The EMBO Journal (2010) 29, 2205–2216 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing

Ryan N Jackson[1], A Alejandra Klauer[2,4], Bradley J Hintze[1,4], Howard Robinson[3], Ambro van Hoof[2] and Sean J Johnson[1,*]

[1]Department of Chemistry and Biochemistry, Utah State University, Logan, UT, USA, [2]Department of Microbiology and Molecular Genetics, University of Texas Health Science Center-Houston, Houston, TX, USA and [3]Department of Biology, Brookhaven National Laboratory, Upton, NY, USA

The essential RNA helicase, Mtr4, performs a critical role in RNA processing and degradation as an activator of the nuclear exosome. The molecular basis for this vital function is not understood and detailed analysis is significantly limited by the lack of structural data. In this study, we present the crystal structure of Mtr4. The structure reveals a new arch-like domain that is specific to Mtr4 and Ski2 (the cytosolic homologue of Mtr4). *In vivo* and *in vitro* analyses demonstrate that the Mtr4 arch domain is required for proper 5.8S rRNA processing, and suggest that the arch functions independently of canonical helicase activity. In addition, extensive conservation along the face of the putative RNA exit site highlights a potential interface with the exosome. These studies provide a molecular framework for understanding fundamental aspects of helicase function in exosome activation, and more broadly define the molecular architecture of Ski2-like helicases.

The EMBO Journal (2010) **29,** 2205–2216. doi:10.1038/emboj.2010.107; Published online 28 May 2010
Subject Categories: RNA; Structural Biology
Keywords: exosome activation; RNA helicase; RNA processing; TRAMP; X-ray crystallography

Introduction

RNA surveillance is a fundamental quality control process that prevents aberrantly transcribed or modified RNAs, poorly assembled ribonucleoprotein complexes (RNPs), and unneeded RNAs from interfering with normal cellular gene expression and regulation (Lebreton and Seraphin, 2008; Houseley and Tollervey, 2009). Defects in RNA processing have been linked to a variety of disease states, including neurodegenerative diseases and cancer (Bruserud, 2007; Nelson and Keller, 2007). A critical component of nuclear RNA surveillance is the nuclear exosome complex that contains two exoribonuclease active sites and an endoribonuclease site (Allmang *et al*, 1999b; Jensen and Moore, 2005; Lacava *et al*, 2005; Vanacova *et al*, 2005; Wyers *et al*, 2005; Liu *et al*, 2006; Dziembowski *et al*, 2007; Lebreton and Seraphin, 2008; Lebreton *et al*, 2008; Anderson and Wang, 2009; Houseley and Tollervey, 2009; Lykke-Andersen *et al*, 2009; Schaeffer *et al*, 2009; Schneider *et al*, 2009).

Although the exosome contains the catalytic activities for RNA processing and degradation, its *in vivo* activity requires many cofactors, including Mtr4. Mtr4 (for mRNA transport; also known as Dob1) was identified in a genetic screen for yeast mutants that accumulated polyadenylated RNAs in the nucleus (Liang *et al*, 1996), and was subsequently observed to affect rRNA processing and other exosome functions (de la Cruz *et al*, 1998). Mtr4 may affect some exosome functions in isolation, but is also the largest component of the TRAMP complex, a three-protein complex containing a poly(A) polymerase (Trf4 or Trf5), a putative RNA-binding protein (Air1 or Air2), and Mtr4 (Jensen and Moore, 2005; Lacava *et al*, 2005; Vanacova *et al*, 2005; Wyers *et al*, 2005; Anderson and Wang, 2009). The TRAMP complex identifies RNA substrates and adds a short poly(A) tail to the 3′ end, thereby promoting their degradation (Lacava *et al*, 2005; Anderson and Wang, 2009). The substrates of TRAMP are wide ranging and include tRNAs, snoRNAs, snRNAs, ncRNAs, rRNAs, some mRNAs, and cryptic unstable transcripts (Allmang *et al*, 1999a; van Hoof *et al*, 2000; Milligan *et al*, 2005; Buhler *et al*, 2007; Houseley *et al*, 2007, 2008; Reis and Campbell, 2007; Wang *et al*, 2008). Mtr4 and the exosome are also involved in chromatin remodelling (Houseley *et al*, 2007, 2008; Reis and Campbell, 2007; San Paolo *et al*, 2009) and in normal processing of various RNAs that require 3′ end trimming to reach a mature state (e.g. 5.8S rRNA, snoRNAs, snRNAs, and some mRNAs; de la Cruz *et al*, 1998; Allmang *et al*, 1999a; van Hoof *et al*, 2000). A homologue of Mtr4, Ski2, is also involved in the activation of the exosome, but is located in the cytoplasm and acts primarily on mRNA transcripts (Lebreton and Seraphin, 2008). Mtr4 and Ski2 homologues have been identified in a wide variety of eukaryotes including humans (Anderson and Wang, 2009; Houseley and Tollervey, 2009), indicating that their roles in RNA surveillance, processing, and decay are conserved.

Mtr4 is essential for yeast growth (Liang *et al*, 1996) and provides a critical link between polyadenylation of RNA substrates by Trf4 and degradation by the exosome. A point mutant that disrupts helicase activity of Mtr4 results in the accumulation of polyadenylated RNAs that are no longer removed by the RNA degradation machinery (Wang *et al*, 2008). *In vitro* analysis indicates that Mtr4 has RNA-dependent ATPase and helicase activity (Bernstein *et al*, 2008; Wang *et al*, 2008). The role of Mtr4 in exosome-mediated RNA decay presumably includes unwinding of RNA secondary structure and/or displacement of proteins associated with RNP complexes to present a 'clean' substrate to the exosome (Lebreton and Seraphin, 2008; Houseley and Tollervey, 2009). In addition, it has been suggested that the ATPase activity of

*Corresponding author. Department of Chemistry and Biochemistry, Utah State University, 0300 Old Main Hill, Logan, UT 84322-0300, USA. Tel.: +1 435 797 2089; Fax: +1 435 797 3390;
E-mail: sean.johnson@usu.edu
[4]These authors contributed equally to this work

Received: 9 February 2010; accepted: 4 May 2010; published online: 28 May 2010

Mtr4 and Ski2 may be used to feed RNA substrates into the ring-like structure of the exosome, analogous to ATPases associated with the proteasome (van Hoof and Parker, 1999; Lorentzen and Conti, 2006).

Helicases are ubiquitous and diverse enzymes that are broadly classified into families on the basis of a distinct set of sequence motifs localized in core helicase domains, which are generally the site of nucleic acid and nucleotide binding (Cordin *et al*, 2006; Jankowsky and Fairman, 2007; Singleton *et al*, 2007; Pyle, 2008). Mtr4 is designated as a superfamily II RNA helicase belonging to the Ski2-like family of DExH/D RNA helicases (de la Cruz *et al*, 1999). The most closely related member of this family is the Ski2 protein (38% identity), from which the family derives its name. More distantly related Ski2-like RNA helicases include Brr2 (RNA splicing) and Slh1 (translation initiation; Pena *et al*, 2009).

A molecular understanding of Mtr4 function is significantly limited because no structures are currently available for Mtr4 or Ski2. The only available structures of a Ski2-like RNA helicase are C-terminal fragments of Brr2 (Pena *et al*, 2009; Zhang *et al*, 2009), which have limited sequence similarity to Mtr4. Although Ski2-like DNA helicase structures have been characterized more extensively (Buttner *et al*, 2007; Richards *et al*, 2008; Zhang *et al*, 2008; Oyama *et al*, 2009), the 122-kDa Mtr4 is considerably larger than these DNA helicases (e.g. Hel308 is 79 kDa) and the similarity between these proteins is not apparent beyond the core helicase domains. Only three domains are identifiable in the Mtr4 sequence: the two core helicase domains and an additional C-terminal domain, designated DSHCT (Dob1/Ski2/HelY C-terminal domain; Staub *et al*, 2004), which has no obvious sequence similarity to domains of known structure or function.

In this study, we report the 3.4 Å crystal structure of full-length Mtr4 from *Saccharomyces cerevisiae*. This structure provides, to the best our knowledge, the first view of an exosome-activating cofactor, and also represents the first full-length description of a Ski2-like RNA helicase. The structure clarifies which molecular features are shared throughout the Ski2-like helicase family, and highlights features that are unique to Mtr4 and Ski2. Surprisingly, the Mtr4 structure reveals a prominent domain that adopts a new arch-like appearance. The removal of the arch domain produces a defect in rRNA processing, which resembles the inactivation of the Rrp6 exonuclease component of the exosome. The data presented here suggest that Mtr4 performs an important role in the regulation of exosome activity. Additionally, Mtr4 employs some mechanisms that are general to the Ski2-like family of RNA helicases and others that are unique to Mtr4 and Ski2.

