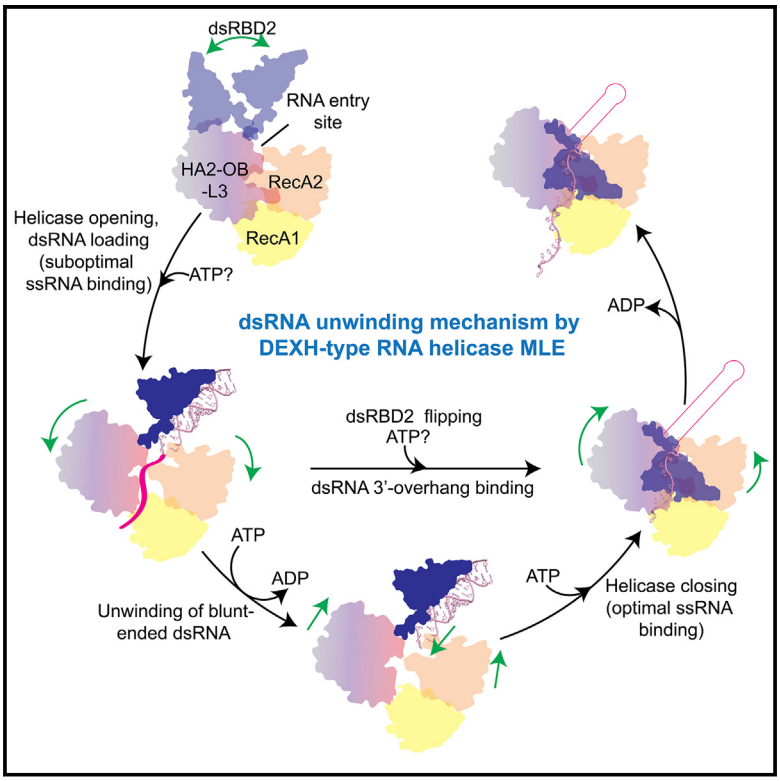
Molecular Cell

Article

Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless

Graphical abstract



Highlights

- Cryo-EM reveals autoregulatory function of dsRBD2 domain in DExH RNA helicase MLE
- Auxiliary domain dsRBD2 aligns dsRNA with MLE's RNA tunnel
- MLE unwinds blunt-end dsRNA and translocates from 3′ to 5′
- These findings reveal essential roles of auxiliary domains in MLE/DHX9 function

Authors

Pravin Kumar Ankush Jagtap, Marisa Müller, Anna E. Kiss, ..., Martin Beck, Peter B. Becker, Janosch Hennig

Correspondence

pravin.jagtap@embl.de (P.K.A.J.), janosch.hennig@embl.de (J.H.)

In brief

Jagtap et al., combining structural and functional studies, show how the auxiliary domain dsRBD2 in DExH RNA helicase MLE regulates the helicase activity by undergoing large-scale conformation changes. These findings contribute to understanding the fundamental mechanics of auxiliary domains in MLE, which serves as a model for its human ortholog and potential therapeutic target, DHX9.

Jagtap et al., 2023, Molecular Cell *83*, 1–16
December 7, 2023
https://doi.org/10.1016/j.molcel.2023.10.026

CellPress

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Molecular Cell

CellPress
OPEN ACCESS

Article

Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless

Pravin Kumar Ankush Jagtap,[1,2,*] Marisa Müller,[3] Anna E. Kiss,[3] Andreas W. Thomae,[3,4] Karine Lapouge,[5] Martin Beck,[1,6] Peter B. Becker,[3] and Janosch Hennig[1,2,7,*]

[1]Structural and Computational Biology Unit, EMBL Heidelberg, Meyerhofstraße 1, 69117 Heidelberg, Germany
[2]Chair of Biochemistry IV, Biophysical Chemistry, University of Bayreuth, Bayreuth, Germany
[3]Molecular Biology Division, Biomedical Center, LMU Munich, 82152 Planegg-Martinsried, Germany
[4]Core Facility Bioimaging at the Biomedical Center, LMU Munich, 82152 Planegg-Martinsried, Germany
[5]Protein Expression and Purification Core Facility, EMBL Heidelberg, 69117 Heidelberg, Germany
[6]Department of Molecular Sociology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany
[7]Lead contact
*Correspondence: pravin.jagtap@embl.de (P.K.A.J.), janosch.hennig@embl.de (J.H.)
https://doi.org/10.1016/j.molcel.2023.10.026

SUMMARY

RNA unwinding by DExH-type helicases underlies most RNA metabolism and function. It remains unresolved if and how the basic unwinding reaction of helicases is regulated by auxiliary domains. We explored the interplay between the RecA and auxiliary domains of the RNA helicase maleless (MLE) from *Drosophila* using structural and functional studies. We discovered that MLE exists in a dsRNA-bound open conformation and that the auxiliary dsRBD2 domain aligns the substrate RNA with the accessible helicase tunnel. In an ATP-dependent manner, dsRBD2 associates with the helicase module, leading to tunnel closure around ssRNA. Furthermore, our structures provide a rationale for blunt-ended dsRNA unwinding and 3′-5′ translocation by MLE. Structure-based MLE mutations confirm the functional relevance of our model for RNA unwinding. Our findings contribute to our understanding of the fundamental mechanics of auxiliary domains in DExH helicase MLE, which serves as a model for its human ortholog and potential therapeutic target, DHX9/RHA.

INTRODUCTION

RNA helicases are essential enzymes for remodeling RNA during a wide range of cellular processes.[1–3] These RNA-dependent ATPases use their helicase activity to interconvert RNA secondary structures[4,5] and are categorized into non-processive (DEAD-box) and processive (DExH-type) helicases.[6,7] In DEAD-box helicases, binding of substrates, ATP, and RNA triggers structural transitions between their RecA1 and RecA2 domains, transitioning between closed "on" and open "off" states upon ATP hydrolysis and concomitant RNA release. Eukaryotic DExH/DEAH helicases process one nucleotide per ATP hydrolysis, and their RecA domains (helicase module) are often flanked by N- and/or C-terminal auxiliary domains,[8–15] which engage in RNA interactions that contribute to substrate selectivity in the biological context.[16,17] Conditional conformational changes of auxiliary domains may install positive or negative autoregulation of the helicase activity.[18–25] Although previous studies have enhanced our understanding of helicase action, detailed molecular insights into how eukaryotic DExH/DEAH helicases recognize substrates and the functional interplay between auxiliary domains and the helicase module during double-stranded RNA (dsRNA) unwinding remain unexplored.

Maleless (MLE) is a well-studied DExH-type RNA helicase,[26] particularly in the context of *Drosophila* dosage compensation, where the male-specific-lethal (MSL) complex enhances X chromosome gene transcription by approximately 2-fold. The MSL complex comprises MLE, the rate-limiting subunit MSL2, the scaffold protein MSL1, epigenetic reader MSL3, histone acetyltransferase "males-absent on the first" (MOF), and one of two long non-coding RNAs roX1 or roX2 ("RNA on the X").[27,28] Lack of roX RNA is lethal as dosage compensation cannot be maintained.[29] During dosage compensation, MLE binds conserved roX-box motifs and remodels roX RNAs into an alternative conformation (alternative stem loop [ASL]), which is then incorporated into the MSL complex (Figure 1A).[30–34] MLE exhibits high-affinity binding to roX RNAs but loosely associates in an RNA-dependent manner with the MSL complex,[33] suggesting a transient interaction and potentially leaving after handing over roX RNA.[34] This RNA remodeling by MLE is essential for MSL complex localization to the X territory of the nucleus.[31,33] Besides dosage compensation, MLE is also involved in RNA editing and siRNA processing.[35,36]

MLE's domain architecture includes two N-terminal dsRNA-binding domains (dsRBDs), followed by the helicase domains (RecA1, RecA2), HA2 and OB-like domains, and a C-terminal glycine-rich region (Figure 1B).[8] This modular architecture segregates ss- and dsRNA-binding properties within MLE, with

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

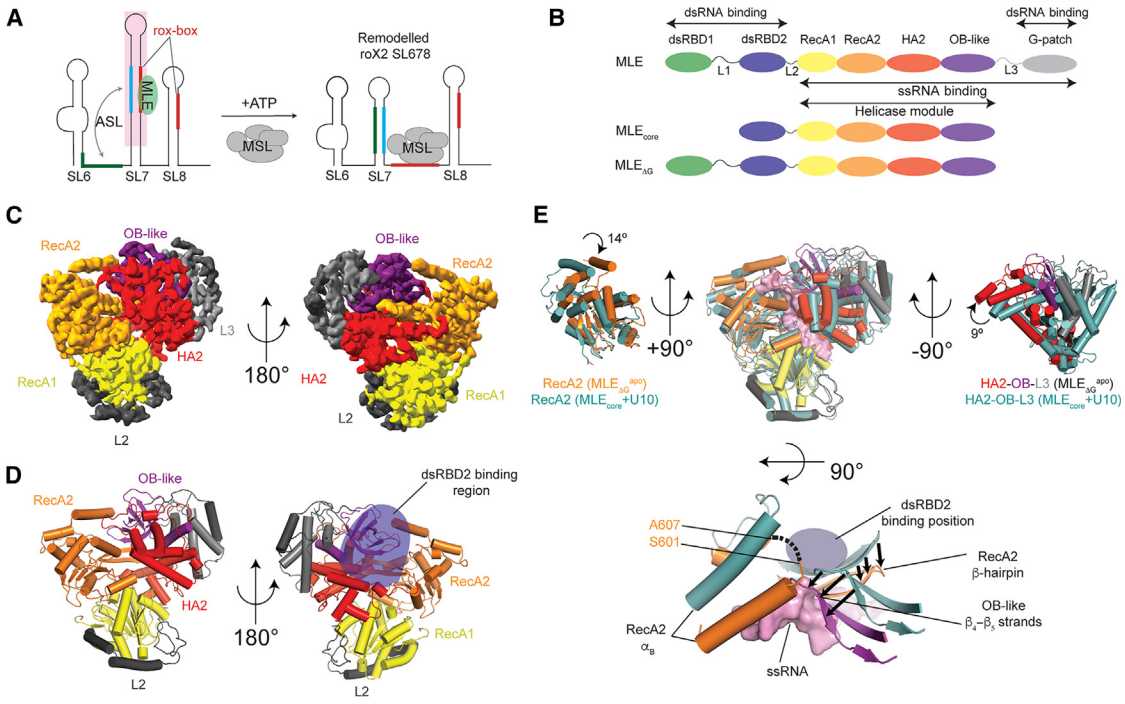


Figure 1. Cryo-EM structure of $MLE_{\Delta G}$ in the apo-state

(A) Schematic of the secondary structure of the biologically relevant three SLs in roX2 RNA located at the 3′ end (SL6, SL7, and SL8). Upon remodeling of roX2 RNA (alternative stem loop [ASL]) by MLE, the MSL complex subunit MSL2 binds to the roX-box region. The region in roX2, of which dsRNA fragments were designed for cryo-EM studies with MLE is shaded in pink.

(B) Domain organization of MLE from *D. melanogaster* and the constructs used for structural studies. The ss- and dsRNA-binding regions in MLE are indicated.

(C and D) (C) Cryo-EM density and (D) structure of $MLE_{\Delta G}^{apo}$. The expected binding position of dsRBD2 on the MLE helicase module, as expected from the MLE_{core} + U10 crystal structure (PDB: 5AOR), is marked with a blue oval.

(E) Structural superposition of $MLE_{\Delta G}^{apo}$ cryo-EM (with respective domain colors) and MLE_{core} + U10 (light teal) crystal structures. Helices are shown as cylinders for simplicity. The structures were superposed with respect to the RecA1 domain. The rotations of the RecA2 domain and the HA2-OB-L3 module are indicated. Black arrows in the top view indicate the inward movement of the RecA2 β-hairpin and the OB-like β_4-β_5 strands, which occludes the ssRNA-binding tunnel in the $MLE_{\Delta G}^{apo}$ structure. The missing loop between the RecA2 α_B and α_C is shown as a dotted line in the $MLE_{\Delta G}^{apo}$ structure. See also Figures S1–S6.

dsRNA binding mediated by the two N-terminal dsRBDs, and single-stranded RNA (ssRNA) binding by the helicase module comprising both RecA domains, the HA2 domain, and the OB-like fold. The unstructured G-patch region binds to both ds- and ssRNA.[37] The region spanning dsRBD2 to the L3 linker forms the compact helicase core (MLE_{core}), comprising the RecA, HA2, OB-like fold domains, and the structured L3 linker, which creates an RNA-binding tunnel for ssRNA during helicase translocation and thus forms MLE's helicase module.[8] Among the two dsRBDs, dsRBD2 is indispensable for helicase activity and localization to the X chromosome.[37]

Prior studies have unveiled the structures of the isolated tandem dsRBDs, both free and upon dsRNA binding and the MLE_{core} with a single-stranded uridyl-homopolymer RNA and ATP transition state analog ADP:AlF_4.[8,38,39] However, these studies did not fully elucidate the RNA recognition and unwinding mechanism, the structural rearrangements during MLE's productive RNA interactions, such as binding of roX stem-loop (SL) structures, unwound ssRNA funneling through the helicase, and the role of auxiliary domains in the entire RNA remodeling process. To address these questions, we conducted structural, biochemical, biophysical, and functional studies supported by a series of cryo-electron microscopy (cryo-EM) structures of MLE at different helicase cycle states.

