biol.* **11**, 346–351.
7. Kataoka, N., Diem, M. D., Kim, V. N., Yong, J. & Dreyfuss, G. (2001). Magoh, a human homolog of *Drosophila mago* nashi protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J.* **20**, 6424–6433.
8. Kim, V. N., Yong, J., Kataoka, N., Abel, L., Diem, M. D. & Dreyfuss, G. (2001). The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. *EMBO J.* **20**, 2062–2068.
9. Le Hir, H., Gatfield, D., Braun, I. C., Forler, D. & Izaurralde, E. (2001). The protein Mago provides a link between splicing and mRNA localization. *EMBO Rep.* **2**, 1119–1124.
10. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997.
11. Degot, S., Le Hir, H., Alpy, F., Kedinger, V., Stoll, I., Wendling, C. *et al.* (2004). Association of the breast cancer protein MLN51 with the exon junction

complex via its speckle localizer and RNA binding module. *J. Biol. Chem.* **279**, 33702–33715.

12. Tange, T. O., Shibuya, T., Jurica, M. S. & Moore, M. J. (2005). Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA*, **11**, 1869–1883.

13. Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B. & Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nature Struct. Mol. Biol.* **12**, 861–869.

14. Silverman, E., Edwalds-Gilbert, G. & Lin, R. J. (2003). DExD/H-box proteins and their partners: helping RNA helicases unwind. *Gene*, **312**, 1–16.

15. Caruthers, J. M. & McKay, D. B. (2002). Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123–133.

16. Cordin, O., Banroques, J., Tanner, N. K. & Linder, P. (2006). The DEAD-box protein family of RNA helicases. *Gene*, **367**, 17–37.

17. Fairman, M. E., Maroney, P. A., Wang, W., Bowers, H. A., Gollnick, P., Nilsen, T. W. & Jankowsky, E. (2004). Protein displacement by DExH/D "RNA helicases" without duplex unwinding. *Science*, **304**, 730–734.

18. Yang, Q. & Jankowsky, E. (2005). ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry*, **44**, 13591–13601.

19. Delagoutte, E. & von Hippel, P. H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Quart. Rev. Biophys.* **36**, 1–69.

20. Shibuya, T., Tange, T. O., Stroupe, M. E. & Moore, M. J. (2006). Mutational analysis of human eIF4AIII identifies regions necessary for exon junction complex formation and nonsense-mediated mRNA decay. *RNA*, **12**, 360–374.

21. Li, Q., Imataka, H., Morino, S., Rogers, G. W., Jr, Richter-Cook, N. J., Merrick, W. C. & Sonenberg, N. (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell. Biol.* **19**, 7336–7346.

22. Reichert, V. L., Le Hir, H., Jurica, M. S. & Moore, M. J. (2002). 5′ exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. *Genes Dev.* **16**, 2778–2791.

23. Fribourg, S., Gatfield, D., Izaurralde, E. & Conti, E. (2003). A novel mode of RBD-protein recognition in the Y14-Mago complex. *Nature Struct. Biol.* **10**, 433–439.

24. Lau, C. K., Diem, M. D., Dreyfuss, G. & Van Duyne, G. D. (2003). Structure of the Y14-Magoh core of the exon junction complex. *Curr. Biol.* **13**, 933–941.

25. Shi, H. & Xu, R. M. (2003). Crystal structure of the *Drosophila* Mago nashi-Y14 complex. *Genes Dev.* **17**, 971–976.

26. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M. & Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218–229.

27. Radermacher, M. (1988). Three-dimensional reconstruction of single particles from random and non-random tilt series. *J. Electron Microsc. Tech.* **9**, 359–394.

28. Penczek, P. A., Grassucci, R. A. & Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, **53**, 251–270.

29. Quillin, M. L. & Matthews, B. W. (2000). Accurate calculation of the density of proteins. *Acta Crystallog. sect. D*, **56**, 791–794.

30. Story, R. M., Li, H. & Abelson, J. N. (2001). Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proc. Natl Acad. Sci. USA*, **98**, 1465–1470.

31. Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A. *et al.* (1998). Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure*, **6**, 89–100.

32. Chen, J. Z. & Grigorieff, N. (2006). SIGNATURE: a particle selection system for molecular electron microscopy. *J. Struct. Biol.* In the press.

33. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.

34. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199.

35. McRee, D. E. (1999). XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.

36. Caruthers, J. M., Johnson, E. R. & McKay, D. B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl. Acad. Sci.* **97**, 13080–13085.

37. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.