Results and discussion

Structure determination

The full-length Mtr4 protein from *S. cerevisiae* was expressed in *Escherichia coli* using a construct containing a six-histidine affinity tag at the N-terminus. The recombinant protein was purified by Ni-chelate, heparin affinity and gel filtration chromatography. Crystallization conditions were identified and optimized to obtain a 3.4-Å native data set and 3.6-Å selenomethionine (Se-met) substituted data set. As no effective molecular replacement model was available, the Se-met data was used to solve the phase problem using single-wavelength anomalous dispersion (SAD) methods. The initial model was built using the SAD maps and subsequently refined against the 3.4 Å native data to a final R/R_{free} of 29.3/33.3% (Table I).

Table I Data collection, phasing and refinement statistics

	Mtr4 Se-met	Mtr4 native
Data collection		
Space group	$P3_121$	$P3_121$
Cell dimensions		
a, *b*, *c* (Å)	133.4, 133.4, 192.1	133.5, 133.5, 190.9
α, β, γ (deg)	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution (Å)	30–3.6 (3.73–3.60)[a]	30–3.4 (3.52–3.40)
R_{sym}	0.120 (0.689)	0.059 (0.494)
$I/\sigma I$	18.5 (1.9)	31.0 (1.9)
Completeness (%)	97.6 (78.9)	94.0 (59.9)
Redundancy	13.5 (7.6)	10.0 (5.8)
Refinement		
Resolution (Å)		30–3.4
No. reflections		26 097
R_{work}/R_{free}		0.293/0.333
No. atoms		
Protein		6487
Ligand/ion		20
Water		0
B-factors		
Protein		157.9
Ligand/ion		171.2
RMSD		
Bond lengths (Å)		0.011
Bond angles (deg)		1.530

[a]Values in parentheses are for highest-resolution shell.
One crystal was used for each data set.

The electron-density maps are quite clean and continuous, considering the modest resolution of the data (Figure 1). The gaps in the current model include 85 residues at the N-terminus and 11 loops of varying length for which no interpretable electron density is observed (Figure 1D). The anomalous signal from selenomethionine sites, which constitute ~3% of the Mtr4 sequence and are relatively well distributed throughout the protein, was used to assign amino-acid sequence. The alignments of the Hel308 (Buttner *et al*, 2007; Richards *et al*, 2008) and Hjm (Oyama *et al*, 2009) DNA helicase structures to the Mtr4 structure were used to further clarify sequence assignment and connectivity through portions of the structure. Consequently, we were able to confidently model 85% (918 residues) of the Mtr4 sequence and four phosphate molecules into the electron density.

Overall description

The crystal structure reveals that Mtr4 is composed of five distinct domains (Figure 1). The base of the structure contains four domains (domains 1, 2, 3, and 4) assembled in a circular manner, with a channel of ~12 Å diameter passing through the centre. In addition, a fifth domain is a large (265 residue) new structure that arches over the other four domains. This domain is inserted into the middle of domain 3 and is therefore named domain 3a, or the 'arch' domain. A brief description of each of these domains follows.

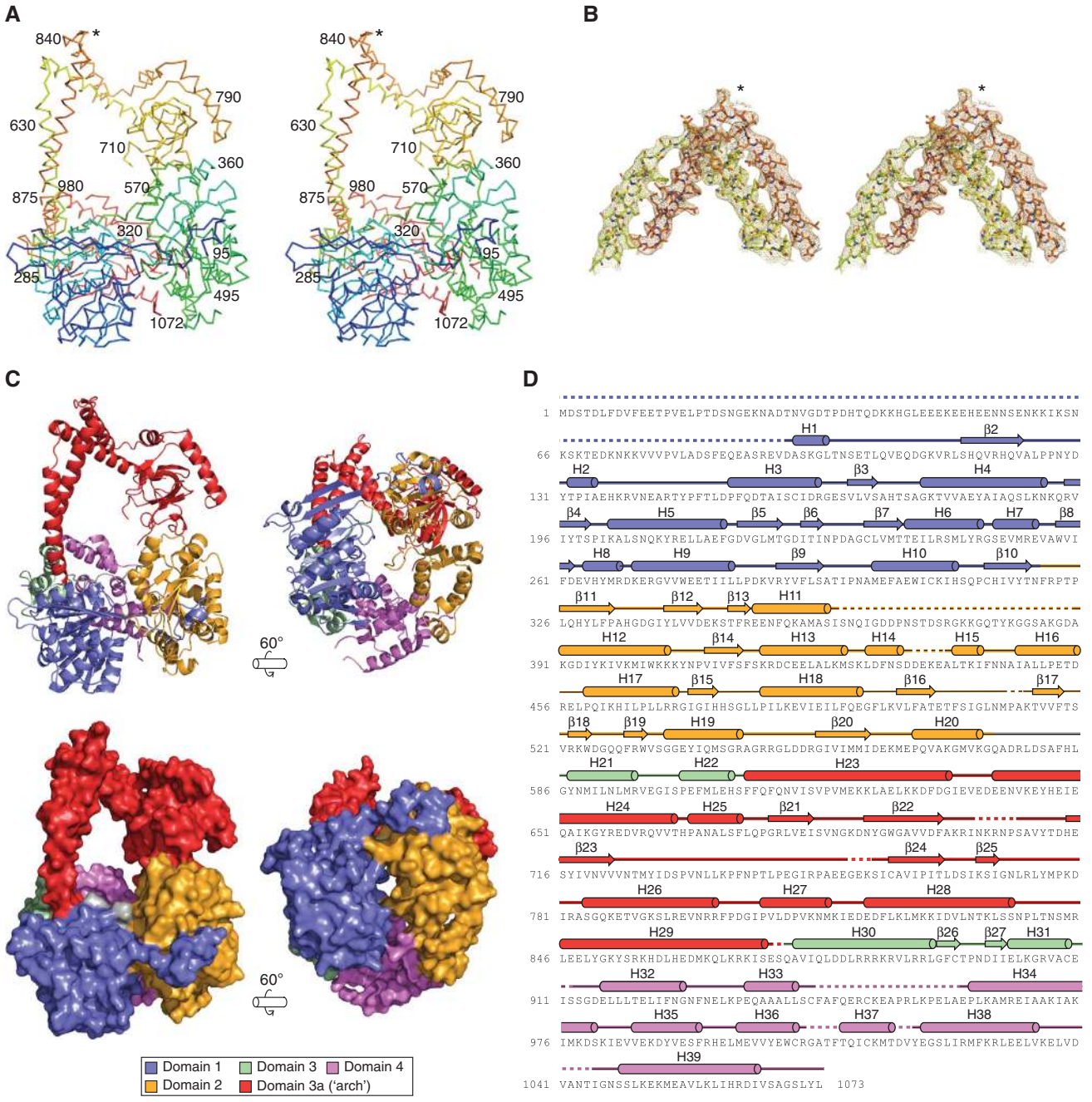

Figure 1 Structural overview of Mtr4. (**A**) Stereo view of the Mtr4 structure. A Cα trace is shown with rainbow colouring from N- (blue) to C-termini (red). Residue numbers are indicated. (**B**) Representative stereo view of the $2f_o f_c$ electron density map, contoured at 1 σ. The (*) corresponds to the same position marked in (**A**). (**C**) Ribbon and surface representations of the Mtr4 structure, coloured by structural domains (see also Supplementary Figure S1). (**D**) Mtr4 primary sequence and secondary structure. Helices and strands are numbered as indicated. Dashed lines denote gaps in the model due to lack of electron density. The first strand in domain 1 (β1) is not indicated, but presumably comes from part of the 85 residues at the N-terminus of Mtr4; the precise sequence could not be determined because no electron density is observed connecting it to the rest of the domain.

Domains 1 and 2 are the core helicase domains. Domains 1 and 2 are canonical RecA-like domains that are each composed of a central β-sheet surrounded by α-helices (Supplementary Figure S1). As observed in other helicase structures (Cordin *et al*, 2006), eight signature sequence motifs are located in a cleft at the interface of domains 1 and 2 (Supplementary Figure S1A). These motifs are typically involved in nucleotide and nucleic acid binding, and hydrolysis of ATP. Domain 1 is an unusually extended β-structure composed of 10 strands which, to our knowledge, is the longest sheet currently reported for superfamily 2 helicase structures (Supplementary Figure S1B). Secondary structure predictions of other Mtr4 and Ski2 species suggest that an extended sheet is a common feature of this protein family. The structural or functional role of this extended sheet is unclear, but the additional strands do contribute to the surface area along the base of the structure. Similar to domain 1, the β-sheet of domain 2 is rather large, containing eight

strands. Extending off the second strand (β2) of domain 1 are a loop and a short helix that span the cleft and pack with domain 2.

Domain 3 adopts a winged helix fold. An ordered 14-residue loop extends across the face of domain 1, connecting domain 2 to domain 3. Domain 3 adopts a winged helix-like fold composed of helices H21, H22, and H30, followed by a β-hairpin (Supplementary Figure S1C). The fold is structurally similar to the winged helix domains of Hel308 (RMSD = 1.97 Å) and Hjm (RMSD = 2.04 Å). An additional helix (H31) is positioned C-terminal to the winged helix fold and forms part of the channel wall that runs through the base of the protein.