RESULTS

MLE ssRNA-binding tunnel is inaccessible in apo and nucleotide-bound states

A previously published crystal structure of MLE_{core} with ssRNA of 10 uridines (U10) and ATP transition state analog ADP:AlF_4 (MLE_{core}^{ATP} + U10) was interpreted as a closed state of the enzyme after unwinding the dsRNA substrate but still bound to ssRNA.[8] To monitor conformational rearrangements during unwinding, we determined seven cryo-EM structures. First, we determined the structure of $MLE_{\Delta G}$ (lacking only the C-terminal unstructured G-patch region; Figure 1B) in the apo-state ($MLE_{\Delta G}^{apo}$). Its architecture resembles the MLE_{core}^{ATP} + U10 crystal structure[8] (Figures 1C, 1D, and S1–S4; Table 1), adopting a trefoil arrangement formed by RecA1, RecA2 domains, and the HA2-OB-L3 module. However, dsRBD2, which stably interacts with the helicase module in the MLE_{core}^{ATP} + U10 state, and dsRBD1 are not visible in the EM density map, suggesting the flexibility of dsRBD1,2 orientations.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

CellPress
OPEN ACCESS

Table 1. Cryo-EM data collection, refinement, and validation statistics

	$MLE_{\Delta G}^{apo}$	$MLE_{\Delta G}^{ATP}$	$MLE_{\Delta G}^{ATP}$ + U10	$MLE_{\Delta G}^{ATP}$ + UUC	$MLE_{\Delta G}^{ATP}$ + SL7, dts1	$MLE_{\Delta G}^{ATP}$ + SL7, dts2	$MLE_{\Delta G}^{ADP}$ + UUC	$MLE_{\Delta G}^{ADP}$ + SL7
EMDB ID	15935	15933	15931	15932	15934		17703	17711
PDB ID	8B9L	8B9J	8B9G	8B9I	8B9K		8PJB	8PJJ
Data collection and processing								
Magnification	105,000	105,000	105,000	105,000	105,000	105,000	105,000	130,000
Voltage (kV)	300	300	300	300	300	300	300	300
Electron exposure (e^{-}/Å^2)	50.5	51.7	61.2	48.5	47.9	51.7	49.4	52.6
Defocus range (μm)	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0	−0.8 to −2.0
Pixel size (Å)	0.822	0.822	0.822	0.822	0.822	0.822	0.822	0.932
Initial particle images (no.)	1,470,793	3,199,267	4,503,260	1,807,333	1,763,219	184,188	2,704,419	1,952,320
Final particle images (no.)	270,599	234,365	559,766	476,446	325,471		79,756	61,690
Map resolution (Å)	3.9	3.5	2.9	3.0	4.0		3.6	4.2
FSC threshold	0.143	0.143	0.143	0.143	0.143		0.143	0.143
Map resolution range (Å)	2.9–6.8	2.9–6.8	2.5–5.9	1.8–3.0	3.5–8.6		2.4–8.1	2.8–9.0
Refinement								
Initial model used (PDB code)	5AOR	$MLE_{\Delta C}$	5AOR	5AOR	5AOR, 6I3R		5AOR	5AOR, 6I3R
Model resolution (Å)	3.9	3.5	3.0	2.9	4.0		3.6	4.2
Model resolution range (Å)	420.8–3.9	420.8–3.5	420.8–3.0	420.8–2.9	526.0–4.0		420.8–3.6	526.0–4.2
Map sharpening *B* factor (Å^2)	−154	−137	−128	−134	−203		−133	−143
Model composition								
Non-hydrogen atoms	6,827	6,879	7,801	8,036	8,522		7,903	8,353
Protein residues	861	864	953	984	966		967	947
Nucleotide residues	0	0	10	11	40		11	40
Ligands	0	3	3	3	2		2	1
B factors (Å^2)								
Protein	45.3	45.3	40.5	41.1	92.1		40.6	365.1
Ligand	–	43.2	21.6	30.1	20		–	281.8
Nucleotide	–	–	52.1	45.57	53.0		67.9	234.4
RMSDs								
Bond lengths (Å)	0.003	0.004	0.004	0.004	0.003		0.003	0.003
Bond angles (°)	0.69	0.79	0.628	0.68	0.6		0.66	0.74
Validation								
MolProbity score	2.2	2.1	1.7	2.1	2.8		1.7	2.32
Clashscore	14.3	13.3	9.0	11.7	16.8		8.9	19.03
Poor rotamers (%)	0.3	0.4	1.3	2.3	4.7		0.4	0.2
Ramachandran plot								
Favored (%)	90.74	91.94	97.46	96.32	89.09		96.25	90.32
Allowed (%)	8.44	7.59	2.43	3.58	10.4		3.64	9.36
Disallowed (%)	0.82	0.47	0.11	0.10	0.52		0.1	0.32

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

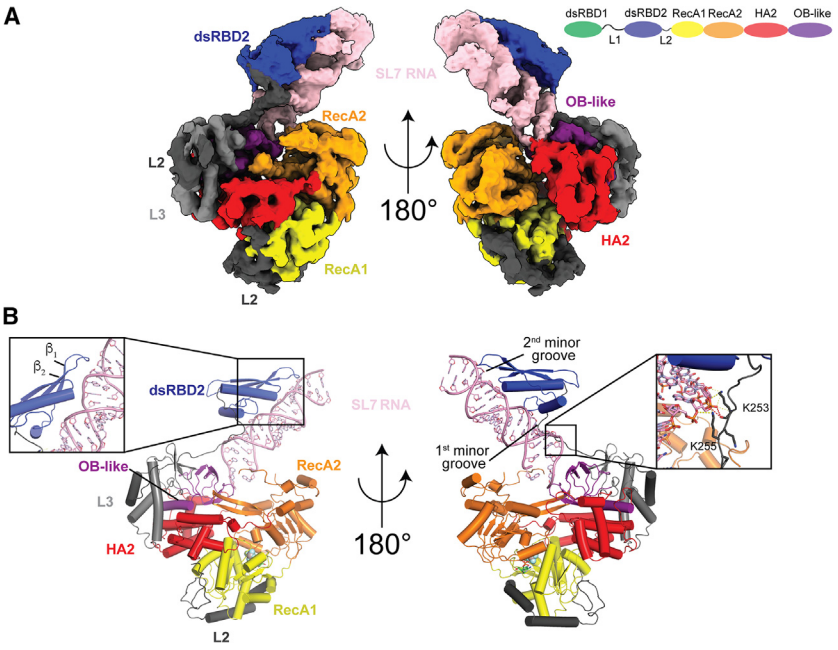


Figure 2. Structure of MLE$_{\Delta G}$ in complex with dsRNA

(A) Cryo-EM density map of MLE$_{\Delta G}$ + SL7 with respective domain colors. The dsRNA is shown in pink.

(B) Structure of MLE$_{\Delta G}$ + SL7. Zoomed-in panels show key interactions between the RNA and different domains of MLE. The two minor grooves in the dsRNA are indicated. Yellow and red dashed lines show potential polar and electrostatic contacts, respectively, between RNA and dsRBD2-helicase module linker. See also Figures S1–S7.

Comparison of the MLE$_{\Delta G}^{apo}$ and MLE$_{core}^{ATP}$ + U10 structures revealed that ssRNA cannot fit the entry tunnel, clashing with several structural elements of MLE$_{\Delta G}^{apo}$ including the RecA2-α_B and α_B-α_C loop (Figure 1E; Video S1; for a detailed description, see Figure S4C). In the MLE$_{\Delta G}^{apo}$ structure, the helicase module interface to dsRBD2 seen in the MLE$_{core}^{ATP}$ + U10 structure is solvent exposed, and RecA2 residues 602–606 interacting with dsRBD2 in the MLE$_{core}^{ATP}$ + U10 structure are not visible and presumably flexible. The tip of the RecA2 β_4-α_4 β-hairpin is pushed into the RNA path of the tunnel (Figure 1E). Thus, the conformation of the RNA-binding tunnel in the MLE$_{\Delta G}^{apo}$ state does not support ssRNA binding. The cryo-EM structure of MLE$_{\Delta G}$ in complex with ADP:AlF$_4$ (MLE$_{\Delta G}^{ATP}$) showed only minor conformational changes compared with the MLE$_{\Delta G}^{apo}$ structure (Figures S1–S3, S4A, S4B, S5, and S6). The distances between the center-of-mass of the RecA1, RecA2, and HA2-OB-L3 module changed within ~1 Å, and binding of ATP does not lead to an opening of the helicase module to allow ssRNA binding (Figure S6). This contrasts with other DExH RNA helicases, including spliceosome-associated DEAH helicases Prp43 and Prp22, which show conserved residues involved in RNA and ATP binding and yet reside in an open conformation competent for ssRNA binding in the absence of adenosine nucleotides.[9,10,40–44] In Prp22, the presence of adenosine nucleotides stabilizes the closed conformation of the helicase module, regardless of the presence of RNA.[40] Thus, the MLE$_{\Delta G}^{apo}$ and MLE$_{\Delta G}^{ATP}$ structures do not explain the structural transitions necessary for dsRNA unwinding to arrive at the MLE$_{core}^{ATP}$ + U10 state.

Structure of MLE in complex with dsRNA reveals the opening of the helicase module upon dsRNA binding

The discussed structures exclude MLE conformations when bound to dsRNA. MLE initially recognizes roX RNA stem-loop 7 (SL7) dsRNA (Figure 1A),[34] potentially involving MLE dsRBD1,2 domains, which were studied previously only in isolation.[37–39] Therefore, dsRNA was predicted to bind to the MLE helicase module in a perpendicular position relative to the ssRNA, involving structural changes of RecA2-α_B and α_B-α_C loop to expose the RNA entry tunnel.[8,39]

To investigate dsRNA binding to MLE, we attempted to obtain a cryo-EM structure of MLE$_{\Delta G}$ in complex with roX2 SL678 (Figure 1A). However, sample heterogeneity and complex dissociation were observed in electro mobility shift assays (EMSAs) and on cryo-EM grids, respectively (Figure S7A). To circumvent this, we designed an SL derived from SL7 of roX2 with a 3′ single-stranded extension consisting of the UxUUU ssRNA motif shown to bind MLE in the RNA tunnel[8] (called "SL7" hereafter, Figure S7B). MLE$_{\Delta G}$ bound to this RNA with an affinity of 16 ± 1.4 nM in a 1:1 stoichiometry (Figures S7C and S7D).

We obtained a cryo-EM structure of MLE$_{\Delta G}$ in complex with SL7 RNA and ADP:AlF$_4$ (MLE$_{\Delta G}^{ATP}$ + SL7, Figures 2A, 2B, S1–S3, and S4B; Table 1). The dsRBD2 domain is now visible in the EM density, bound to dsRNA and aligning it with the ssRNA-binding tunnel instead of the previously predicted perpendicular conformation (Figures 2A, 2B, and S8A–S8C).[39] The dsRBD2 mediates canonical interactions with the dsRNA. The loop connecting dsRBD2 β_1-β_2 strands and the α_1 helix bind to the 2nd and 1st minor grooves, respectively, and the N-terminal region of α_2 helix binds to the next major groove (counting from the proximal base to the helicase module).[38,39,45] Key residues in the L2 linker cause structural constraints and define the range for dsRNA contacts to be within the first register of the minor-major-minor groove (Figure 2B).

Notably, dsRNA recognition by MLE involves dsRBD2 and the N-terminal 26 residues of the L2 linker, providing a continuous patch of a positive electrostatic surface potential that interacts with the negatively charged dsRNA backbone (Figures S8D and S8E). This positively charged surface continues into the RNA entry tunnel (Figure S8E). Highly conserved L2 linker lysines (K253-K256) are flexible in the absence of dsRNA[8] but guide the 3′ end of the dsRNA into the entry tunnel of the helicase module (Figures 2B, S8D, and S8F), emphasizing their importance for helicase function.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

CellPress
OPEN ACCESS

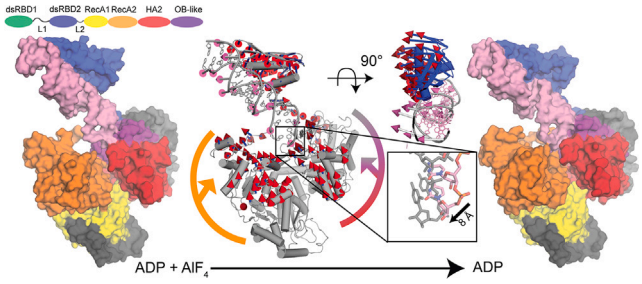


Figure 3. Conformational rearrangements in dsRNA-bound MLE upon ATP hydrolysis

The $MLE_{\Delta G}^{ATP}$ + SL7 (left) and $MLE_{\Delta G}^{ADP}$ + SL7 (right) structures are shown as surface representations. Movement of dsRNA, dsRBD2, RecA2 domain, and HA2-OB-L3 module upon ATP to ADP transition is shown with C_α vectors from blue to red on the cartoon representation of $MLE_{\Delta G}^{ATP}$ + SL7 structure (in gray, with RecA2 and HA2-OB-L3 module shown with orange and red-purple-gray arcs). The top view shows the movement of dsRNA-dsRBD2 toward the back of the protein. The RecA2 domain and HA2-OB-L3 module moves up, toward the dsRNA, and the 3′ strand at the bound end of the dsRNA moves down toward the RNA-binding tunnel in the helicase module (shown in inset), rationalizing 3′ to 5′ directional movement of MLE on RNA. See also Figures S1–S6 and S9.

Conformational rearrangement of the helicase module upon ATP hydrolysis

To investigate the effect of ATP hydrolysis on MLE's helicase module conformation, we determined the cryo-EM structure of $MLE_{\Delta G}$ bound to ADP and SL7 ($MLE_{\Delta G}^{ADP}$ + SL7, Figures 3, S1–S3, and S4B). This structure closely aligns with that in complex with ADP + AlF_4, with minor domain rotations (overall root-mean-square deviation [RMSD]: 1.2 Å). The RecA2 and HA2-OB-L3 modules rotate by 9° and 6°, whereas dsRNA and the dsRBD2 rotate by 6° and 32° with respect to $MLE_{\Delta G}^{ATP}$ + SL7 (Figure S9). Notably, the RecA2 domain and HA2-OB-L3 module move upward toward the dsRNA's 5′ end, relative to the RecA1 domain, and the dsRNA's 3′ end moves into the helicase module by ~8 Å (Figure 3; Video S2). MLE has been shown to unwind blunt-ended RNA duplexes, albeit less efficiently, relying on dsRBD2.[8] These structural changes in $MLE_{\Delta G}^{ADP}$ + SL7 complex compared with $MLE_{\Delta G}^{ATP}$ + SL7 may explain how the 3′ end is ratcheted into the helicase module upon ATP hydrolysis, facilitating directional 3′-5′ translocation on the dsRNA.

Conformational rearrangement of dsRBD2 is required for ssRNA binding to the helicase module

Our $MLE_{\Delta G}$ structures in complex with the SL7 RNA reveal large conformational changes upon dsRNA binding. To accommodate dsRNA at the ssRNA entry tunnel, the RecA2 domain and HA2-OB-L3 module rotate by 14° and 19°, respectively, away from the tunnel compared with the $MLE_{\Delta G}$ structure (Figures 4A and S10A; Video S3). These rotations cause the retraction of the HA2 α_6-α_7 loop, α_{10} helix, and the RecA2 α_2/α_3 helices, allowing the 3′ end of the dsRNA to approach the entry tunnel. Furthermore, the RecA2 α_B helix is not visible in the EM density, suggesting it is flexible, thereby allowing it to accommodate the 5′ end of the dsRNA (Figures 4B and S10B).