Domain 4 resembles the seven-helix bundle of Hel308. The C-terminal domain of Mtr4 is an eight-helix bundle comprising residues 912–1073, and corresponds to the previously annotated DSHCT domain (Staub *et al*, 2004). The Mtr4 structure is the first report of a DSHCT domain. Unexpectedly, the structure closely resembles the seven-helix bundle structures from Brr2 (Pena *et al*, 2009; Zhang *et al*, 2009), Hel308 (Buttner *et al*, 2007; Richards *et al*, 2008), and Hjm (Oyama *et al*, 2009). Indeed, the structural similarity between Mtr4 and Hel308 allowed us to identify connectivity between several helices in domain 4 that would have otherwise remained undetermined due to weak electron density in several loop regions.

Domain 3a—the arch domain—is unique to Mtr4 and Ski2. The most striking and unexpected feature of the Mtr4 structure is a large, 265-residue domain that arches over the rest of the Mtr4 structure and accounts for approximately 25% of the Mtr4 protein (Figures 1 and 2). This 'arch' domain is inserted between the second and third helices of domain 3 (H22 and H30), and replaces the loop that is typically observed in other Ski2-like helicases. The arch domain is composed of two distinct features: (1) the arms and (2) the fist.

The 'arm' (H23 and H29) and 'forearm' (H24 and H28) are each composed of an ascending and descending helix that form antiparallel left-handed coiled coils (Figure 2A). Two short loops at the 'elbow' are located between the arm and the forearm, allowing the structure to make a sharp ∼120° turn. This marked bend is primarily responsible for the arch-like appearance of the domain. Several conserved features are observed at the elbow. A proline (Pro839) facilitates formation of the bend, and similar residues are found in the same region throughout Mtr4 species, although the absolute position varies slightly (Figure 2A). Hydrophobic packing by several conserved residues (Val637, Val643, Tyr646, Leu840, Leu846, and Tyr853) seems to stabilize the bend. In addition, a potential salt bridge forms between the only invariant residues in this region: Glu640 and Lys856. The low resolution of the data precludes an absolute assignment of side-chain conformation, but these residues are clearly positioned in a manner that would allow a salt bridge. Significantly, these residues are only brought into proximity because of the sharp bend at the elbow. A physical interaction between the two residues is currently the best explanation for the absolute conservation observed at these sites. The accumulation of conserved features at the elbow suggests that the bend is a genuine feature of the arch domain that cannot be dismissed as a crystallographic artefact.

The forearm extends into a globular α–β 'fist' that contains a central β-sheet and sits directly above domain 2 (Figures 1 and 2A). The core of the fist is structurally similar to the ribosomal protein L14e of *Sulfolobus solfataricus* (RMSD = 2.72 Å; Figure 2B and Supplementary Figure S2). Both Mtr4 and the L14e protein of *S. solfataricus* adopt a five-strand fold followed by a C-terminal helix. Substantial crystal packing is observed along helices H25 and H26 of the fist. In the absence of additional structural data, it is not clear what influence these packing interactions have on the overall fold of the arch domain, but the most likely impact is on the orientation of the fist.

It is important to note that the internal packing features observed in the arch (e.g. non-crystallographic packing) do not necessarily preclude motion at the elbow or other conformational rearrangements in the arch. The loops at the elbow may actually facilitate conformational flexibility. Indeed, given the proximity of the fist to the predicted path of RNA (see below), we anticipate conformational changes in the arch upon binding of substrates or protein cofactors. Rigid-body motions in the arch may also result from conformational changes in other domains (e.g. upon ATP hydrolysis). Such motions, if they occur, are likely to have significant mechanistic implications. Additional crystallographic and biophysical analyses are therefore needed to characterize the potential dynamics of the arch domain.

The arch domain appears to be a new domain, both in terms of structure and sequence. Comparable arch-like domains have not been observed in other helicase structures. Secondary structure matching (SSM) (Krissinel and Henrick, 2004) and DALI (Holm *et al*, 2008) searches using the arch as a query structure failed to identify similar structures in the protein data bank (although some similarity exists between the fist of the arch and the L14e protein). Similarly, a BLAST (Altschul *et al*, 1990) search of the arch sequence exclusively yields Mtr4 sequences. We note, however, that Ski2 contains a region of similar length (residues 830–1083), which also has limited similarity to other known sequences. Predicted secondary structure analysis indicates this region in Ski2 adopts a fold that resembles the Mtr4 arch (Figure 2A). Although the sequence similarity within the arch domain of Mtr4 and Ski2 is low (∼34%), it does include several residues that are critical for maintaining the overall structure in Mtr4 (Supplementary Figure S3). We conclude that, despite the low sequence similarity, Mtr4 and Ski2 both contain an arch-like domain. The domain is not observed in other Ski2-like RNA or DNA helicases, and is therefore unique to the Mtr4/Ski2 subfamily of Ski2-like helicases (Figure 2C).

Molecular architecture of Ski2-like RNA helicases

The Mtr4 structure defines the fundamental molecular architecture of Ski2-like helicases. The structure clearly demonstrates that domains 3 and 4 are characteristic features of Ski2-like family members (Figure 2C). This observation is particularly significant because sequence analysis failed to identify either region of Mtr4 as similar to other Ski2-like proteins. Domain 3 was not recognized as a winged helix due to the insertion of the 265-residue arch domain between the second and third helices (H22 and H30) of the motif

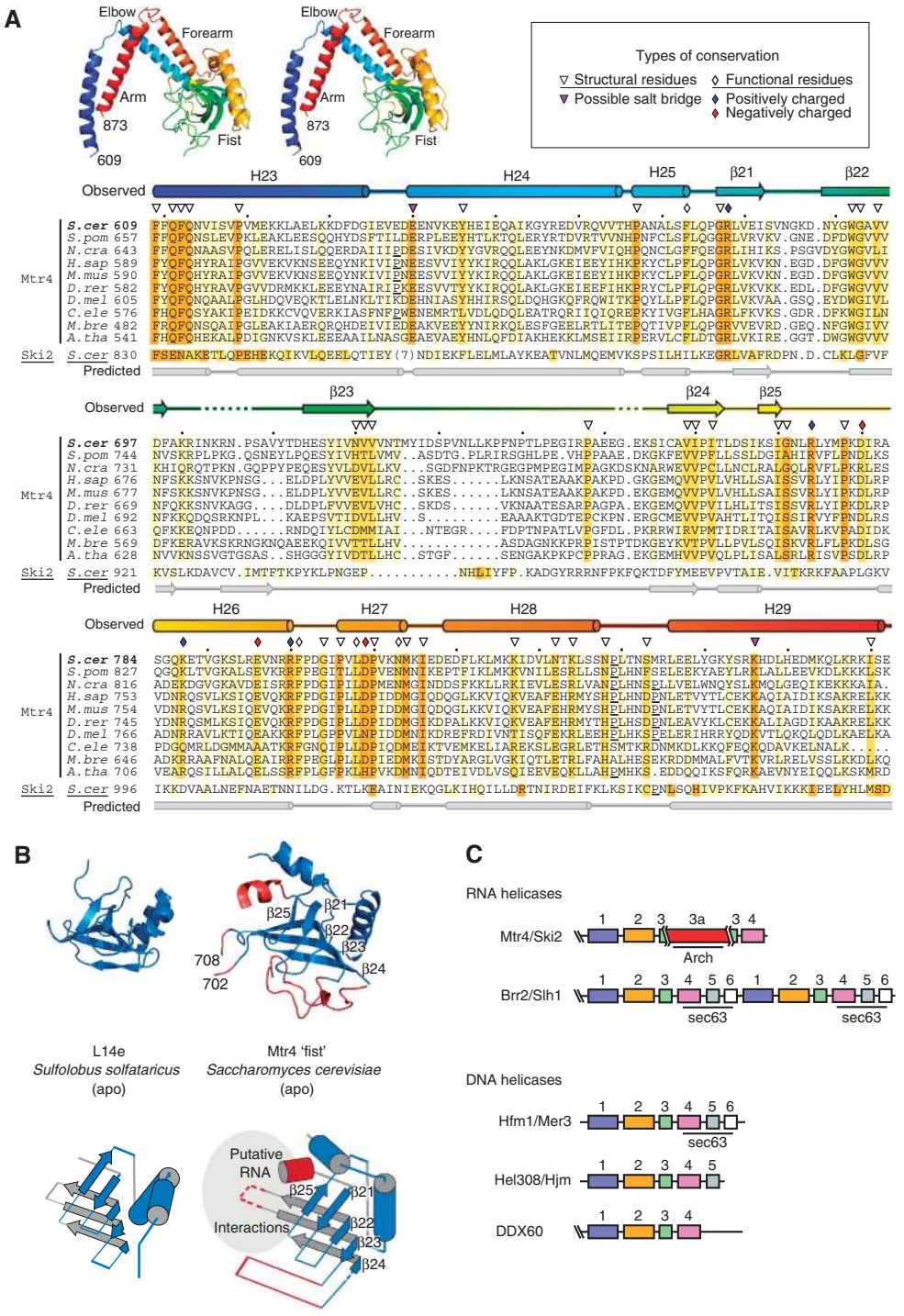

Figure 2 Domain 3a—the 'arch' domain. (**A**) Sequence alignment of the arch domain. Conservation was calculated for Mtr4 sequences using Consurf (Landau *et al*, 2005), based on 108 Mtr4 sequences (orange, strictly conserved residues; yellow, similar residues). Conservation was similarly calculated for 42 Ski2 sequences and is displayed below the Mtr4 sequences. Observed (Mtr4) and predicted (Ski2) secondary structure for the arch is indicated. The colouring of the Mtr4 secondary structure corresponds to the stereo figure of the arch shown above. Two types of conservation are observed: residues that have a role in maintaining the overall fold or structure of the arch (▼), and those that have a potential functional role (◆). Proline residues that are predicted to influence the bend at the elbow are underlined. (S.cer, *Saccharomyces cerevisiae*; S.pom, *Schizosaccharomyces pombe*; N.cra, *Neurospora crassa*; H.sap, *Homo sapiens*; M.mus, *Mus musculus*; D.rer, *Danio rerio*; D.mel, *Drosophila melanogaster*; C.ele, *Caenorhabditis elegans*; M.bre, *Monosiga brevicollis*; A.tha, *Arabidopsis thaliana*) (**B**) Comparison of the Mtr4 fist with the L14e ribosomal protein. Structures (top) and topology diagrams (bottom) are shown. Similar features between the structures are coloured blue. The central fold of the L14e protein from *S. solfataricus* (PDB 2joy) is structurally similar to the fist of Mtr4 and was used to guide model building in the fist. A more extensive comparison is shown in Supplementary Figure S2. (**C**) Domain arrangement of Ski2-like RNA and DNA helicases. Domains 1, 2, 3, and 4 are characteristic of all identified Ski2-like helicases. Sequence and secondary structure analysis indicates that the arch domain is unique to the Mtr4/Ski2 sub-family. The function of DDX60 as an RNA or DNA helicase has not been demonstrated.

(Figure 1C). However, the structural similarity of this domain to other Ski2-like helicases is strong. Similarly, the eight-helix bundle of the Mtr4 domain 4 is clearly related to that of the Brr2 and Hel308 structures (Figure 3C). While our manuscript was under revision, the structure of a DEAH-box RNA helicase, Prp43, was published (He *et al*, 2010), which also contains domains 1–4. This suggests that domains 1–4 may be a common feature in other helicases beyond the Ski2-like RNA and DNA helicase family. Each subfamily of the Ski2-like helicases also retains unique architectural features. The most striking distinction between Mtr4/Ski2 and the other subfamilies is the presence of the large arch domain, which accounts for a quarter of the entire Mtr4 protein.

Model for RNA binding

The structural similarity between Mtr4 (excluding the arch domain) and Hel308 allows us to model the likely RNA-binding site for Mtr4 (Figure 3C). The model was built by aligning the Hel308 DNA-bound structure onto domains 1, 2, 3, and 4 of the Mtr4 structure. On the basis of this model, it is expected that RNA enters Mtr4 near the interface of the fist and domain 2. Single-stranded RNA then makes a 90° turn to pass through a channel formed by domains 1, 2, 3, and 4, and exits out of the base of the structure. The immediate path of a displaced strand is predicted to extend between domain 2 and the fist, possibly in the direction of the central hole of the arch. However, the trajectory of the displaced strand is unclear, in part because the duplex region likely arises from

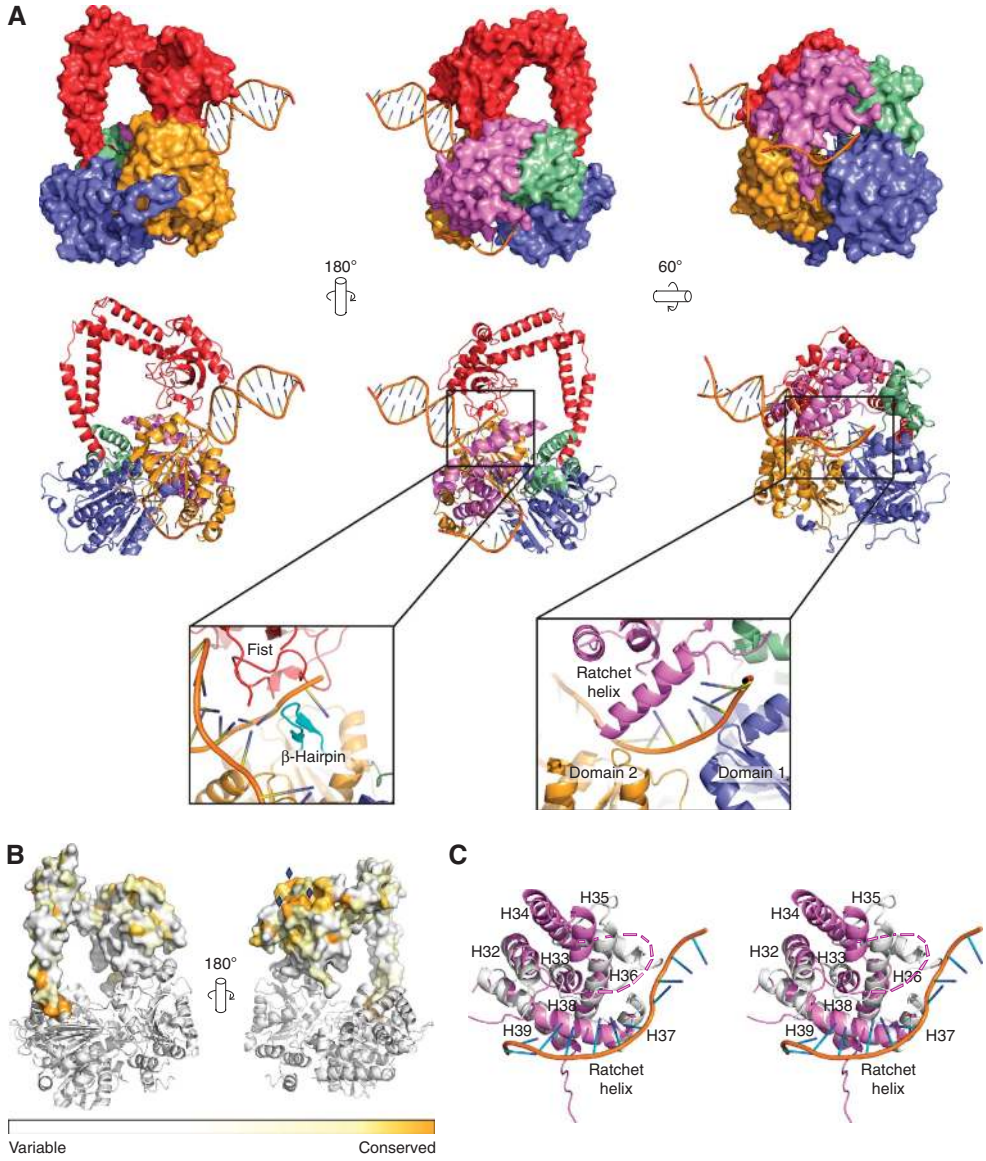

Figure 3 RNA-binding model. (**A**) Nucleic acid from the DNA-bound Hel308 structure (PDB 2p6r) was superimposed on the Mtr4 structure to model the likely RNA-binding site. Surface (top) and cartoon (middle) representations are shown, coloured by domain as in Figure 1. Inset images highlight specific structural features observed in Mtr4 that are associated with strand separation and translocation in Hel308, including the β-hairpin (from domain 2) and ratchet helix (domain 4). In addition, extensive interactions are predicted along the fist of the arch. (**B**) Sequence conservation mapped on the surface of the arch. The colour scheme of conservation is the same as in Figure 2A. The position of three strictly conserved arginine residues that potentially interact with RNA is indicated (◆). (**C**) Stereo view of domain 4. Domain 4 of Mtr4 is an eight-helix bundle and is designated as a DSHCT domain (purple). The domain from Mtr4 is superimposed with the DNA-bound seven-helix bundle of Hel308 (white) to highlight the structural similarity between these two domains. An 18-residue disordered loop in Mtr4 is drawn as a dashed line. Helix H38 of Mtr4 corresponds to the ratchet helix of Hel308.

secondary structures formed by a single strand of RNA, which may preclude large separation of the opposing strands.

The Mtr4 structure reveals all of the general structural components that are required for nucleic acid binding, strand melting and strand translocation (Figure 3A). Specific features include: (1) a β-hairpin (residues 521–532 of domain 2) that is involved in melting duplex nucleic acid, (2) the conserved motifs along the interface of domains 1 and 2 are positioned to interact with the phosphate backbone of the unwound strand, and (3) a ratchet helix. The consistency between Mtr4 and Hel308 (Buttner *et al*, 2007) in these three conserved structural features along the putative RNA-binding path suggests that all Ski2-like helicases likely promote strand displacement and translocation using the same basic mechanisms.