However, we only observed density for two nucleotides from SL7's 5′ end of the ssRNA extension leading up to the entry tunnel in $MLE_{\Delta G}^{ATP}$ + SL7 and $MLE_{\Delta G}^{ADP}$ + SL7 structures, although the designed ssRNA extension was intended to fit into the tunnel entirely (Figures 2 and 3). Nevertheless, EMSAs confirmed that $MLE_{\Delta G}$ can bind both parts of SL7 RNA ($SL7^{dsRNA}$ and UUC RNA; Figure S7B), albeit with lower affinity compared with complete SL7 RNA (Figure S10C). To understand the interactions between $MLE_{\Delta G}$ and ssRNA, we determined cryo-EM structures of $MLE_{\Delta G}$ in complex with ADP:AlF_4 and either U10-mer ssRNA ($MLE_{\Delta G}^{ATP}$ + U10) or UUC ssRNA ($MLE_{\Delta G}^{ATP}$ + UUC) (Figures S1–S3, S4A, S4B, and S11A–S11D). Both structures superpose with the previous MLE_{core} + U10 crystal structure[8] with RMSDs of 0.7 and 1.4 Å, respectively, and we observed density for U10 and UUC ssRNAs in the RNA-binding tunnel (Figures S11A and S11C). Contrary to the $MLE_{\Delta G}^{apo}$ structure, dsRBD2 now occupies the same position on the helicase module as in the MLE_{core}^{ATP} + U10 structure (Figures S11B and S11D). Furthermore, dsRBD1 and the linker connecting dsRBD1,2 are not visible, suggesting that this region remains flexible in the presence of ssRNA.

Thus, $MLE_{\Delta G}$ can bind ssRNAs, but the SL7 ssRNA moiety is not bound in the $MLE_{\Delta G}^{ATP}$ + SL7 structure due to clashes of RecA2 and OB-like β-hairpins with the putative binding sites of nucleotides 3–7 (Figure S11E). Comparison of $MLE_{\Delta G}^{ATP}$ + SL7 and $MLE_{\Delta G}^{ATP}$ + UUC structures revealed that for ssRNA insertion into the helicase module and transition from the dsRNA- to ssRNA-bound state, the RecA2 domain and HA2-OB-L3 module must move inward toward the ssRNA-binding tunnel accompanied by rotations of 16° and 17°, respectively (Figure 4C; Video S4). These domain motions are facilitated by an extensive conformational change of dsRBD2 (Figures 4D and S12A). To bind ssRNA, dsRBD2 flips back onto the helicase module with a displacement of 60 Å while rotating by 146°. For simplicity, we call the conformation of dsRBD2 in the $MLE_{\Delta G}^{ATP}$ + SL7 structure "flipped-out" ("out" in construct names) and in complex with ssRNA as "flipped-in" ("in" in construct names). In the flipped-in conformation, dsRBD2 acts as a brace and connects the RecA2 domain and the HA2-OB-L3 modules by forming an intricate network of hydrogen bonds and hydrophobic interactions, which stabilize the ssRNA-bound closed state of MLE (Figures S12B and S12C).

To investigate the effect of ATP hydrolysis on $MLE_{\Delta G}$ bound to ssRNA, we determined the cryo-EM structure of MLE in complex with ADP and UUC ssRNA ($MLE_{\Delta G}^{ADP}$ + UUC). This structure superposed with $MLE_{\Delta G}^{ATP}$ + UUC with an overall RMSD of 0.7 Å but showed minor conformational changes at the 5′ end of the ssRNA and backbone of the RNA at positions 5–7 (Figures S13A and S13B). The RNA backbone at position 6–7 of the RNA bend is recognized by a conserved loop of RecA2 (via His691 and Gln693) that was referred to as the "hook" loop or motif H.[8] Previous mutation of the hook-loop residues impaired the helicase activity but not RNA binding of MLE.[8] However, a higher resolution will be needed to confidently conclude that such minor movements may facilitate RNA translocation through the helicase tunnel.

Our structural analyses suggest that dsRBD2 conformation change ("in" vs. "out") governs the preference of MLE for ds- or ssRNA and that it, in turn, should cause the release of dsRNA and vice versa. To validate this, we performed competition

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

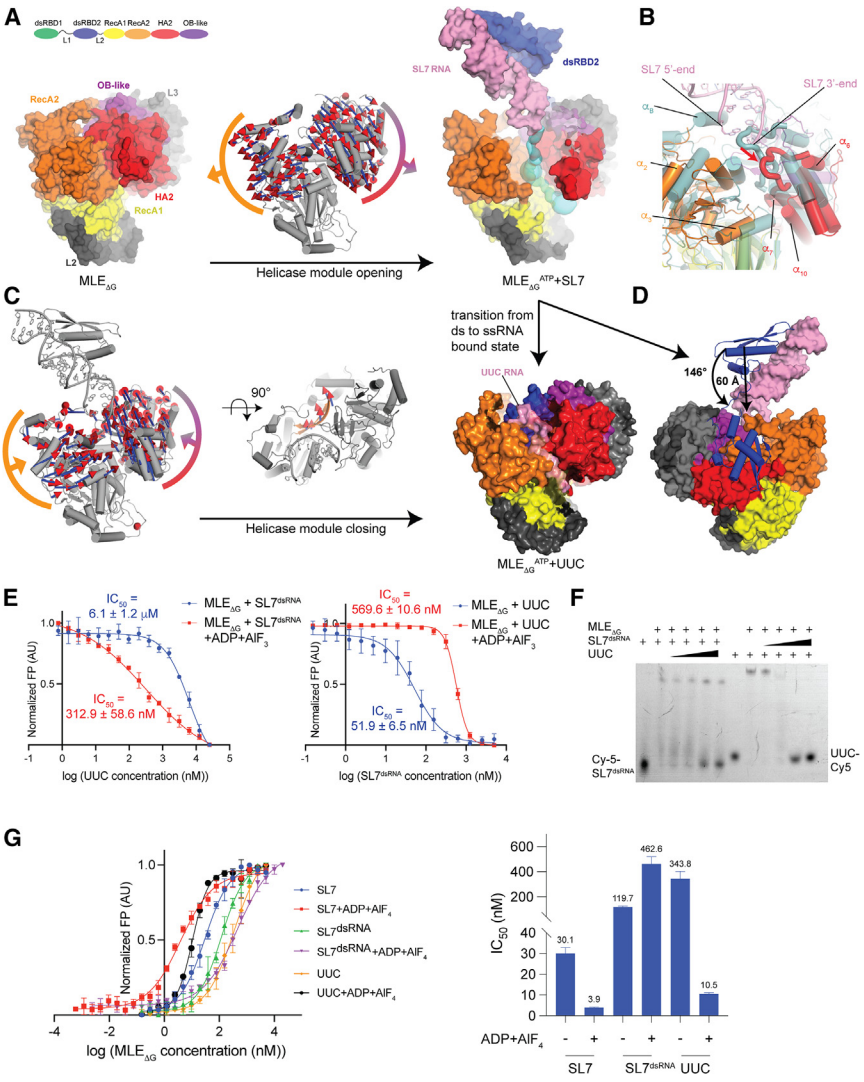


Figure 4. Opening and closing of MLE helicase module upon binding to dsRNA and ssRNA

(A) The MLE$_{\Delta G}$ and MLE$_{\Delta G}^{ATP}$ + SL7 structures are shown as surface representations. Movement of the RecA2 domain and the HA2-OB-L3 module upon dsRNA binding and helicase module opening is shown with C_α vectors from blue to red on the cartoon representation of MLE$_{\Delta G}$ structure (in gray, with RecA2 and HA2-OB-L3 module shown with orange and red-purple-gray arcs). The formed RNA-binding tunnel in MLE$_{\Delta G}^{ATP}$ + SL7 structures is shown in cyan. The neighboring regions around the RNA-binding tunnel are shown with a transparent surface.

(B) Zoom-in of the RNA-binding tunnel showing structural changes in RecA2 and HA2 domains upon dsRNA binding to MLE$_{\Delta G}$ (with respective domain colors) compared with MLE$_{\Delta G}$ (cyan) structure. The loop between α_6-α_7 helices from the OB-like domain is shown with a thicker radius for clarity with its displacement indicated by a red arrow.

(C) Closing of the helicase module upon transition from SL7 dsRNA to UUC ssRNA-bound state. The movement of RecA2 domain and HA2-OB-L3 module upon transition from ds to ssRNA-bound state is shown with C_α vectors from blue to red on the cartoon representation of MLE$_{\Delta G}^{ATP}$ + SL7 structure (in gray). The 90° top view depicts the movement of RecA2 β-hairpin (orange) out of the ssRNA-binding tunnel. The UUC ssRNA-bound state of MLE$_{\Delta G}$ is shown as surface representation with ssRNA in pink and neighboring region around ssRNA in transparent for clarity.

(D) Flipping of the dsRBD2 onto the helicase module is shown on the MLE$_{\Delta G}^{ATP}$ + SL7 structure upon transition from ssRNA- to dsRNA-bound state.

(E) Competition FP assays showing inhibition of SL7^{dsRNA} binding by increasing the concentration of UUC ssRNA and UUC RNA binding by increasing the concentration of SL7^{dsRNA} in presence and absence of transition state analog.

(F) EMSA experiments showing release of SL7^{dsRNA} by increasing the concentration of UUC RNA and UUC RNA by increasing the concentration of SL7^{dsRNA}.

(G) FP assay curves (left) and binding affinity (right) of SL7, SL7^{dsRNA}, and UUC ssRNA to MLE in presence and absence of ATP transition state analogs (ADP:AlF$_4$).

Data are represented as mean ± standard deviation in (E) and (G). See also Figures S1–S6, S10, and S12.

experiments with the double-stranded (SL7^{dsRNA}, Figure S7B) or single-stranded (UUC RNA) region of SL7 using fluorescence polarization (FP) and EMSA (Figures 4E and 4F). The SL7^{dsRNA} contained a 2-nucleotide overhang at the 3′ end, as this is involved in the interaction with dsRBD2-helicase module linker (Figure 2). FP experiments confirmed competitive inhibition of SL7^{dsRNA} binding by UUC ssRNA, which was further enhanced by ADP:AlF$_4$. SL7^{dsRNA} could reciprocally compete with UUC ssRNA, with reduced effectiveness in the presence of ADP:AlF$_4$ (Figure 4E). EMSA experiments confirmed this and showed the release of SL7^{dsRNA} upon increasing the concentration of UUC ssRNA and vice versa, in agreement with our structural analysis and FP assays (Figure 4F).

Our cryo-EM structures suggest that the position of dsRBD2 depends on RNA secondary structure. Competition experiments and structural data imply that dsRBD2 must release dsRNA and then flip back onto the core for MLE to bind the ssRNA. However, MLE tandem dsRBDs bind to dsRNA with nanomolar affinity[38,39] raising the question of how dsRBD2 dissociates from dsRNA. We explored the role of ATP binding by assessing the affinity of MLE$_{\Delta G}$ for SL7, SL7^{dsRNA}, and UUC RNA in the absence and presence of ADP:AlF$_4$ using FP assays (Figure 4G). As expected, SL7 and UUC RNA showed an ∼8- and ∼33-fold increase of affinity in the presence of ADP:AlF$_4$, likely due to enhanced binding of the ssRNA region. However, MLE$_{\Delta G}$ showed a ∼4-fold reduced affinity for SL7^{dsRNA} in the presence of ADP:AlF$_4$. Therefore, in the absence of ATP, MLE shows a ∼3-fold preference for the dsRNA part of SL7 (120 vs. 344 nM), but in the presence of ADP:AlF$_4$, it shows a ∼44-fold preference for ssRNA (463 vs. 10.5 nM), signifying the binding of ATP to the helicase module as a driving force for flipping dsRBD2 onto the helicase module.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

CellPress
OPEN ACCESS

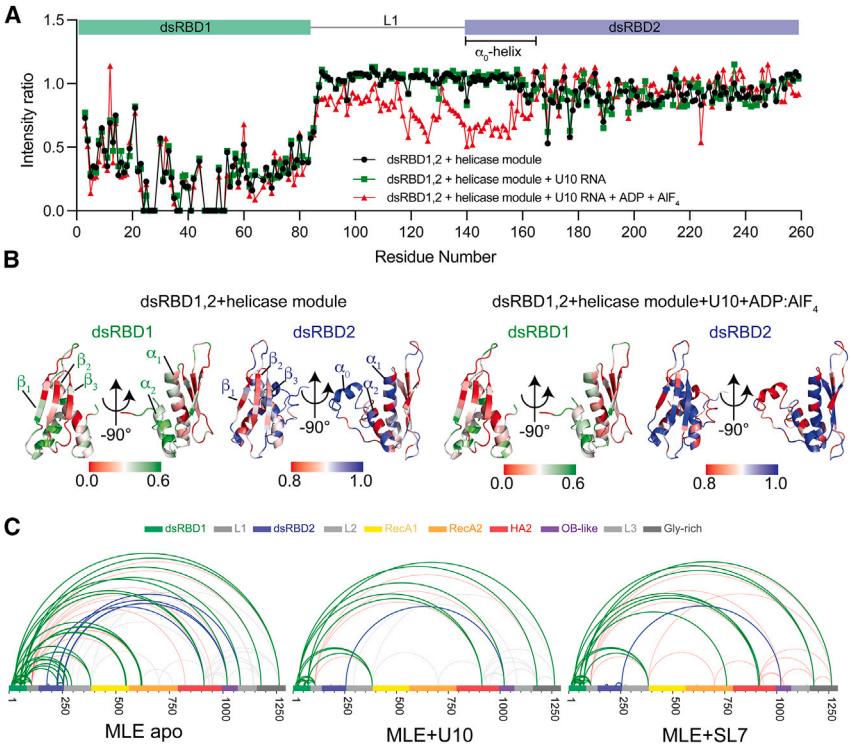


Figure 5. Interaction between MLE dsRBD1,2 and the helicase module

(A) Intensity ratios of dsRBD1,2 upon sequential addition of MLE helicase module (equimolar amount, in *trans*), U10 ssRNA, and ADP:AlF_4 and as quantified from spectra shown in Figure S14A.

(B) Mapping of line broadening and intensity decrease of NMR signals in dsRBD1,2 on the dsRBD1 and dsRBD2 NMR structures (PDB: 6I3R) upon addition of the helicase module, U10 ssRNA, and ADP:AlF_4.