The RNA-binding model also predicts protein–RNA interactions that are unique to Mtr4. The arch domain is poised to interact with RNA primarily in the fist. Conserved, positively charged residues (Arg678, Arg774, and Arg800) are located along one face of the fist and are strong candidates for potential interactions (Figure 3B). Several large loops, which are disordered in the current structure, potentially interact with the incoming RNA. These loops include: a 15-residue loop (699–713) in the fist, a 29-residue region (361–389) in domain 2, and an 18-residue loop (943–961) in domain 4. Each of these disordered regions contains conserved, positively charged residues that could interact with a negatively charged RNA backbone. We note that the temperature-sensitive Cys942Tyr mtr4-1 mutation (Liang *et al*, 1996) resides in domain 4 at the base of the disordered loop, and potentially affects interactions with incoming RNA.

The localization of the fist near the predicted entry point for RNA and adjacent to the β-hairpin of domain 2 suggests that the arch may influence strand separation or resolution of RNA/RNP structures. The coiled-coil scaffold of the arms of the arch may actually provide the structural rigidity needed to resolve some structures. Alternatively, the arch may regulate RNA access to or navigation through the helicase. In this model, the arch could fill a role analogous to domain 5 of other Ski2-like helicases, which acts as a 'molecular brake' to regulate processivity along various substrates (Richards *et al*, 2008).

The RNA-binding model suggests that the large hole formed by the arms of the arch is probably not the primary interface for RNA binding. Consistent with this model, little conservation of charge or sequence is observed along this surface (Figures 2A and 3B). The predicted path of RNA binding is also a significant distance away from the winged helix of domain 3. Although winged helix domains are often associated with nucleic acid binding (Gajiwala and Burley, 2000), it is unclear whether this domain interacts with RNA.

The bottom surface of Mtr4 is a potential interaction site for the exosome

To identify regions of potential surface interactions, conservation among Mtr4 sequences was mapped onto the Mtr4 structure using Consurf (Landau *et al*, 2005; Figure 4A). Strong conservation is observed throughout the base of the structure and is particularly concentrated along the bottom surface near the putative RNA exit site. The conservation is less pronounced in the arch domain and is generally limited to one surface of the fist that includes a cluster of highly conserved arginines (also see Figure 3B). When Ski2 conservation is mapped onto the Mtr4 structure, a similar pattern of strong conservation is observed along the base of the structure with even more limited conservation in the arch (Figure 4B). In contrast to Mtr4 and Ski2, the surface of Hel308 is much less conserved, especially near the nucleic acid exit site (Figure 4C). Thus, the conservation along the base of Mtr4 and Ski2 seems to be a unique feature of this subfamily of Ski2-like helicases. This suggests that the conserved surface has a role that is common to Mtr4 and Ski2, but is distinct from canonical helicase function. One intriguing possibility is that the exosome interacts with Mtr4 (and Ski2) along the base of the structure. This possibility is consistent with the observation that Mtr4 co-purifies with the human exosome (Chen *et al*, 2001). Such a direct interaction would position RNA emerging from the base of Mtr4 adjacent to the site at which RNA enters the exosome. We note that the base of Mtr4 is roughly equivalent in diameter to the entry surface of the exosome.

The current data provide few clues for the location of Trf4–Air2 assembly on Mtr4. As polyadenylation of RNA substrates is likely to precede interactions with Mtr4, the protein–protein binding site is expected to be nearer the putative RNA entry site rather than at the conserved exit site of Mtr4. The arch domain is an attractive target for Trf4–Air2 assembly because of its proximity to the entry site and its distinctive scaffold-like appearance. However, the lack of conservation in this region yields few clues to potential interfaces.

The arch is important for exosome function in vivo

To investigate the function of the arch domain, an Mtr4-archless mutant was constructed by inserting a four-residue linker between H22 and H30 of domain 3 (Supplementary Figure S4). Thus, the archless mutant essentially resembles domains 1, 2, 3, and 4 of Hel308. Not surprisingly, purified Mtr4-archless protein retains both RNA-dependent ATPase activity and ATP-dependent helicase activity (Figure 5). Furthermore, the activity is comparable with full-length Mtr4 (Figure 5B).

To address whether the arch is functionally important, we generated yeast expression plasmids for wild-type and archless Mtr4. A strain containing the Mtr4-archless version is viable, but grows significantly slower than the control wild-type strain (Figure 5D). This slow growth phenotype is not simply due to reduced expression of Mtr4-archless, as western blot analysis indicates that wild-type and archless Mtr4 are expressed at similar levels (Figure 5E). Thus, although the arch domain of Mtr4 is not required for *in vitro* ATPase or helicase activity, it is important for proper Mtr4 function *in vivo*.

To analyse whether the growth defect of the Mtr4-archless strain reflected a defect in exosome-mediated functions, we analysed the processing of 5.8S rRNA and the degradation of the 5′ETS, which are products of a 35S polycistronic precursor (Figure 5F). The 5.8S rRNA is generated from a 7S precursor by the exosome and Mtr4 (de la Cruz *et al*, 1998). The Mtr4-archless strain accumulated a distinct RNA species that is 30 nucleotides (nt) longer than the normal 5.8S rRNA (Figure 5G). We conclude that the arch is required for exosome-mediated rRNA processing. The 5.8S + 30 accumulation phenotype of the Mtr4-archless mutant is similar to that seen in *rrp6*Δ strains (Figure 5G). In contrast, point

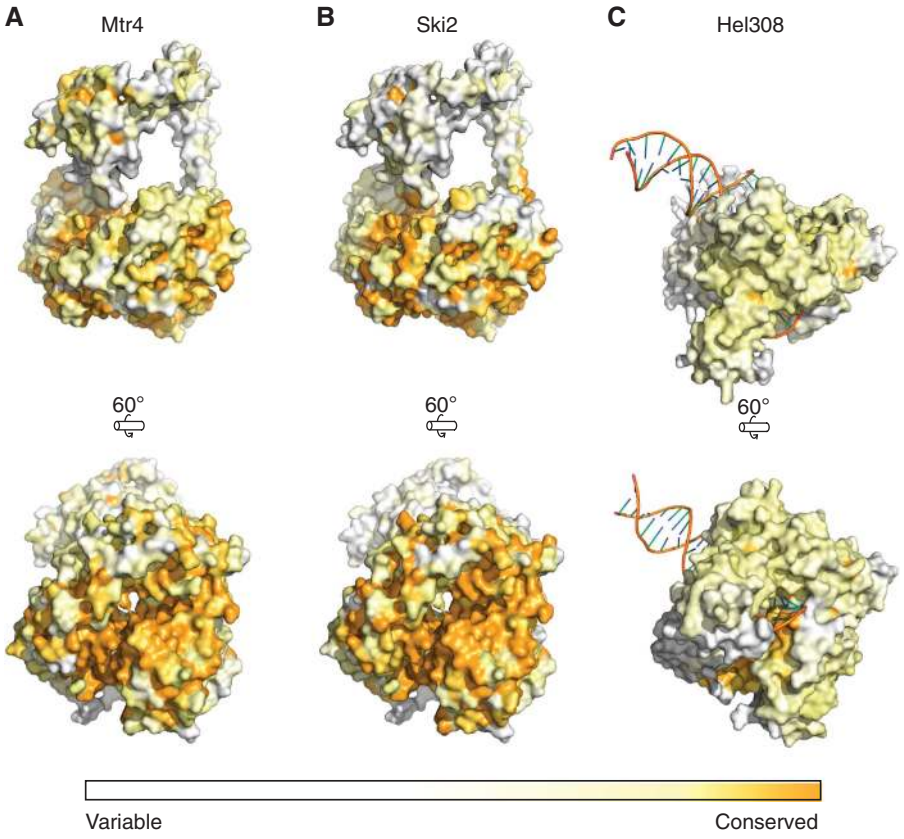

Figure 4 Surface conservation. Evolutionary conservation scores were calculated and mapped onto the protein surface using Consurf (Landau *et al*, 2005). Conservation is depicted by colour gradient ranging from variable (white) to highly conserved (orange). (**A**) Mtr4 conservation mapped onto the Mtr4 structure. (**B**) Ski2 conservation mapped onto the Mtr4 structure. (**C**) Hel308 conservation mapped onto the Hel308 structure (PDB 2p6r). The strong conservation observed along the base of Mtr4 and Ski2 suggests that this is a potential interaction surface. See Supplementary Figure S3 for a more detailed alignment of Mtr4 and Ski2 sequences.

mutations in the core exosome do not lead to the accumulation of this species, but instead cause the accumulation of heterogeneous species that range in length between 7S and 5.8S + 30 (Allmang *et al*, 1999a). On the basis of these observations it has previously been suggested that removal of the last 30 nt of the 5.8S precursor specifically requires Rrp6 and cannot be performed by the core exosome (Briggs *et al*, 1998). The Mtr4-archless phenotype demonstrates that the arch domain is also required for final processing of remaining 30 nt of the 5.8S precursor.

The 5′ ETS is the 5′ most part of the 35S rRNA precursor, which is degraded by the exosome after cleavage from the precursor. Probing a northern blot with a probe for the 5′ETS revealed a pattern in archless that was very similar to that seen in *rrp6*Δ strain. Both strains accumulate about four-fold higher levels of the 5′ ETS compared with wild type. For both the 5.8S rRNA processing and 5′ETS degradation assays, the phenotype of archless mtr4 did not resemble that of point mutants inactivating either of the other two catalytic activities of the exosome (i.e. the *rrp44-D551N* mutant lacking Rrp44 exoribonuclease activity and the *rrp44-D171A* mutant lacking Rrp44 endoribonuclease activity; Figure 5). Thus, our northern blot results suggest a possible role of the arch in activation of Rrp6, although they do not rule out the possibility that the arch also affects the activity of Rrp44 on some substrates.

Consistent with a role for Mtr4 in activating Rrp6, a recent study indicates that the TRAMP complex can stimulate the activity of Rrp6 *in vitro* in a manner that is independent of the helicase or poly(A) polymerase activities of TRAMP (Callahan and Butler, 2009). Notably, TRAMP enhancement of Rrp6 activity was observed even in a TRAMP complex depleted of detectable levels of Mtr4. However, other cofactors associated with the 5.8S rRNA during processing, including the core exosome, were not present in the *in vitro* analysis of TRAMP-mediated Rrp6 activation. The mechanism for Rrp6 activation and the role of the Mtr4 arch domain in that process therefore remains unclear. One possibility is that the arch indirectly enhances Rrp6 activity *in vivo* through other protein–protein interactions (e.g. Trf4 and/or Air2). Alternatively, the arch may assist in making the extra 30 nt of the 5.8S precursor accessible to Rrp6 *in vivo*. This last possibility is especially interesting in light of recent structural and biochemical data indicating that the distance from the top of the core exosome to the active site of the exonuclease Rrp44 is 30 nt (Bonneau *et al*, 2009; i.e. the same as the number left in 5.8S + 30). Previous models for RNA processing have not addressed how the core exosome would stop once it reaches + 30 and how the 5.8S + 30 species is released from the core exosome to be targeted to Rrp6. We speculate that Mtr4 could help stop exosomal decay of processed substrates. The Mtr4-archless data further

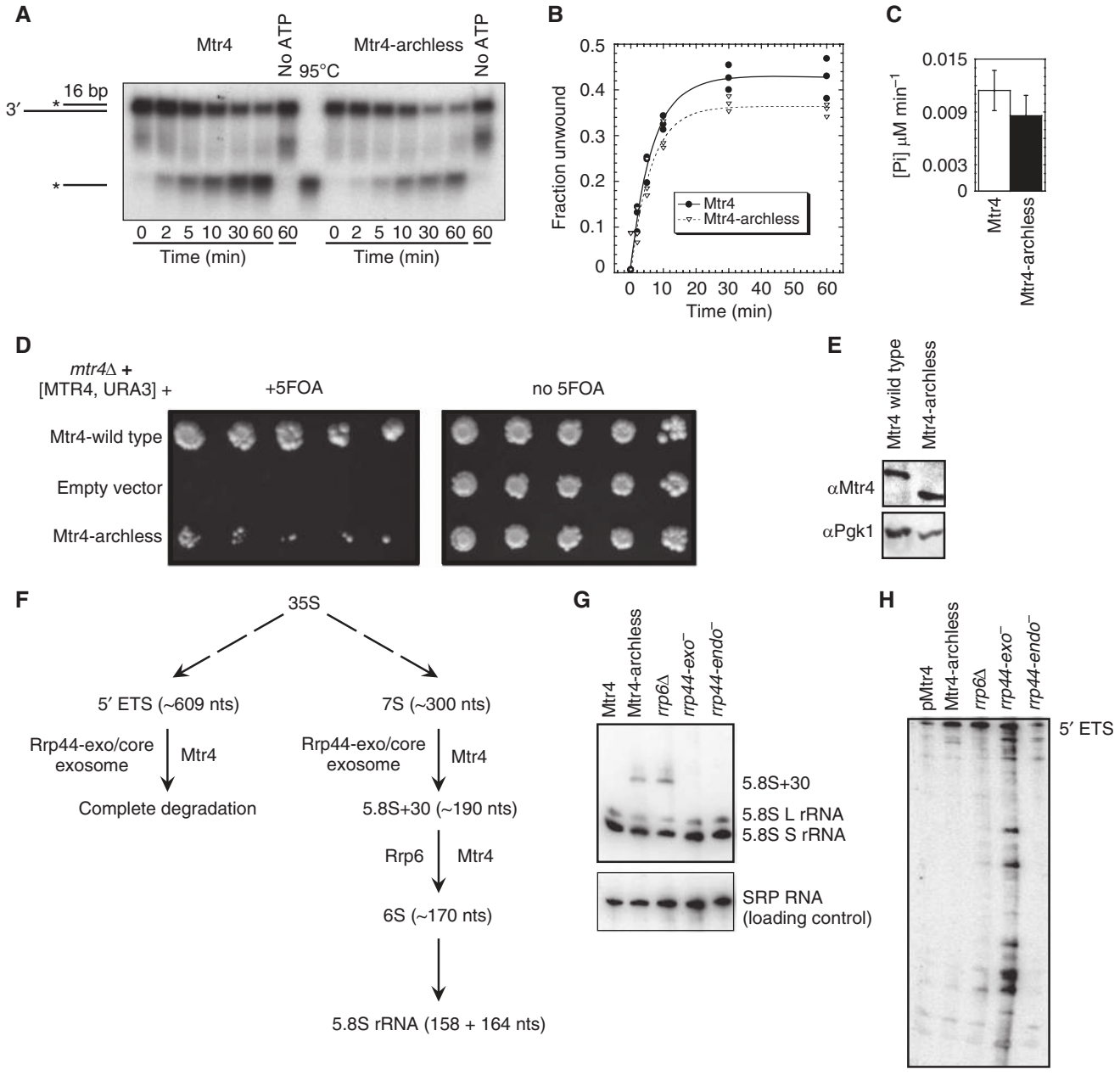

Figure 5 The arch domain of Mtr4 is required for RNA processing and degradation, but not for helicase or ATPase activity. (**A**) Displacement of a radiolabelled single-stranded RNA from a 16-bp duplex with a 3′-single-stranded overhang by full-length Mtr4 and the Mtr4-archless mutant, as observed on a non-denaturing polyacrylamide gel. Illustrations on the left describe the mobility of duplex and single-stranded RNA through the gel; the asterisk represents the ^{32}P label. Aliquots were taken at the time points indicated after the addition of ATP. Lane 8 displays the complete dissociation of the duplex after heating an aliquot to 95°C, and lanes 7 and 15 show a reaction without ATP after 60 min. (**B**) Time courses for the fraction of displaced (unwound) RNA compared with total RNA is shown. Closed circles indicate Mtr4 activity; open triangles indicate Mtr4-archless activity. A solid line representing Mtr4 and a dashed line representing Mtr4-archless were fit to the data points as a first-order reaction (Wang *et al*, 2008), allowing for the determination of reaction amplitudes ($A = 0.427 \pm 0.010$, Mtr4; $A = 0.364 \pm 0.016$, Mtr4-archless) and observed unwinding rate constants ($k'_{unw} = 0.066 \pm 0.005$/min, Mtr4; $k'_{unw} = 0.058 \pm 0.008$/min, Mtr4-archless). (**C**) Enhancement of ATPase activity after introduction of RNA for Mtr4 and Mtr4-archless. The rate was determined using a malachite green assay that monitors the rate of release of inorganic phosphate [Pi] from ATP over time. Background values (activity in the absence of RNA) have been subtracted from the total rate. (**D**) The arch domain is required for optimal growth rate. The indicated plasmids were introduced into a yeast strain that had the *MTR4* gene deleted from the chromosome, and that also contained a plasmid encoding wild-type Mtr4 with a URA3 selectable marker. Growth on 5-FOA plates (left) selects for cells that have lost the URA3 plasmid, and thus shows that Mtr4-archless confers a slow growth phenotype. (**E**) A western blot probed with anti-Mtr4 antibodies (top) and re-probed with anti- Pgk1 antibodies as a loading control shows that archless and wild-type Mtr4 are expressed at similar levels. (**F**) The 5′ ETS and 5.8S rRNA are degraded and processed through actions of the nuclear exosome and Mtr4. The 5′ ETS and 7S rRNA precursor are downstream products of 35S rRNA precursor processing events. 5.8S rRNA is generated through processing of 7S and 5.8S + 30 intermediates and the 5′ ETS is completely degraded. (**G**) The 5.8S + 30 intermediate accumulates in an Mtr4-archless mutant and an *rrp6* knockout strain but not in strains that lack Rrp44 exonuclease or endonuclease activity. Shown is a northern blot probed with a ^{32}P-radiolabelled oligonucleotide specific for the 5.8S rRNA (top) and re-probed for the RNA subunit of the signal recognition particle (bottom). (**H**) The 5′ ETS signal is four-fold greater in the Mtr4-archless, *rrp6* knockout, and *rrp44*-exo$^-$ strains, but only the *rrp44*-exo$^-$ strain shows decay intermediates. Same northern blot, as shown in (**G**), re-probed with a ^{32}P-radiolabelled oligonucleotide specific for the 5′ ETS is shown. Supplementary Figure S4 provides the structural justification for design of the Mtr4-archless mutant used in these studies.