(C) Interdomain crosslinks within MLE in the apo-state, in complex with U10 ssRNA or with SL7 dsRNA, are presented as arches. Crosslinks originating from dsRBD1, linker, and the dsRBD2 are shown in green, red, and blue for MLE, respectively. The crosslinking pattern suggests non-specific interactions between dsRBD1 and the helicase module, supporting the observed line broadening in the NMR spectrum and a dynamic exchange with the helicase module. See also Figures S12B and S12C and Tables S1 and S2.

The interactions of dsRBD1,2 with the helicase module are dynamic

Although our construct used for cryo-EM studies included dsRBD1, we did not observe any EM density, indicating its flexible orientation relative to the helicase module. This aligns with previous findings suggesting that dsRBD1 has weak dsRNA-binding affinity and a flexible linker connecting dsRBD1,2.[37,38] To explore potential transient dsRBD1-helicase module interactions that could not be captured in cryo-EM data, we monitored changes in the ^{1}H-^{15}N HSQC NMR spectrum of ^{15}N-labeled dsRBD1,2 upon addition of unlabeled helicase module in *trans*. Surprisingly, dsRBD1 peaks displayed severe line broadening compared with minor line broadening in the linker and dsRBD2, indicating that dsRBD1 interacts stronger with the helicase module than dsRBD2, contrary to our cryo-EM and previously available structures (Figures 5A, 5B, and S14A; Table S1).[8]

To investigate if RNA and nucleotides influence dsRBD1,2's interaction with the helicase module, we performed NMR titrations with U10 ssRNA and ADP:AlF_4. Although ssRNA alone had little impact, the presence of ssRNA and ADP:AlF_4 caused line broadening in linker residues and the α_0 helix of dsRBD2. The observed line broadening is consistent with our previous findings that this dynamic α_0 helix does not bind RNA but interacts with the helicase module.[8,38] By contrast, line broadening of dsRBD1 residues is unaffected by the presence of ADP:AlF_4 and ssRNA (Figures 5A and S14A; Table S1).

To reconcile the NMR and cryo-EM data regarding dsRBD-helicase interactions, we employed crosslinking mass spectrometry (CL-MS) in both the apo- and RNA-bound states (Figure 5C; Table S2). In the apo-state, 7 lysines from dsRBD1 and a single lysine from the linker region crosslink with 8 lysines from the helicase module and the G-patch. However, dsRBD2 lysine residues only crosslinked with three lysines, K936, K1020, and K1081, located in the HA2, OB-like, and L3 regions of the helicase module. Two of these lysines (K1020 and K1081) are within crosslinking distance of the previously known binding site of dsRBD2, as demonstrated by the cryo-EM structure of $MLE_{\Delta G}$ (Figure S14B). In the presence of U10 ssRNA or SL7 dsRNA, the dsRBD1-linker formed fewer crosslinks, but with similar distribution. On the other hand, K256 from dsRBD2 formed only one specific interdomain crosslink with K1020 of the OB-like domain. From our cryo-EM structures, this crosslink is possible if MLE binds to either ss or dsRNA. In summary, the data support our cryo-EM structures and suggest that dsRBD1-linker interacts dynamically with different parts of the helicase module, whereas dsRBD2 has limited flexibility in the absence of RNA and assumes a specific conformation in the presence of RNA.

Structure-based mutations of MLE affect RNA binding and helicase activity *in vitro*

We next sought to validate our structural findings through structure-based mutations. First, to investigate the role of lysine residues 253–256 of linker L2 in guiding RNA entry into the helicase module, we mutated them to glutamates ($MLE_{\Delta G}^{linker}$). Second, to test whether dsRBD2 conformation changes are essential for RNA unwinding, we designed mutations to lock MLE in either the open, dsRBD2 flipped-out conformation (competent for dsRNA binding), or the closed dsRBD2 flipped-in conformation (competent for ssRNA binding). To stabilize the closed conformation ($MLE_{\Delta G}^{in}$), we engineered a cysteine bridge between dsRBD2 and the helicase module (Figure S12B), whereas to force the open conformation ($MLE_{\Delta G}^{out}$), we mutated 13 residues on the helicase module to alanines, disrupting hydrogen bonds between dsRBD2 and the helicase module (Figure S12B). All mutants maintained their proper fold and stability (Figures S15A–S15C).

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

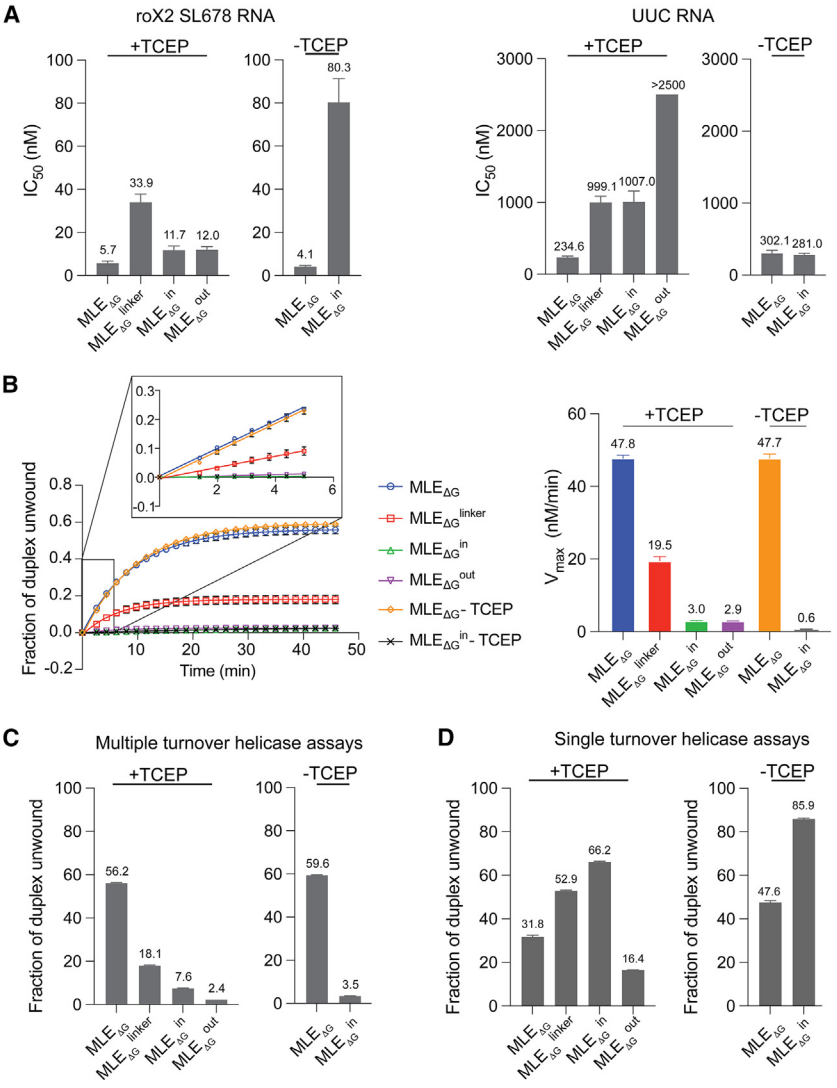


Figure 6. Biophysical and biochemical characterization of MLE mutants

(A) Affinity of $MLE_{\Delta G}$ wild type and mutants for roX2 SL678 and UUC RNA shown as half-maximal inhibitory concentration (IC_{50}) as determined by FP assays.

(B) Multiple-turnover helicase assays with $MLE_{\Delta G}$ and its mutants along with the fitting in the linker region to derive the maximum initial velocity (V_{max}) and the derived V_{max} values are shown. The time point at which ATP was added was adjusted to zero after recording baseline fluorescence for 3 min and the assays were recorded in triplicates. Every 2nd data point is shown in the helicase assay curve for simplicity. Errors that are smaller than the symbol of the data points are not displayed.

(C and D) (C) Fraction of duplex RNA unwound in multiple-turnover and (D) single-turnover helicase assays for $MLE_{\Delta G}$ and its mutants. Data are represented as mean ± standard deviation. See also Figure S17.

In FP binding assays with predominantly double-stranded roX2 SL678 or single-stranded UUC RNA (Figure 6A), $MLE_{\Delta G}^{linker}$ showed a ~4- to 6-fold reduced affinity for both RNAs, suggesting the importance of the positive surface charge in the dsRBD2-helicase module linker for RNA binding. Without reducing agent, $MLE_{\Delta G}^{in}$ showed a ~20-fold decreased binding affinity to SL678 compared with $MLE_{\Delta G}$ (80 vs. 4.1 nM), which was restored with tris(2-carboxyethyl)phosphine (TCEP) (11.7 vs. 5.7 nM). $MLE_{\Delta G}^{in}$ binding to UUC RNA remained largely unaffected compared with $MLE_{\Delta G}$ (280 vs. 302 nM), whether or not TCEP was present. By contrast, $MLE_{\Delta G}^{out}$ showed a severe loss of ssRNA-binding affinity (>2,500 nM) compared with $MLE_{\Delta G}$ (234 nM). Thus, these experiments confirm that $MLE_{\Delta G}^{in}$ preferentially binds ssRNA and $MLE_{\Delta G}^{out}$ binds dsRNA.

To assess MLE's helicase activity, we confirmed that the mutants retained their ATPase activity using 1D NMR experiments (Figure S16). $MLE_{\Delta G}$ did not show significant ATPase activity in the absence of RNA but was stimulated by the addition of SL678, as expected as the ATPase activity of DExH helicases has been shown to be stimulated by RNA.[46–49] The mutants showed a slight decrease in ATPase activity, and in the presence of TCEP, $MLE_{\Delta G}^{in}$ regained full ability to hydrolyze ATP equivalent to $MLE_{\Delta G}$, suggesting the release of dsRBD2 from the helicase module. Overall, these results indicate that all MLE mutants can bind and hydrolyze ATP.

Next, we assessed the real-time helicase activity of these mutants using a model duplex RNA derived from SL7 under multiple (RNA excess) and single (protein excess) turnover conditions (Figures 6B–6D and S7B). In multiple-turnover assays, $MLE_{\Delta G}^{linker}$ showed a modest 2.5-fold reduced activity compared with $MLE_{\Delta G}$ (19.5 ± 1.2 vs. 47.8 ± 0.8 nM/min). This decreased activity likely arises from weaker RNA substrate binding (Figure S17B). $MLE_{\Delta G}^{in}$ showed negligible activity in the absence of TCEP (0.6 ± 0.03 nM/min), which recovered to a minor extent in the presence of TCEP (3 ± 0.08 nM/min). This inability of $MLE_{\Delta G}^{in}$ to fully regain activity equivalent to wild type in the presence of TCEP could arise from transient formation of the cysteine bridge during the helicase cycle with dsRBD2 in the "in" conformation, limiting product release in a closed helicase conformation. $MLE_{\Delta G}^{out}$ also exhibited weak helicase activity (2.9 ± 0.09 nM/min). The decreased activity of the mutants was also reflected in a lower fraction of duplex unwound under multiple-turnover conditions where $MLE_{\Delta G}$ could unwind 56% of the dsRNA whereas $MLE_{\Delta G}^{linker}$, $MLE_{\Delta G}^{in}$, and $MLE_{\Delta G}^{out}$ showed only 18%, 3.5% (without TCEP), and 2.5% fraction of duplex RNA unwinding in 30 min, respectively.

In single-turnover experiments, $MLE_{\Delta G}$ showed a complex behavior compared with the mutants, making it difficult to estimate the total RNA unwound fraction and suggesting possible differences in the helicase mechanism between $MLE_{\Delta G}$ and its

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

mutants (Figure S17C). Only $MLE_{\Delta G}^{out}$ showed a decreased activity, as shown by a lower fraction of RNA duplex unwound. $MLE_{\Delta G}^{linker}$ and $MLE_{\Delta G}^{in}$ showed equivalent or higher fraction of duplex unwinding compared with $MLE_{\Delta G}$.

In summary and considering RNA-binding data, $MLE_{\Delta G}^{linker}$ is helicase active but has a general defect in RNA binding. $MLE_{\Delta G}^{out}$ is helicase activity-deficient, and $MLE_{\Delta G}^{in}$, despite high activity in single-turnover experiments, exhibits limited activity due to inhibited product release caused by the forced closed conformation of the helicase module.

The conformation of dsRBD2 is relevant for X chromosome territory formation and RNA binding *in vivo*

Next, we assessed the contribution of dsRBD2 and its adjacent linker region to MLE's *in vivo* function. MLE colocalizes with the MSL complex and roX RNA on the male X chromosome, where they bind to MRE sequences at active genes. Faithful localization to the X chromosome is revealed as staining of a coherent X chromosome territory by immunofluorescence microscopy. MLE binding to the MSL2-marked territory depends on its interaction with roX2,[8] and conversely, coherent MSL2-territory staining depends on the presence of MLE and its ATPase and helicase activities.[50]

We generated non-clonal S2 cell lines stably expressing GFP-tagged, full-length MLE variants MLE_{fl}^{WT} and MLE_{fl}^{linker}, and, in an independent approach, MLE_{fl}^{WT}, MLE_{fl}^{in}-eGFP, and MLE_{fl}^{out}-eGFP. The *mle* transgenes had been rendered RNAi-resistant by synonymous codon adaptation, allowing us to monitor their function after the depletion of endogenous MLE. RNAi directed against irrelevant glutathione-S-transferase (GST) sequences served as control (Figures S18A and S18B). The rescue capability of GFP-tagged transgenes was monitored by qualitative and quantitative immunofluorescence microscopy (Table S3). The pool of cells expressing low levels of GFP-tagged transgene (GFP^{low}), corresponding to sub-physiological levels of MLE, served as internal negative control.

As expected, depletion of MLE by RNAi impaired X chromosome territory formation, which also translated into an overall drop of MSL2 levels (panel S2 *mle* RNAi) (Figure 7A). Expression of MLE_{fl}^{WT}-eGFP restores MSL2-marked territories, the transgene localizes to X chromosome territories, and MSL2 nuclear levels are increased relative to the GFP^{low} cell pool (Figures 7A and 7B). MLE_{fl}^{linker}-eGFP cells, however, displayed an intermediate phenotype, and MSL2 showed a speckled and diffuse nuclear localization and no coherent X territory enrichment, suggesting that the positive surface charge in the L2 linker is important for MLE's RNA binding and unwinding function and is required for X chromosome enrichment.