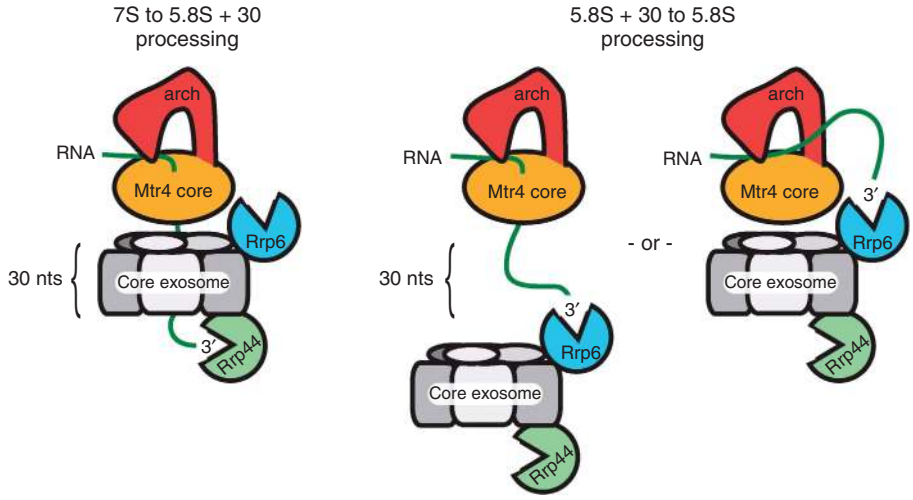

Figure 6 Model for 5.8S rRNA processing. Initial processing of the 7S precursor is performed by Rrp44 in association with the core exosome, leaving a 5.8S + 30 nt RNA. Final processing is performed by Rrp6. Mtr4 is proposed to help make the 5.8S + 30 substrate accessible to Rrp6, possibly by removal of the RNA from the core exosome. Although Rrp6 is depicted in association with the core exosome during processing of the 5.8S + 30 substrate, such an interaction may not be required, as previously suggested (Callahan and Butler, 2008).

raise the possibility that the arch domain (and/or proteins associated with it) may facilitate the retraction of 5.8S + 30 intermediates out of the exosome core to make them accessible for Rrp6 (Figure 6).

Conclusion

The full-length Mtr4 structure presented here represents, to the best of our knowledge, the first complete view of a Ski2-like RNA helicase, and is also the first full-length structure for any RNA helicase of comparable size. The structure reveals a variety of features that suggest that while Mtr4 retains canonical helicase function, it also contains added functionality that extends beyond that of previously characterized helicases. The new arch domain, in particular, seems to have an important role in processing of RNA substrates. The structural and biochemical data presented here suggest numerous avenues towards a more detailed mechanistic understanding of exosome-mediated RNA processing and decay, as well as a broader view of the activities of Ski2-like helicases.

Materials and methods

Construct design

The full-length Mtr4 DNA sequence from *S. cerevisiae* was inserted into a pET 151-D-topo *E. coli* expression vector (Invitrogen) with an N-terminal, TEV-protease cleavable, hexahistidine tag. The removal of the his tag by TEV cleavage leaves two additional amino acids at the N-terminus of the native sequence. The archless Mtr4 mutant was constructed by replacing protein residues 615–878 with a four-residue ELST linker (Supplementary Figure S4).

Protein expression

Mtr4 protein was recombinantly expressed in an *E. coli* BL21(DE3)-codon + RIL cell line (Stratagene). Protein expression was induced using an autoinduction protocol (Studier, 2005). Autoinduction yielded a greater amount of protein per media volume (15–20 mg/l) than traditional IPTG induction methods. Growth at room temperature allowed for greater Mtr4 solubility than that observed for 37°C growth. Cells were collected at 20 000 r.p.m. and stored at −80°C.

Protein purification

Cell lysis was performed by manually breaking up the frozen cell pellet followed by lysozyme treatment and sonication. Ni-affinity, heparin affinity, and gel filtration chromatography techniques were used to purify the protein. All purification buffers contained 5% glycerol, 50 mM HEPES (pH 7.5), and 2 mM beta-mercaptoethanol. Salt concentration was varied in the buffers through the chromatography procedures. The final buffer conditions of purified protein were those described above plus 160 mM NaCl.

Activity assays

The ATPase activity was determined using a colorimetric malachite green assay that detects the release of inorganic phosphate [Pi] after ATP hydrolysis. The method was adapted from the procedure described by Bernstein *et al* (2008) and is described in detail in Supplementary data.

Helicase activity was determined essentially as described by Wang *et al*, (2008). Briefly, the assay measures displacement of a ^{32}P-labelled 16-nt RNA from a complementary RNA strand containing an additional 25-nt 3′ overhang. Displacement is observed on a non-denaturing polyacrylamide gel. The detailed protocol is described in Supplementary data.

Yeast complementation assay

The MTR4-wild-type or MTR4-archless plasmids containing a LEU2 selectable marker were transformed into an MTR4 deletion strain complemented with a MTR4-wild-type copy plasmid containing a URA3 selectable marker. To test whether the arch was essential for growth, the resulting transformants were grown in Synthetic Complete-LEU (SC-LEU) liquid media overnight at 30°C. The liquid cultures were then serially diluted and spotted onto plates containing 5-fluoro-orotic acid (5-FOA; to counter select against the URA3 plasmid with MTR4-wild-type) or control plates (SC-LEU). For more details regarding yeast plasmids and methods see Supplementary data.

Western blot analysis

The MTR4-wild-type and MTR4-archless strains that grew on 5-FOA were grown in YPD, total protein was isolated and analysed using antibodies against Mtr4 at a 1:5000 dilution. Western blots were stripped and re-probed with antibodies against PGK1 (Molecular Probes) to control for equal loading.

Northern blot analysis

The MTR4-wild-type and MTR4-archless strains were grown in YPD, total RNA was isolated, resolved on polyacrylamide gels, and probed with 5′ ^{32}P-labelled oligonucleotides for 5.8S processing defects (5′-TTTCGCTGCGTTCTTCATC-3′), 5′ETS degradation defects (5′-CGAACGACAAGCCTACTCG-3′) and for the RNA subunit of the signal recognition particle (5′-GTCTAGCCGCGAGGAAGG-3′).

Crystallization

Crystallization was performed using standard vapour diffusion methods. Crystals were grown at 4°C in 2.4 M ammonium dihydrogen phosphate, and 0.5 M Tris base (pH 8.5) at a 1:2 protein:well drop ratio, using 10 mg/ml his-tagged Mtr4 protein. Although TEV-cleaved protein is amenable to crystallization in the same conditions, there was no improvement in the diffraction of the crystals. To provide phasing data for structure solution, selenomethionine-substituted (Se-met) Mtr4 was expressed in *E. coli* using a modified autoinduction protocol (Studier, 2005). Se-met purification and crystallization was the same as described for the native protein.

Data collection and structure determination

Crystallographic data were collected to 3.4 Å (native) and 3.6 Å (Se-met) on beamline X29 at the National Synchrotron Light Source (NSLS; Table I). Data were processed using HKL2000 (Otwinowski and Minor, 1997). The Mtr4 crystals belong to space group $P3_121$ ($a = 133.4$ Å, $b = 133.4$ Å, $c = 191.9$ Å; $\alpha = 90°$, $\beta = 90°$, $\gamma = 120°$) and contain one molecule in the asymmetric unit (Matthews coefficient = 3.9; 68% solvent). Phases were determined by the single-wavelength anomalous dispersion method using Se-met substituted Mtr4. The programs SOLVE (Terwilliger and Berendzen, 1999) and RESOLVE (Terwilliger, 2000), as implemented in the PHENIX software package (Adams *et al*, 2002), were used to identify selenium positions (30 out of 32 potential sites were identified) and to calculate maps to 3.6 Å. The maps were subsequently extended to 3.4 Å using the native data.

Modelling and refinement

The resulting electron density maps were of sufficient quality to manually build an initial backbone trace using Coot (Emsley and Cowtan, 2004). The Hel308 DNA helicase structures (PDB 2va8 and 2p6r; Buttner *et al*, 2007; Richards *et al*, 2008) were used to clarify ambiguous connectivity between secondary structures and to improve the model in domains 1, 2, 3 and 4. The Robetta structure prediction server (Kim *et al*, 2004) was used to identify homology between the fist of the arch domain and the archaeal L14e protein from *S. solfataricus* (2joy). The L14e structure provided critical guidance for tracing the map in this region. Secondary structure predictions obtained from PSIPRED (Jones, 1999) provided useful guidance for modelling secondary structure. Methionine residues constitute ~3% of the protein and are quite evenly distributed throughout the sequence. Consequently, methionine positions identified from the Se-met data provided critical place markers to assign sequence to 95% of the final structure.