Next, we tested the relevance of dsRBD2 conformations in MLE_{fl}^{in}-eGFP and MLE_{fl}^{out}-eGFP-expressing cells. As described above, we observed robust complementation of endogenous MLE loss in MLE_{fl}^{WT}-eGFP cells (Figures 7C and 7D). A comparable phenotype was observed in MLE_{fl}^{in}-eGFP-expressing cells, which seems in contrast to the observed deficiencies of this mutant *in vitro* (Figure 6). However, due to the intracellular reducing environment, the intended disulphide bonds in this mutant are most likely unstable *in vivo*, rendering MLE_{fl}^{in}-eGFP a wild type like variant. This result also demonstrates that the cysteine substitutions per se do not alter the physiological function of the enzyme. By contrast, we observed a strong phenotype in MLE_{fl}^{out}-eGFP-expressing cells (Figures 7C and 7D). In control *gst* RNAi-treated cells, endogenous MLE still localized to MSL2-stained X territories, whereas MLE_{fl}^{out}-eGFP did not colocalize and exhibited a speckled nuclear localization. In the absence of endogenous MLE, the MLE_{fl}^{out}-eGFP mutant failed to rescue X chromosome territory formation and restore physiological MSL2 protein levels. The data confirm that the helicase-deficient enzyme does not function *in vivo*.

MLE binds with high selectivity to tandem SL structures and poly(A)/U-rich ssRNA sequence motifs, which are present in critical parts of roX2.[30,31,34] We reasoned that the observed phenotypes in MLE_{fl}^{linker}-eGFP and MLE_{fl}^{out}-eGFP-expressing cells are the consequence of compromised roX2 recognition and incorporation into the MSL complex (Figures 7A–7D). Therefore, we assessed substrate binding in MLE_{fl}-eGFP-expressing cells by RNA immunoprecipitation (RIP) under native conditions. These experiments were performed in the presence of endogenous MLE to maintain a stable pool of MLE substrates in the cells. RIP of the abundant non-coding 7SK or GAPDH mRNA served as negative control. As expected, roX2 immunoprecipitated well with MLE_{fl}^{WT}-eGFP (Figures 7E, 7F, and S18C–S18F). MLE_{fl}^{linker}-eGFP displayed a moderate reduction in immunoprecipitated roX2 (Figures 7E, S18C, and S18D), in good agreement with the decrease in affinity to ssRNA/dsRNA observed *in vitro* (Figure 6A). In cells, diminished roX2 binding translates into an intermediate complementation efficiency (Figures 7A and 7B), which suggests that the linker adjacent to dsRBD2 contributes to guide the RNA substrate into the helicase tunnel. The MLE_{fl}^{in}-eGFP mutant, which in cells is considered wild type like, enriched roX2 well, even to a greater extent than MLE_{fl}^{WT}-eGFP (Figures 7F, S18E, and S18F). The MLE_{fl}^{out}-eGFP mutant, however, showed severely diminished roX2 binding, suggesting that disrupting dsRBD2-helicase module interactions causes helicase deficiency and limits faithful roX2 recognition and stable integration into the MSL complex (Figures 7F, S18E, and S18F). This observation is further strengthened by the fact that the dsRBD2 mutants are immunoprecipitated at higher levels than MLE_{fl}^{WT}-eGFP due to a larger fraction of mutant transgene-expressing cells (Figure S18F). Because roX2 is rendered unstable in the absence of functional MLE,[51,52] we propose that the observed moderate ability of MLE_{fl}^{linker}-eGFP and failure of MLE_{fl}^{out}-eGFP, respectively, to complement MLE loss is explained by a deficiency of roX2 incorporation into the dosage compensation complex (requiring dsRNA binding and helicase activity), leading to degradation of the RNA. In addition to roX1 and roX2, MLE binds—presumably transiently—to a small subset of additional substrates that harbor similar secondary structure and sequence determinants.[30,33,34] To test if other substrates are recognized in the same way, we quantified the enrichment of RpS29 with the MLE_{fl}-eGFP transgenes (Figures 7E, 7F, S18C, and S18E).[34] MLE_{fl}^{WT}-eGFP and MLE_{fl}^{in}-eGFP indeed enriched RpS29, albeit at lower levels than roX2, confirming our previous observation. By contrast, the MLE_{fl}^{linker}-eGFP and MLE_{fl}^{out}-eGFP mutants did not enrich RpS29, suggesting that a comparable recognition mode applies to at least two substrates of MLE, roX2, and RpS29.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

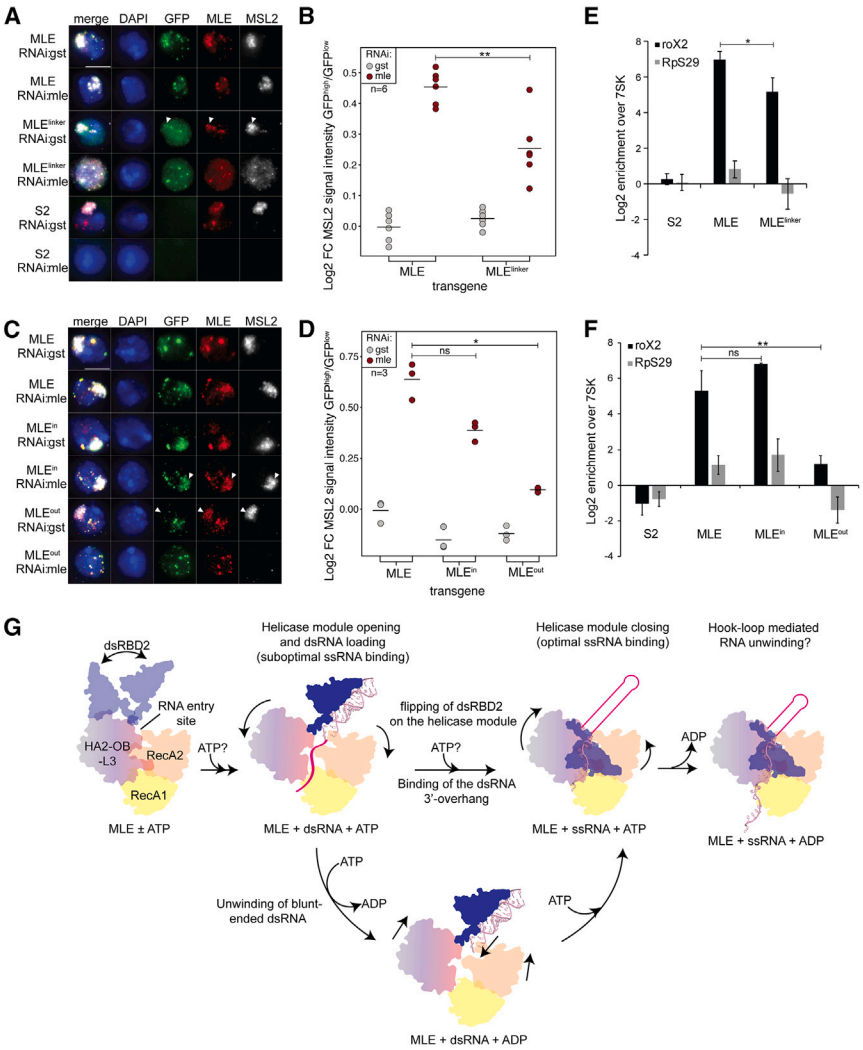


Figure 7. The conformation of dsRBD2 is relevant for X chromosome territory formation and RNA binding *in vivo*

(A and C) Representative immunofluorescence images of S2 cells stably expressing RNAi-resistant wild type or mutant MLE_{fl}-eGFP. Cells were treated with control (*gst*) dsRNA (endogenous and transgenic MLE present) or with dsRNA targeting endogenous *mle* (endogenous MLE depleted and transgenic MLE present). Panel "S2" shows immunofluorescence of non-transfected S2 control cells, which were treated the same way. DNA is shown in blue, transgenic MLE_{fl}-eGFP in green, MLE in red, and MSL2 in gray. Arrowheads mark X territories. Scale bar represents 5 μm.

(B and D) Dot plots showing the immunofluorescence-based complementation assay from (A) and (C). Each dot represents the $\log_2$-fold change of the population median MSL2 mean nuclear signal between GFP^{high} (expressed MLE_{fl}-eGFP wild-type or mutant transgene) and GFP^{low} cells per biological replicate. The number of cells included in the quantification is given in Table S3. p values were calculated using a two-tailed paired t test. ns, not significant; *p < 0.05; **p < 0.01.

(E and F) RNA immunoprecipitation (RIP) of GFP-tagged MLE_{fl}-wild type and mutants from stable S2 cells. S2 represents a GFP-RIP control using non-transfected S2 cell extract. Relative enrichment (IP/input) of roX2 and RpS29 transcripts was analyzed by RT-qPCR and is presented normalized to unbound 7SK. Error bars represent the standard deviation for three (E) and four (F) independent replicates. p values were calculated using a two-tailed paired t test. ns, not significant; *p < 0.05; **p < 0.01.

(G) A model for *cis* regulation of RNA binding and helicase activity by accessory domains in MLE. As the order of RNA and ATP binding by MLE remains unclear, the ATP binding by MLE is shown with a question mark. See also Figure S18 and Table S3.

DISCUSSION

The simple statement that RNA helicases unwind dsRNA contrasts with the observed complexity of domain organization and functional specification of many such enzymes. MLE, for example, exhibits remarkable substrate selectivity, favoring substrates with RNA secondary structure (pairs of SLs with alternative base pairing potential) with small linear motifs that are exposed only upon unwinding (the U-rich roX-box).[34] For the unwinding reaction to be productive, MLE needs to control the helicase reaction, to hold on to the remodeled RNA, and finally transfer the remodeled roX to the MSL complex.

We imagine that MLE's substrate selectivity and context-dependent helicase activity are mediated by its auxiliary domains, similar to nucleosome remodeling ATPases, where multiple auxiliary domains install conditional autoinhibition, substrate-mediated activation, and proof-reading.[53]

Our new structures reveal dramatic conformational changes of MLE upon binding to substrate dsRNA, leading to the conditional opening of the channel that binds the ssRNA products. Combined with previous studies, these structures allow us to propose a model for the helicase cycle of MLE. Below, we highlight the salient features of the model, as depicted in Figure 7G.

Hallmarks of the model

In our model (Figure 7G) the helicase module initially adopts a closed conformation that prevents non-substrate RNA binding. Connected by flexible linkers, the two dsRBD domains are free to sample the immediate vicinity for potential substrate RNAs. The distance between dsRBD2 and L2 linker ensures productive binding of captured RNAs by the helicase module, either at the 3′ end of dsRNA or within single-stranded loops, and dsRBD2 aligns the dsRNA with the helicase tunnel. Recognition of the terminal RNA nucleotides induces domain movement and opens the RNA-binding tunnel. In case of blunt-ended dsRNA substrates, particularly during the initial stages of RNA unwinding upon ATP hydrolysis, these conformational changes would facilitate directional 3′-5′ translocation of MLE on the dsRNA and entry of ssRNA into the helicase tunnel.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

For dsRNA substrates with 3′-ssRNA overhangs, the ssRNA component may not be optimally bound because of structural elements from the helicase module obstructing the RNA tunnel. Previous studies have indicated dsRBD2-independent binding of the helicase module to the single-stranded U20 substrates, albeit with lower affinity.[8] Flipping of dsRBD2 back onto the helicase module pulls back regions of RecA2 and HA2 domains, enabling optimal ssRNA binding in the tunnel that is conducive to helicase activity. Although the order of ATP and RNA binding by MLE remains uncertain, from our data, the dsRBD2 transition from "out" to "in" conformation seems to be driven by ATP binding to the helicase module. This conformational change is predicted to force the helicase module into the RNA-bound closed conformation, promoting helicase activity and stable binding of the ssRNA product.

Translocation mechanism of MLE

DExH helicases typically unwind RNA duplexes with 3′-5′ directionality with a requirement for a single-stranded 3′ overhang for substrate loading.[16,54] MLE, however, has the unusual ability to unwind dsRNAs with blunt ends dependent on the presence of the dsRBD2 domain.[8] Our cryo-EM structures of MLE in complex with SL7 RNA and ATP transition state analog or ADP shed light on the mechanism behind this unusual property. The continuous positively charged surface of dsRBD2 and L2 linker serves to hold the dsRNA, aligning it with the 3′ end and priming it for entry into the active site (Figure S8E). The limited linker length restricts dsRBD2's interaction with the base of larger RNA stems (Figure 2B). Initially, a few nucleotides are unwound with lower efficiency,[8] followed by dsRBD2 transitioning to the "in" conformation and efficient RNA unwinding. Notably, the dsRNA observed in our structures in complex with the SL7 RNA has a two-nucleotide overhang, which may dynamically occur in blunt-ended dsRNA substrates since the pairing of terminal bases is often weak and flexible.

A hook-loop or motif H from the RecA2 domain has been identified as crucial for RNA unwinding of substrates with ssRNA overhangs.[8] Relative domain motions of RecA2 with respect to RecA1-HA2 domains upon ATP hydrolysis were proposed to shift this motif toward the 5′ end of the RNA, rationalizing the mechanical movement of the 3′ end into the helicase tunnel. However, we observed only minor structural changes in MLE bound to ADP or ATP and ssRNA, requiring further high-resolution structures for deciphering MLE's translocation mechanism.

Auxiliary domains regulate access to the RNA tunnel in MLE

Although recent structures of DExH/DEAH RNA helicases within the spliceosome context have been determined, they do not contain dsRNA substrates. A distinct structure of the archael DExH helicase Hel308 bound to dsRNA exists, but it significantly differs in domain architecture and arrangement from its eukaryotic counterparts.[55]

Previous DExH/DEAH RNA helicase structures revealed that binding of the ATP transition state analog opens the ssRNA-binding tunnel by rearrangements of C-terminal domains.[9] By contrast, we did not observe substantial conformational changes upon binding of ADP:AlF_4, suggesting a different, ATP-independent mechanism for opening the ssRNA-binding tunnel. Such a mechanism has primarily been seen in the DEAH-type helicase DHX36, which unwinds G-quadruplexes.[56] Our $MLE_{\Delta G}$ structure complexed with dsRNA provides a snapshot of how dsRNA enters the RNA tunnel within this class of helicases: Besides their role in direct RNA binding, HA2 and the OB-like domains serve as a gate that requires large domain rearrangements to regulate RNA entry into the RNA-binding tunnel. The specific architecture of auxiliary domains relative to the helicase module undoubtedly contributes to substrate selectivity and regulation among helicases.

Selective binding of RNA by MLE could occur early during RNA recognition

In our MLE structures with ssRNAs, observed in this and previous studies,[8] the first two nucleotides of ssRNA bend toward and interact with the RecA2 and dsRBD2 domains. However, in the dsRNA-bound structure, the first nucleotide contacts the OB-like domain (Figure 2B). Examining the orientation of the RNA 3′ end in the dsRNA-bound structure and from its superposition with ssRNA-bound structures suggests that the 3′ end of dsRNA in the RNA tunnel follows a path starting from the position corresponding to the 3rd nucleotide in ssRNA-bound structures (Figure S11E). The recognition of this 3rd nucleotide in ssRNA-bound structures is specific and involves residues from the RecA2 and OB-like domains.[8] Thus, these auxiliary domains might serve as a specificity checkpoint, preventing non-substrate RNA binding. Previous mutations of H1032 and K1033 from the OB-like β_3-β_4 loop strongly reduced MLE's affinity not only to ssRNA but also to dsRNA.[8] This is explained by our MLE-dsRNA structure, as these two residues contact the 3′ end of dsRNA (Figure 2B). This mechanism, where auxiliary domains shape the RNA preferences, has recently been reported in the bacterial DExH helicase HrpB.[57]

Role of dsRBD2 in MLE beyond dsRNA recognition

Previously, dsRBD2's role in MLE function was solely associated with dsRNA binding.[8,37–39] Our structure-guided mutations helped us distinguish dsRBD2's contributions to dsRNA and ssRNA binding from its role in modulating the helicase activity. We found that the association of dsRBD2 with the helicase module triggers the closed conformation of MLE. This mode of action is distinct from its binding to dsRNA, where dsRBD2 must dissociate from the helicase module to allow it to open (Figure 3).

In MLE, the proposed gating mechanism relies on the conformational changes of dsRBD2, acting as a *cis*-regulatory element. Recently, a similar helicase activation mechanism, albeit in *trans*, was observed for Prp43, where a G-patch-containing protein acts as a brace between RecA2 and the winged helix domains, connecting two mobile parts of the helicase. *Trans* activation by related G-patch proteins has also been proposed for Prp2, DHX15, and DHX35.[58–60] These proteins, like many RNA helicases, require *trans* factors to enhance their inherent low helicase activity.[16,61,62] The comparison between different DExH helicases beautifully illustrates how evolution finds different solutions to a structural problem. MLE itself has a G-patch that binds RNA and is required for full helicase activity.[37] If helicase activity is controlled by direct interaction of the G-patch with another

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

domain in MLE or with the substrate RNA remains to be explored.

In striking contrast to many other helicases, such as Prp43, MLE shows a profound selectivity for uridine-rich sequences, forming base-specific contacts within the ssRNA channel. This raises the question of how such tightly bound ssRNA would be released during the helicase cycle and handed over to the MSL complex. We suggest a plausible mechanism: ssRNA may be released either stochastically or in a regulated manner when dsRBD2 detaches from the MLE helicase module, thereby forcing its opening. It is tempting to speculate that such a mechanism could be directly connected to the "handover" of unwound roX RNA to MSL2, which has been shown to also bind the U-rich roX-box.[34]

Potential role of dsRBD1 in mediating protein-protein interactions in MLE

Our NMR, CL-MS analysis, and cryo-EM structures showed that dsRBD1 interacts transiently and non-specifically with the helicase module (Figure 5) and does not bind RNA in the structures of this study. Although dsRBD1 was shown to be required for the localization of MLE and the MSL complex at the X chromosome, mutation of dsRBD1 RNA-binding residues did not interfere with this.[37–39] Our data support earlier suggestions of a potential role for dsRBD1 in mediating protein-protein interactions during MSL complex assembly. These factors could be the MSL1/MSL2 module, which is sufficient for the recruitment of MLE-roX2 to the MSL complex.[34] This would not be surprising, as several RNA-binding domains have been shown to have evolved away from their primary role of RNA binding and gained protein-protein interaction capabilities important for biological function.[63,64]

Implications for human DHX9

Understanding the mechanics of MLE helicase function also sheds light on the function of its human ortholog, DHX9/RNA helicase A (RHA). DXH9 and MLE show a similar domain organization with 51% sequence identity within structured domains. Conceivably, DHX9 is thus regulated by similar dsRBD2-mediated substrate-gating mechanisms. DHX9 is a multi-functional protein with roles in transcription regulation, mRNA translation, and microRNA (miRNA) processing.[65–68] It is also involved in the replication of various viruses, such as hepatitis C and HIV, and directly interacts with inverted *Alu* transposon elements to suppress the RNA processing defects arising due to *Alu* transposon insertions.[69–71] Considering its role in tumor cell maintenance, targeting of DHX9 could provide a novel chemotherapeutic approach.[72,73] The previously identified inhibitor aurintricarboxylic acid is a general, non-specific inhibitor of protein-RNA interactions.[74,75] Our study suggests that interfering with the interactions between dsRBD2 and the helicase module would provide a viable alternative strategy for drug development.

Limitations of the study

Our cryo-EM structure of the MLE-SL7 complex showed only the dsRNA region of SL7, despite expectations of simultaneous binding to the ssRNA portion based on our biochemical assays. We speculate that the simultaneous binding of MLE to both regions might be a transient state not easily captured by cryo-EM. Alternatively, MLE with dsRBD2 in an "out" conformation bound to the ssRNA region may exist in solution but could either dissociate or not enter the cryo-EM grid due to charge distribution issues caused by the exposed dsRNA region.

In our observations, changes in MLE and ssRNA conformations between ADP and ATP-bound states were minor. However, the resolution of our structures is above 3 Å, and higher-resolution data would be necessary to validate these conformational changes and elucidate structural alterations during ATP hydrolysis and MLE translocation on ssRNA. Considering that inorganic phosphate is present after ATP hydrolysis, the MLE complex with ssRNA, ADP, and inorganic phosphate could represent a more physiologically relevant post-catalytic state, potentially involving more substantial structural changes than observed between ATP and ADP-bound states. Moreover, the use of a high concentration of competitor strands in our multiple-turnover helicase assays might act as inhibitors of helicase activity. Therefore, these assays need to be interpreted as comparisons between wild-type and mutant proteins.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Bacterial strains
  - Insect Cells
  - Cell culture
- METHOD DETAILS
  - Protein expression and purification
  - *In vitro* transcription
  - Electromobility shift assays (EMSAs)
  - Fluorescence polarisation assays
  - Sample preparation for cryo-EM
  - Cryo-EM grid preparation and imaging
  - Cryo-EM data processing
  - Model building
  - NMR experiments
  - Cross-linking mass spectrometry (CLMS)
  - CLMS data analysis
  - Circular dichroism spectroscopy
  - Size exclusion chromatography – multi-angle light scattering (SEC-MALS)
  - Protein thermal stability measurements
  - Real-time fluorescence RNA helicase assay
  - Generation of stably transfected cell lines
  - *In vivo* RNA immunoprecipitation (RIP)
  - RNAi interference and immunocytochemistry
  - Fluorescence microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2023.10.026.

ACKNOWLEDGMENTS

We acknowledge support from Felix Weis and Wim Hagen for cryo-EM data collection and Bernd Simon for NMR data collection. We thank Raffaella Villa for providing RNAi-resistant *mle* cDNA and Silke Krause for technical support. We acknowledge the Protein Analytics Unit at the Biomedical Center, LMU Munich, for providing equipment, services, expertise, and assistance with data analysis. The work was supported by the EU Marie Curie Actions Cofund EIPOD fellowship to P.K.A.J., Deutsche Forschungsgemeinschaft (DFG) through grant Be1140/9-1 to P.B.B., an Emmy-Noether Fellowship to J.H. (HE7291/1-1), and grant HE7291/6-1 to J.H.

AUTHOR CONTRIBUTIONS

Conceptualization, P.K.A.J. and J.H.; funding acquisition, J.H. and P.B.B.; investigation, P.K.A.J., M.M., A.E.K., A.W.T., and K.L.; software, P.K.A.J. and A.W.T.; supervision, J.H., P.B.B., and M.B.; writing – original draft, P.K.A.J., J.H., P.B.B., and M.M.; writing – review and editing, P.K.A.J., M.M., A.E.K., A.W.T., K.L., M.B., P.B.B., and J.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 12, 2022
Revised: July 27, 2023
Accepted: October 18, 2023
Published: December 7, 2023

REFERENCES

1. Bleichert, F., and Baserga, S.J. (2007). The long unwinding road of RNA helicases. Mol. Cell *27*, 339–352. https://doi.org/10.1016/j.molcel.2007.07.014.
2. Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. Trends Biochem. Sci. *36*, 19–29. https://doi.org/10.1016/j.tibs.2010.07.008.
3. Jarmoskaite, I., and Russell, R. (2014). RNA helicase proteins as chaperones and remodelers. Annu. Rev. Biochem. *83*, 697–725. https://doi.org/10.1146/annurev-biochem-060713-035546.
4. Lorsch, J.R., and Herschlag, D. (1998). The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. Biochemistry *37*, 2194–2206. https://doi.org/10.1021/bi9724319.
5. Iggo, R.D., and Lane, D.P. (1989). Nuclear protein p68 is an RNA-dependent ATPase. EMBO J. *8*, 1827–1831. https://doi.org/10.1002/j.1460-2075.1989.tb03577.x.
6. Bourgeois, C.F., Mortreux, F., and Auboeuf, D. (2016). The multiple functions of RNA helicases as drivers and regulators of gene expression. Nat. Rev. Mol. Cell Biol. *17*, 426–438. https://doi.org/10.1038/nrm.2016.50.
7. Bohnsack, K.E., Yi, S., Venus, S., Jankowsky, E., and Bohnsack, M.T. (2023). Cellular functions of eukaryotic RNA helicases and their links to human diseases. Nat. Rev. Mol. Cell Biol. *24*, 749–769. https://doi.org/10.1038/s41580-023-00628-5.
8. Prabu, J.R., Müller, M., Thomae, A.W., Schüssler, S., Bonneau, F., Becker, P.B., and Conti, E. (2015). Structure of the RNA helicase MLE reveals the molecular mechanisms for uridine specificity and RNA-ATP coupling. Mol. Cell *60*, 487–499. https://doi.org/10.1016/j.molcel.2015.10.011.
9. Tauchert, M.J., Fourmann, J.B., Lührmann, R., and Ficner, R. (2017). Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. eLife *6*, e21510. https://doi.org/10.7554/eLife.21510.
10. Schmitt, A., Hamann, F., Neumann, P., and Ficner, R. (2018). Crystal structure of the spliceosomal DEAH-box ATPase Prp2. Acta Crystallogr. D Struct. Biol. *74*, 643–654. https://doi.org/10.1107/S2059798318006356.
11. He, Y., Andersen, G.R., and Nielsen, K.H. (2010). Structural basis for the function of DEAH helicases. EMBO Rep. *11*, 180–186. https://doi.org/10.1038/embor.2010.11.
12. Cordin, O., Hahn, D., and Beggs, J.D. (2012). Structure, function and regulation of spliceosomal RNA helicases. Curr. Opin. Cell Biol. *24*, 431–438. https://doi.org/10.1016/j.ceb.2012.03.004.
13. Schütz, P., Wahlberg, E., Karlberg, T., Hammarström, M., Collins, R., Flores, A., and Schüler, H. (2010). Crystal structure of human RNA helicase A (DHX9): structural basis for unselective nucleotide base binding in a DEAD-box variant protein. J. Mol. Biol. *400*, 768–782. https://doi.org/10.1016/j.jmb.2010.05.046.
14. Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y., and Leulliot, N. (2010). Prp43p contains a processive helicase structural architecture with a specific regulatory domain. EMBO J. *29*, 2194–2204. https://doi.org/10.1038/emboj.2010.102.
15. Wagner, J.D., Jankowsky, E., Company, M., Pyle, A.M., and Abelson, J.N. (1998). The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. EMBO J. *17*, 2926–2937. https://doi.org/10.1093/emboj/17.10.2926.
16. Ozgur, S., Buchwald, G., Falk, S., Chakrabarti, S., Prabu, J.R., and Conti, E. (2015). The conformational plasticity of eukaryotic RNA-dependent ATPases. FEBS Journal *282*, 850–863. https://doi.org/10.1111/febs.13198.
17. Rudolph, M.G., and Klostermeier, D. (2015). When core competence is not enough: functional interplay of the DEAD-box helicase core with ancillary domains and auxiliary factors in RNA binding and unwinding. Biol. Chem. *396*, 849–865. https://doi.org/10.1515/hsz-2014-0277.
18. Absmeier, E., Becke, C., Wollenhaupt, J., Santos, K.F., and Wahl, M.C. (2017). Interplay of cis- and trans-regulatory mechanisms in the spliceosomal RNA helicase Brr2. Cell Cycle *16*, 100–112. https://doi.org/10.1080/15384101.2016.1255384.
19. Absmeier, E., Wollenhaupt, J., Mozaffari-Jovin, S., Becke, C., Lee, C.T., Preussner, M., Heyd, F., Urlaub, H., Lührmann, R., Santos, K.F., and Wahl, M.C. (2015). The large N-terminal region of the Brr2 RNA helicase guides productive spliceosome activation. Genes Dev. *29*, 2576–2587. https://doi.org/10.1101/gad.271528.115.
20. Absmeier, E., Rosenberger, L., Apelt, L., Becke, C., Santos, K.F., Stelzl, U., and Wahl, M.C. (2015). A noncanonical PWI domain in the N-terminal helicase-associated region of the spliceosomal Brr2 protein. Acta Crystallogr. D Biol. Crystallogr. *71*, 762–771. https://doi.org/10.1107/S1399004715001005.
21. Collins, R., Karlberg, T., Lehtiö, L., Schütz, P., van den Berg, S., Dahlgren, L.G., Hammarström, M., Weigelt, J., and Schüler, H. (2009). The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch. J. Biol. Chem. *284*, 10296–10300. https://doi.org/10.1074/jbc.C900018200.
22. Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., and Doudna, J.A. (2016). Autoinhibitory interdomain interactions and subfamily-specific extensions redefine the catalytic core of the human DEAD-box protein DDX3. J. Biol. Chem. *291*, 2412–2421. https://doi.org/10.1074/jbc.M115.700625.
23. Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. Mol. Cell *41*, 693–703. https://doi.org/10.1016/j.molcel.2011.02.010.
24. Gowravaram, M., Bonneau, F., Kanaan, J., Maciej, V.D., Fiorini, F., Raj, S., Croquette, V., Le Hir, H., and Chakrabarti, S. (2018). A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner. Nucleic Acids Res. *46*, 2648–2659. https://doi.org/10.1093/nar/gky040.

25. Santos, K.F., Jovin, S.M., Weber, G., Pena, V., Lührmann, R., and Wahl, M.C. (2012). Structural basis for functional cooperation between tandem helicase cassettes in Brr2-mediated remodeling of the spliceosome. Proc. Natl. Acad. Sci. USA *109*, 17418–17423. https://doi.org/10.1073/pnas.1208098109.