PHENIX (Adams *et al*, 2002) was used to perform individual b-factor, positional, and TLS refinement. Owing to the low resolution of the data, extensive secondary structure restraints were used to maintain proper geometry during refinement. The final model is a mixed α-helical (50%) and β-sheet (15%) structure containing 918 out of 1073 total residues and four phosphates. Gaps in the current model include 85 residues of the native sequence at the N-terminus and 11 loops of varying length throughout the rest of the sequence (Figure 1). Complete side chains were modelled when supported by the electron density. All other residues were modelled as alanine or serine, as deemed appropriate. The structure was refined to a final R_{work}/R_{free} of 29.3/33.3%. Ramachandran statistics calculated using MolProbity (Davis *et al*, 2007) indicate 74.8% of the residues are in the favoured region with 3.8% outliers. Figures were generated with PyMOL (DeLano, 2002).

Sequence alignment and conservation analysis

Sequence alignments and conservation scores were calculated with ConSurf (Landau *et al*, 2005). Using the arch sequence from *S. cerevisiae* Mtr4 as a query, 108 unique Mtr4 sequences were identified for analysis. A similar search with the putative arch sequence from *S. cerevisiae* Ski2 retrieved 42 unique sequences.

Accession codes

The atomic coordinates and structure factors for Mtr4 have been deposited with Protein Data Bank accession code 3L9O.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Dr Christopher Hill at the University of Utah for access to crystallization robotics. We also thank Dr Patrick Linder for generously providing antibodies against Mtr4 and members of the Johnson and van Hoof labs for insightful comments. The research was supported by the USU Center for Integrated Biosystems (RNJ), the Eccles Foundation (BJH), a USU New Faculty Research Grant (SJJ) and NIH grant GM 069900 (AvH). Financial support for use of the NSLS comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the U.S. Department of Energy, and from the National Center for Research Resources of the National Institutes of Health.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58:** 1948–1954

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D (1999a) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J* **18:** 5399–5410

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P (1999b) The yeast exosome and human PM-Scl are related complexes of 3′→5′ exonucleases. *Genes Dev* **13:** 2148–2158

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410

Anderson JT, Wang X (2009) Nuclear RNA surveillance: no sign of substrates tailing off. *Crit Rev Biochem Mol Biol* **44:** 16–24

Bernstein J, Patterson DN, Wilson GM, Toth EA (2008) Characterization of the essential activities of Saccharomyces cerevisiae Mtr4p, A 3′→5′ helicase partner of the nuclear exosome. *J Biol Chem* **283:** 4930–4942

Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell* **139:** 547–559

Briggs MW, Burkard KT, Butler JS (1998) Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3′ end formation. *J Biol Chem* **273:** 13255–13263

Bruserud O (2007) Introduction: RNA and the treatment of cancer. *Curr Pharm Biotechnol* **8:** 318–319

Buhler M, Haas W, Gygi SP, Moazed D (2007) RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. *Cell* **129:** 707–721

Buttner K, Nehring S, Hopfner KP (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol* **14:** 647–652

Callahan KP, Butler JS (2008) Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. *Nucleic Acids Res* **36:** 6645–6655

Callahan KP, Butler JS (2009) The TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. *J Biol Chem* **285:** 3540–3547

Chen CY, Gherzi R, Ong SE, Chan EL, Raijmakers R, Pruijn GJ, Stoecklin G, Moroni C, Mann M, Karin M (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107:** 451–464

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

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall III WB, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35:** W375–W383

de la Cruz J, Kressler D, Linder P (1999) Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci* **24:** 192–198

de la Cruz J, Kressler D, Tollervey D, Linder P (1998) Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3′ end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J* **17:** 1128–1140

DeLano WL (2002) *The PyMOL Molecular Graphics System*. Palo Alto: DeLano Scientific

Dziembowski A, Lorentzen E, Conti E, Seraphin B (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat Struct Mol Biol* **14:** 15–22

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132

Gajiwala KS, Burley SK (2000) Winged helix proteins. *Curr Opin Struct Biol* **10:** 110–116

He Y, Andersen GR, Nielsen KH (2010) Structural basis for the function of DEAH helicases. *EMBO Rep* **11:** 180–186

Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24:** 2780–2781

Houseley J, Kotovic K, El Hage A, Tollervey D (2007) Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. *EMBO J* **26:** 4996–5006

Houseley J, Rubbi L, Grunstein M, Tollervey D, Vogelauer M (2008) A ncRNA modulates histone modification and mRNA induction in the yeast *GAL* gene cluster. *Mol Cell* **32:** 685–695

Houseley J, Tollervey D (2009) The many pathways of RNA degradation. *Cell* **136:** 763–776

Jankowsky E, Fairman ME (2007) RNA helicases—one fold for many functions. *Curr Opin Struct Biol* **17:** 316–324

Jensen TH, Moore C (2005) Reviving the exosome. *Cell* **121:** 660–662

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292:** 195–202

Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32:** W526–W531

Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60:** 2256–2268

Lacava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121:** 713–724

Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33:** W299–W302

Lebreton A, Seraphin B (2008) Exosome-mediated quality control: substrate recruitment and molecular activity. *Biochim Biophys Acta* **1779:** 558–565

Lebreton A, Tomecki R, Dziembowski A, Seraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* **456:** 993–996

Liang S, Hitomi M, Hu YH, Liu Y, Tartakoff AM (1996) A DEAD-box-family protein is required for nucleocytoplasmic transport of yeast mRNA. *Mol Cell Biol* **16:** 5139–5146

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127:** 1223–1237

Lorentzen E, Conti E (2006) The exosome and the proteasome: nano-compartments for degradation. *Cell* **125:** 651–654

Lykke-Andersen S, Brodersen DE, Jensen TH (2009) Origins and activities of the eukaryotic exosome. *J Cell Sci* **122:** 1487–1494

Milligan L, Torchet C, Allmang C, Shipman T, Tollervey D (2005) A nuclear surveillance pathway for mRNAs with defective polyadenylation. *Mol Cell Biol* **25:** 9996–10004

Nelson PT, Keller JN (2007) RNA in brain disease: no longer just 'the messenger in the middle'. *J Neuropathol Exp Neurol* **66:** 461–468

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, Carter JCW, Sweet RM (eds), Vol. 276: Macromolecular Crystallography, part A, pp 307–326. New York: Academic Press

Oyama T, Oka H, Mayanagi K, Shirai T, Matoba K, Fujikane R, Ishino Y, Morikawa K (2009) Atomic structures and functional implications of the archaeal RecQ-like helicase Hjm. *BMC Struct Biol* **9:** 2

Pena V, Jovin SM, Fabrizio P, Orlowski J, Bujnicki JM, Luhrmann R, Wahl MC (2009) Common design principles in the spliceosomal RNA helicase Brr2 and in the Hel308 DNA helicase. *Mol Cell* **35:** 454–466

Pyle AM (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* **37:** 317–336

Reis CC, Campbell JL (2007) Contribution of Trf4/5 and the nuclear exosome to genome stability through regulation of histone mRNA levels in *Saccharomyces cerevisiae*. *Genetics* **175:** 993–1010

Richards JD, Johnson KA, Liu H, McRobbie AM, McMahon S, Oke M, Carter L, Naismith JH, White MF (2008) Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J Biol Chem* **283:** 5118–5126

San Paolo S, Vanacova S, Schenk L, Scherrer T, Blank D, Keller W, Gerber AP (2009) Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. *PLoS Genet* **5:** e1000555

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat Struct Mol Biol* **16:** 56–62

Schneider C, Leung E, Brown J, Tollervey D (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res* **37:** 1127–1140

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76:** 23–50

Staub E, Fiziev P, Rosenthal A, Hinzmann B (2004) Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. *Bioessays* **26:** 567–581

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41:** 207–234

Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* **56:** 965–972

Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55:** 849–861

van Hoof A, Lennertz P, Parker R (2000) Yeast exosome mutants accumulate 3′-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol Cell Biol* **20:** 441–452

van Hoof A, Parker R (1999) The exosome: a proteasome for RNA? *Cell* **99:** 347–350

Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A new yeast poly(a) polymerase complex involved in RNA quality control. *PLoS Biol* **3:** e189

Wang X, Jia H, Jankowsky E, Anderson JT (2008) Degradation of hypomodified tRNA (iMet) *in vivo* involves RNA-dependent ATPase activity of the DExH helicase Mtr4p. *RNA* **14:** 107–116

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121:** 725–737

Zhang L, Xu T, Maeder C, Bud LO, Shanks J, Nix J, Guthrie C, Pleiss JA, Zhao R (2009) Structural evidence for consecutive Hel308-like modules in the spliceosomal ATPase Brr2. *Nat Struct Mol Biol* **16:** 731–739

Zhang X, Nakashima T, Kakuta Y, Yao M, Tanaka I, Kimura M (2008) Crystal structure of an archaeal Ski2p-like protein from *Pyrococcus horikoshii* OT3. *Protein Sci* **17:** 136–145