26. Kuroda, M.I., Kernan, M.J., Kreber, R., Ganetzky, B., and Baker, B.S. (1991). The maleless protein associates with the X chromosome to regulate dosage compensation in Drosophila. Cell *66*, 935–947.

27. Samata, M., and Akhtar, A. (2018). Dosage compensation of the X chromosome: A complex epigenetic assignment involving chromatin regulators and long noncoding RNAs. Annu. Rev. Biochem. *87*, 323–350. https://doi.org/10.1146/annurev-biochem-062917-011816.

28. Lucchesi, J.C., and Kuroda, M.I. (2015). Dosage compensation in Drosophila. Cold Spring Harb. Perspect. Biol. *7*, a019398. https://doi.org/10.1101/cshperspect.a019398.

29. Deng, X., and Meller, V.H. (2006). roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. Genetics *174*, 1859–1866. https://doi.org/10.1534/genetics.106.064568.

30. Ilik, I.A., Maticzka, D., Georgiev, P., Gutierrez, N.M., Backofen, R., and Akhtar, A. (2017). A mutually exclusive stem-loop arrangement in roX2 RNA is essential for X-chromosome regulation in Drosophila. Genes Dev. *31*, 1973–1987. https://doi.org/10.1101/gad.304600.117.

31. Maenner, S., Müller, M., Fröhlich, J., Langer, D., and Becker, P.B. (2013). ATP-dependent roX RNA remodeling by the helicase maleless enables specific association of MSL proteins. Mol. Cell *51*, 174–184. https://doi.org/10.1016/j.molcel.2013.06.011.

32. Militti, C., Maenner, S., Becker, P.B., and Gebauer, F. (2014). UNR facilitates the interaction of MLE with the lncRNA roX2 during Drosophila dosage compensation. Nat. Commun. *5*, 4762. https://doi.org/10.1038/ncomms5762.

33. Ilik, I.A., Quinn, J.J., Georgiev, P., Tavares-Cadete, F., Maticzka, D., Toscano, S., Wan, Y., Spitale, R.C., Luscombe, N., Backofen, R., et al. (2013). Tandem stem-loops in roX RNAs act together to mediate X chromosome dosage compensation in Drosophila. Mol. Cell *51*, 156–173. https://doi.org/10.1016/j.molcel.2013.07.001.

34. Müller, M., Schauer, T., Krause, S., Villa, R., Thomae, A.W., and Becker, P.B. (2020). Two-step mechanism for selective incorporation of lncRNA into a chromatin modifier. Nucleic Acids Res. *48*, 7483–7501. https://doi.org/10.1093/nar/gkaa492.

35. Cugusi, S., Li, Y., Jin, P., and Lucchesi, J.C. (2016). The drosophila helicase MLE targets hairpin structures in genomic transcripts. PLoS Genet. *12*, e1005761. https://doi.org/10.1371/journal.pgen.1005761.

36. Reenan, R.A., Hanrahan, C.J., and Ganetzky, B. (2000). The mle(napts) RNA helicase mutation in drosophila results in a splicing catastrophe of the para Na+ channel transcript in a region of RNA editing. Neuron *25*, 139–149.

37. Izzo, A., Regnard, C., Morales, V., Kremmer, E., and Becker, P.B. (2008). Structure-function analysis of the RNA helicase maleless. Nucleic Acids Res. *36*, 950–962. https://doi.org/10.1093/nar/gkm1108.

38. Ankush Jagtap, P.K.A., Müller, M., Masiewicz, P., von Bülow, S., Hollmann, N.M., Chen, P.C., Simon, B., Thomae, A.W., Becker, P.B., and Hennig, J. (2019). Structure, dynamics and roX2-lncRNA binding of tandem double-stranded RNA binding domains dsRBD1,2 of Drosophila helicase maleless. Nucleic Acids Res. *47*, 4319–4333. https://doi.org/10.1093/nar/gkz125.

39. Lv, M., Yao, Y., Li, F., Xu, L., Yang, L., Gong, Q., Xu, Y.Z., Shi, Y., Fan, Y.J., and Tang, Y. (2019). Structural insights reveal the specific recognition of roX RNA by the dsRNA-binding domains of the RNA helicase MLE and its indispensable role in dosage compensation in Drosophila. Nucleic Acids Res. *47*, 3142–3157. https://doi.org/10.1093/nar/gky1308.

40. Hamann, F., Enders, M., and Ficner, R. (2019). Structural basis for RNA translocation by DEAH-box ATPases. Nucleic Acids Res. *47*, 4349–4362. https://doi.org/10.1093/nar/gkz150.

41. Tauchert, M.J., Fourmann, J.B., Christian, H., Lührmann, R., and Ficner, R. (2016). Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. Acta Crystallogr. F Struct. Biol. Commun. *72*, 112–120. https://doi.org/10.1107/S2053230X15024498.

42. Caruthers, J.M., Johnson, E.R., and McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proc. Natl. Acad. Sci. USA *97*, 13080–13085. https://doi.org/10.1073/pnas.97.24.13080.

43. Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. Nature *472*, 238–242. https://doi.org/10.1038/nature09862.

44. Theissen, B., Karow, A.R., Köhler, J., Gubaev, A., and Klostermeier, D. (2008). Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. Proc. Natl. Acad. Sci. USA *105*, 548–553. https://doi.org/10.1073/pnas.0705488105.

45. Masliah, G., Barraud, P., and Allain, F.H. (2013). RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cell. Mol. Life Sci. *70*, 1875–1895. https://doi.org/10.1007/s00018-012-1119-x.

46. Tanaka, N., and Schwer, B. (2006). Mutations in PRP43 that uncouple RNA-dependent NTPase activity and pre-mRNA splicing function. Biochemistry *45*, 6510–6521. https://doi.org/10.1021/bi052656g.

47. Tanaka, N., and Schwer, B. (2005). Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22. Biochemistry *44*, 9795–9803. https://doi.org/10.1021/bi050407m.

48. Schwer, B., and Guthrie, C. (1991). PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. Nature *349*, 494–499. https://doi.org/10.1038/349494a0.

49. Kim, S.H., Smith, J., Claude, A., and Lin, R.J. (1992). The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. EMBO J. *11*, 2319–2326. https://doi.org/10.1002/j.1460-2075.1992.tb05291.x.

50. Gu, W., Wei, X., Pannuti, A., and Lucchesi, J.C. (2000). Targeting the chromatin-remodeling MSL complex of Drosophila to its sites of action on the X chromosome requires both acetyl transferase and ATPase activities. EMBO J. *19*, 5202–5211. https://doi.org/10.1093/emboj/19.19.5202.

51. Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Mol. Cell *4*, 117–122. https://doi.org/10.1016/s1097-2765(00)80193-8.

52. Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. (2000). Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. Curr. Biol. *10*, 136–143.

53. Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell Biol. *18*, 407–422. https://doi.org/10.1038/nrm.2017.26.

54. Pyle, A.M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. Annu. Rev. Biophys. *37*, 317–336. https://doi.org/10.1146/annurev.biophys.37.032807.125908.

55. Büttner, K., Nehring, S., and Hopfner, K.P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. Nat. Struct. Mol. Biol. *14*, 647–652. https://doi.org/10.1038/nsmb1246.

56. Chen, M.C., Tippana, R., Demeshkina, N.A., Murat, P., Balasubramanian, S., Myong, S., and Ferré-D'Amaré, A.R. (2018). Structural basis of

G-quadruplex unfolding by the DEAH/RHA helicase DHX36. Nature *558*, 465–469. https://doi.org/10.1038/s41586-018-0209-9.

57. Hausmann, S., Geiser, J., Vadas, O., Ducret, V., Perron, K., and Valentini, M. (2020). Auxiliary domains of the HrpB bacterial DExH-box helicase shape its RNA preferences. RNA Biol. *17*, 637–650. https://doi.org/10.1080/15476286.2020.1720376.

58. Studer, M.K., Ivanović, L., Weber, M.E., Marti, S., and Jonas, S. (2020). Structural basis for DEAH-helicase activation by G-patch proteins. Proc. Natl. Acad. Sci. USA *117*, 7159–7170. https://doi.org/10.1073/pnas.1913880117.

59. Bohnsack, K.E., Ficner, R., Bohnsack, M.T., and Jonas, S. (2021). Regulation of DEAH-box RNA helicases by G-patch proteins. Biol. Chem. *402*, 561–579. https://doi.org/10.1515/hsz-2020-0338.

60. Warkocki, Z., Schneider, C., Mozaffari-Jovin, S., Schmitzová, J., Höbartner, C., Fabrizio, P., and Lührmann, R. (2015). The G-patch protein Spp2 couples the spliceosome-stimulated ATPase activity of the DEAH-box protein Prp2 to catalytic activation of the spliceosome. Genes Dev. *29*, 94–107. https://doi.org/10.1101/gad.253070.114.

61. Silverman, E., Edwalds-Gilbert, G., and Lin, R.J. (2003). DExD/H-box proteins and their partners: helping RNA helicases unwind. Gene *312*, 1–16. https://doi.org/10.1016/s0378-1119(03)00626-7.

62. Sloan, K.E., and Bohnsack, M.T. (2018). Unravelling the mechanisms of RNA helicase regulation. Trends Biochem. Sci. *43*, 237–250. https://doi.org/10.1016/j.tibs.2018.02.001.

63. Loerch, S., and Kielkopf, C.L. (2016). Unmasking the U2AF homology motif family: a bona fide protein-protein interaction motif in disguise. RNA *22*, 1795–1807. https://doi.org/10.1261/rna.057950.116.

64. Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol. Cell Biol. *8*, 479–490. https://doi.org/10.1038/nrm2178.

65. Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D., and Montminy, M. (1997). RNA helicase A mediates association of CBP with RNA polymerase II. Cell *90*, 1107–1112. https://doi.org/10.1016/s0092-8674(00)80376-1.

66. Anderson, S.F., Schlegel, B.P., Nakajima, T., Wolpin, E.S., and Parvin, J.D. (1998). BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. Nat. Genet. *19*, 254–256. https://doi.org/10.1038/930.

67. Hartman, T.R., Qian, S., Bolinger, C., Fernandez, S., Schoenberg, D.R., and Boris-Lawrie, K. (2006). RNA helicase A is necessary for translation of selected messenger RNAs. Nat. Struct. Mol. Biol. *13*, 509–516. https://doi.org/10.1038/nsmb1092.

68. Robb, G.B., and Rana, T.M. (2007). RNA helicase A interacts with RISC in human cells and functions in RISC loading. Mol. Cell *26*, 523–537. https://doi.org/10.1016/j.molcel.2007.04.016.

69. Tang, H., Gaietta, G.M., Fischer, W.H., Ellisman, M.H., and Wong-Staal, F. (1997). A cellular cofactor for the constitutive transport element of type D retrovirus. Science *276*, 1412–1415. https://doi.org/10.1126/science.276.5317.1412.

70. Isken, O., Grassmann, C.W., Sarisky, R.T., Kann, M., Zhang, S., Grosse, F., Kao, P.N., and Behrens, S.E. (2003). Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. EMBO J. *22*, 5655–5665. https://doi.org/10.1093/emboj/cdg562.

71. Aktaş, T., Avşar Ilık, İ., Maticzka, D., Bhardwaj, V., Pessoa Rodrigues, C., Mittler, G., Manke, T., Backofen, R., and Akhtar, A. (2017). DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. Nature *544*, 115–119. https://doi.org/10.1038/nature21715.

72. Gulliver, C., Hoffmann, R., and Baillie, G.S. (2020). The enigmatic helicase DHX9 and its association with the hallmarks of cancer. Future Sci. OA *7*, FSO650. https://doi.org/10.2144/fsoa-2020-0140.

73. Lee, T., Paquet, M., Larsson, O., and Pelletier, J. (2016). Tumor cell survival dependence on the DHX9 DExH-box helicase. Oncogene *35*, 5093–5105. https://doi.org/10.1038/onc.2016.52.

74. González, R.G., Haxo, R.S., and Schleich, T. (1980). Mechanism of action of polymeric aurintricarboxylic acid, a potent inhibitor of protein–nucleic acid interactions. Biochemistry *19*, 4299–4303. https://doi.org/10.1021/bi00559a023.

75. Cencic, R., Senechal, P., and Pelletier, J. (2015). Establishment of a primary screening assay for the DHX9 helicase. Comb. Chem. High Throughput Screen. *18*, 855–861. https://doi.org/10.2174/1386207318666151019093433.

76. Straub, T., Neumann, M.F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P.B. (2005). Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. Chromosoma *114*, 352–364. https://doi.org/10.1007/s00412-005-0020-x.

77. Hallacli, E., Lipp, M., Georgiev, P., Spielman, C., Cusack, S., Akhtar, A., and Kadlec, J. (2012). Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in *Drosophila*. Mol. Cell *48*, 587–600. https://doi.org/10.1016/j.molcel.2012.09.014.

78. Quinn, J.J., Ilik, I.A., Qu, K., Georgiev, P., Chu, C., Akhtar, A., and Chang, H.Y. (2014). Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. Nat. Biotechnol. *32*, 933–940. https://doi.org/10.1038/nbt.2943.

79. Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS Genet. *4*, e1000302. https://doi.org/10.1371/journal.pgen.1000302.

80. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). Phenix: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221. https://doi.org/10.1107/S0907444909052925.

81. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/S0907444904019158.

82. Forné, I., Ludwigsen, J., Imhof, A., Becker, P.B., and Mueller-Planitz, F. (2012). Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. Mol. Cell. Proteomics *11*, M111.012088. https://doi.org/10.1074/mcp.M111.012088.

83. Mueller-Planitz, F. (2015). Crossfinder-assisted mapping of protein crosslinks formed by site-specifically incorporated crosslinkers. Bioinformatics *31*, 2043–2045. https://doi.org/10.1093/bioinformatics/btv083.

84. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

85. Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530. https://doi.org/10.1016/j.jsb.2012.09.006.

86. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

87. Skinner, S.P., Fogh, R.H., Boucher, W., Ragan, T.J., Mureddu, L.G., and Vuister, G.W. (2016). CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. J. Biomol. NMR *66*, 111–124. https://doi.org/10.1007/s10858-016-0060-y.

88. Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696. https://doi.org/10.1002/prot.20449.

89. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

90. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

91. Grimm, M., Zimniak, T., Kahraman, A., and Herzog, F. (2015). xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. Nucleic Acids Res. *43*, W362–W369. https://doi.org/10.1093/nar/gkv463.

92. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. *50*, D543–D552. https://doi.org/10.1093/nar/gkab1038.

93. Qiu, C., Liu, W.-Y., and Xu, Y.-Z. (2015). Fluorescence Labeling of Short RNA by Oxidation at the 3′-End. In RNA Nanotechnology and Therapeutics: Methods and Protocols, P. Guo and F. Haque, eds. (Springer), pp. 113–120. https://doi.org/10.1007/978-1-4939-2562-9_8.

94. Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. Nat. Methods *16*, 1146–1152. https://doi.org/10.1038/s41592-019-0580-y.

95. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

96. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25. https://doi.org/10.1002/pro.3235.

97. Harrer, N., Schindler, C.E.M., Bruetzel, L.K., Forné, I., Ludwigsen, J., Imhof, A., Zacharias, M., Lipfert, J., and Mueller-Planitz, F. (2018). Structural architecture of the nucleosome remodeler ISWI determined from cross-linking, mass spectrometry, SAXS, and modeling. Structure *26*, 282–294.e6. https://doi.org/10.1016/j.str.2017.12.015.

98. Tani, H., Fujita, O., Furuta, A., Matsuda, Y., Miyata, R., Akimitsu, N., Tanaka, J., Tsuneda, S., Sekiguchi, Y., and Noda, N. (2010). Real-time monitoring of RNA helicase activity using fluorescence resonance energy transfer in vitro. Biochem. Biophys. Res. Commun. *393*, 131–136. https://doi.org/10.1016/j.bbrc.2010.01.100.

99. Maiato, H., Sunkel, C.E., and Earnshaw, W.C. (2003). Dissecting mitosis by RNAi in Drosophila tissue culture cells. Biol. Proced. Online *5*, 153–161. https://doi.org/10.1251/bpo57.

100. Thomae, A.W., Schade, G.O., Padeken, J., Borath, M., Vetter, I., Kremmer, E., Heun, P., and Imhof, A. (2013). A pair of centromeric proteins mediates reproductive isolation in Drosophila species. Dev. Cell *27*, 412–424. https://doi.org/10.1016/j.devcel.2013.10.001.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-GFP	Roche	Cat# 11814460001; RRID:AB_390913
Rat anti-MLE 6E11	Izzo et al.[37]	N/A
Rabbit anti-MSL2	Straub et al.[76]	N/A
Mouse anti-Lamin (T40)	H. Saumweber	N/A
Donkey anti-mouse-Alexa488	Jackson Immuno Research	Cat# 715-545-151; RRID:AB_2341099
Donkey anti-rat-Cy3	Jackson Immuno Research	Cat# 712-165-153; RRID:AB_2340667
Donkey anti-rabbit-Alexa647	Jackson Immuno Research	Cat# 711-605-152; RRID:AB_2492288
ChromoTek GFP-Trap Agarose	proteintech	Cat# gtak; RRID:AB_2631357
Bacterial and virus strains		
E. coli chemically competent BL21(DE3)	ThermoFisher	C600003
E. coli chemically competent Dh5a-T1R	ThermoFisher	12297016
E. coli electrocompetent DH10Bac	ThermoFisher	10361012
Chemicals, peptides, and recombinant proteins		
Adenylyl-imidodiphosphate AMP-PNP	Roche	10102547001
Bis(sulfosuccinimidyl)suberate BS3 crosslinker	Thermo Fisher	A39266
pCp-Cy5	Jena Bioscience	NU-1706-Cy5
Sf-900™ II SFM	Thermo Fischer	10902088
T4-RNA ligase	Thermo Fischer	EL0021
Gibco Schneider's Drosophila medium	Thermo Fischer	21720001
FBS Supreme, fetal bovine serum	PAN-Biotech	P30-3031
SDB-RPS stage tip material	3 M Empore	PN 2241
Trypsin/Lys-C Mix, Mass Spec Grade	Promega	V5073
Effectene Transfection Reagent	Qiagen	301425
Blasticidin S	gibco	A11139-03
Hygromycin B	ThermoFisher	10687010
SuperScript III First-Strand Synthesis System	ThermoFisher	18080-051
Fast SYBR Green master mix	Applied Biosystems	4385612
Formaldehyde, methanol-free	Polysciences	04018-1
Normal Goat Serum	Jackson Immuno Research	005-000-121
DAPI	ThermoFisher	D1306
VECTASHIELD Antifade Mounting Medium	Vector Laboratories	H-1000-10
Deposited data		
Cryo EM structure of $MLE_{\Delta G}^{apo}$	This study	PDB: 8B9L EMDB ID: 15935
Cryo EM structure of $MLE_{\Delta G}^{ATP}$	This study	PDB: 8B9J EMDB ID: 15933
Cryo EM structure of $MLE_{\Delta G}^{ATP}$+SL7 RNA	This study	PDB: 8B9K EMDB ID: 15934
Cryo EM structure of $MLE_{\Delta G}^{ATP}$+U10 RNA	This study	PDB: 8B9G EMDB ID: 15931
Cryo EM structure of $MLE_{\Delta G}^{ATP}$+UUC RNA	This study	PDB: 8B9I EMDB ID: 15932
Cryo EM structure of $MLE_{\Delta G}^{ADP}$+SL7 RNA	This study	PDB: 8PJJ EMDB ID: 17711

(*Continued on next page*)

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cryo EM structure of $MLE_{\Delta G}^{ADP}$+UUC RNA	This study	PDB: 8PJB EMDB ID: 17703
Microscopy image analysis	This study	Zenodo: https://doi.org/10.5281/zenodo.8378933
Crosslinking mass spectrometry dataset	This study	PRIDE (https://www.ebi.ac.uk/pride): PXD045725
Raw EMSA, Western blot and Immunofluorescence images	This study	Mendeley Data: https://doi.org/10.17632/cwj8zgxvmd.1
Experimental models: Cell lines		
Drosophila melanogaster S2 cells, subclone L2-4	Patrick Heun, Edinburgh	N/A
SF21 cells	Thermo Fisher	11497013
L2-4_MLE-eGFP-rescue	This study	N/A
L2-4_MLE^{in}-eGFP-rescue	This study	N/A
L2-4_MLE^{out}-eGFP-rescue	This study	N/A
L2-4_MLE^{linker}-eGFP-rescue	This study	N/A
Experimental models: Organisms/strains		
DH5α-$T1^{R}$	Invitrogen	12297016
DH10Bac	Thermo Fisher	10361012
BL21 (DE3)	Protein Expression and Purification Core Facility, EMBL, Heidelberg	N/A
Oligonucleotides		
MLE_257_NCOI_FP:TGAGCCATGGCGGATGAGCAG CTGAAGCCATATCCG	Sigma-Aldrich	N/A
dsRBD2_145_NCOI_FP:TGAGCCATGGCGGAGCAG AGGGACATGAACGAAGCGGAG	Sigma-Aldrich	N/A
MLE_FP_NCOI:TGAGCCATGGCGATGGATATAAAAT CTTTTTTGTACCAATTTTGTG	Sigma-Aldrich	N/A
MLE-K253-256E_FP: GAAGAGGATGAG CAGCTGAAGCCA	Sigma-Aldrich	N/A
MLE-K253-256E_RP: TTCCTCCAGAGTT CCACTAAAGGGC	Sigma-Aldrich	N/A
E195C_FP:ATATACACCAGTGGGCCCGTG CCACGCTAGGAGCTTTTTGG	Sigma-Aldrich	N/A
E195C_RP:CCAAAAAGCTCCTAGCGTGGCA CGGGCCCACTGGTGTATAT	Sigma-Aldrich	N/A
S633C_FP:TGGCCATGCTATCCGAATGCG ACGTTAGCTTTGAGC	Sigma-Aldrich	N/A
S633C_RP:GCTCAAAGCTAACGTCGCATT CGGATAGCATGGCCA	Sigma-Aldrich	N/A
R590E602D603E605E608_AAAAA_FP: AGTCGCGCCGAA AGGCCAAGGAAGTGGAGGACGAGGAGCAATTGCTTT CCGCGGCCAAGGCCGAGGCGGCAA TCAACTATAACAAGGTGTG	Sigma-Aldrich	N/A
CACACCTTGTTATAGTTGATTGCCGCCTCGGCCTTG GCCGCGGAAAGCAATTGCTCCTCGTCCTCCA CTTCCTTGGCCTTTCGGCGCGACT	Sigma-Aldrich	N/A
D634A_FP:ATGCTATCCGAATCGGCCGTTAG CTTTGAGCTG	Sigma-Aldrich	N/A
D634A_RP:CAGCTCAAAGCTAACGGCCGA TTCGGATAGCAT	Sigma-Aldrich	N/A

(*Continued on next page*)

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
H746N747T754_AAA_FP:GCGCATG AAGCTCTTTAC TTCAGCTGCCAACCTAACCAGCTACGCCGCAG TTTGGGCAAGCAAAACCAATTTGG	Sigma-Aldrich	N/A
H746N747T754_AAA_RP:CCAAATTGGT TTTGCTTGCCCAAACTGCGGCGTAGCTG GTTAGGTTGGCAGCTGAAGTAAAGAGCTTCATGCGC	Sigma-Aldrich	N/A
E790A_FP:GCCCGCTTCCAAGCGCTAGCGGA CAATCTTACGCCGGAGATG	Sigma-Aldrich	N/A
E790A_RP:CATCTCCGGCGTAAGATTGTCCGCT AGCGCTTGGAAGCGGGC	Sigma-Aldrich	N/A
E835A_FP:CCTCCGGTAGACGCAGTAATCGCA GCTGAGGTGTTGCTTCGCGAG	Sigma-Aldrich	N/A
E835A_RP:CTCGCGAAGCAACACCTCAGCTG CGATTACTGCGTCTACCGGAGG	Sigma-Aldrich	N/A
K1027A_FP:GCGCAAAGTGCTGACTACAGAGTCTG CAGCAGCGTTACTGCACAAAACCTCGG	Sigma-Aldrich	N/A
K1027A_RP:CCGAGGTTTTGTGCAGTAACGCTGCTG CAGACTCTGTAGTCAGCACTTTGCGC	Sigma-Aldrich	N/A
R1057A_FP:CTTCGTTTTCGGCGAGAAGATTGCCA CGCGAGCTGTTTCCTGCAAG	Sigma-Aldrich	N/A
R1057A_RP:CTTGCAGGAAACAGCTCGCGTGG CAATCTTCTCGCCGAAAACGAAG	Sigma-Aldrich	N/A
RNA sequence: SL7mod-UUC: GUGUAAAAUGUUGCU AGCAAAUAUAUAUGCUAGUAACGUU UUACGCCCUCUUUCUUUCUU	Integrated DNA Technologies	N/A
RNA sequence: SL7-up-BHQ-1: GUGUAAAAUGUUGCUAGCA-BHQ1	Biomers	N/A
RNA sequence: SL7-down-6FAM: 6-FAM-UGCUAGUA ACGUUUUACGCCCUCUUUCUUUCUU	Biomers	N/A
RNA sequence: SL7-up: GUGUAAAAUGUUGCUAGCA	Biomers	N/A
RNA sequence: U10 RNA: UUUUUUUUUU	Biomers	N/A
RNA sequence: UUC RNA:CCUCUUUCUUUC	Biomers	N/A
roX2-SL7.fw: GACGTGTAAAATGTTGCAAATTAAG	Biomers	Maenner et al.[31]
roX2-SL7.rv: TGACTGGTTAAGGCGCGTA	Biomers	Maenner et al.[31]
RpS29.fw: AGCGCATCGAAGCATTGATT	Biomers	This study
RpS29.rv: GCGGAAGGTGGTAACTGTTG	Biomers	This study
7SK.fw: GATAACCCGTCGTCATCCAG	Biomers	Hallacli et al.[77]
7SK.rv: AGTAATTCTGCCTGGCGTTG	Biomers	Hallacli et al.[77]
GAPDH.fw: GGAGCCACCTATGACGAAAT	Biomers	Quinn et al.[78]
GAPDH.rv: GTAGCCCAGGATTCCCTTC	Biomers	Quinn et al.[78]
mle RNAi.fw: TTAATACGACTCACTATAGGGAGAA TGGATATAAAATCTTTTTTGTACCAATTTTG	Sigma-Aldrich	Straub et al.[79]

(*Continued on next page*)

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
mle RNAi.rv: TTAATACGACTCACTATAGGGAGAACAGGGCGCATGACTTGCT	Sigma-Aldrich	Straub et al.[79]
gst RNAi.fw: TTAATACGACTCACTATAGGGAGAATGTCCCCTATACTAGGTTA	Sigma-Aldrich	Straub et al.[79]
gst RNAi.rv: TTAATACGACTCACTATAGGGAGAACGCATCCAGGCACATTG	Sigma-Aldrich	Straub et al.[79]
Recombinant DNA		
pFastBac-MLE-flag	Maenner et al.[31]	N/A
pFastBac-His-MLE$_{\Delta G}$	This study	N/A
pFastBac-His-MLE (257-1158)	This study	N/A
pFastBac-MLE$_{\Delta G}^{linker}$	This study	N/A
pFastBac-His-MLE$_{\Delta G}^{in}$	This study	N/A
pFastBac-His-MLE$_{\Delta G}^{out}$	This study	N/A
pHsp70-MLE-eGFP-rescue	This study	N/A
pHsp70-MLE^{in}-eGFP-rescue	This study	N/A
pHsp70-MLE^{out}-eGFP-rescue	This study	N/A
pHsp70-MLE^{linker}-eGFP-rescue	This study	N/A
pET-M11-Thioredoxin-dsRBD1,2	Ankush Jagtap et al.[38]	N/A
Software and algorithms		
Phenix	Adams et al.[80]	SBGrid Consortium
Coot	Emsley and Cowtan[81]	SBGrid Consortium
Crossfinder	Forné et al.[82] and Mueller-Planitz[83]	N/A
Cryosparc	Punjani et al.[84]	https://cryosparc.com/
Relion	Scheres[85]	https://github.com/3dem/relion
Pymol	Schrodinger, LLC	https://pymol.org/2/
RStudio[86]		https://www.r-project.org
Image Studio	LI-COR	N/A
Proteome discoverer 2.2	Thermo Fisher	N/A
GraphPad Prism 9	GraphPad software Inc.	https://www.graphpad.com/scientificsoftware/prism/
CcpNMR	Skinner et al.[87] and Vranken et al.[88]	https://ccpn.ac.uk/software/version-2/
Adobe Illustrator	Adobe	https://www.adobe.com/products/illustrator.html
NMRPipe	Delaglio et al.[89]	SBGrid Consortium
Fiji	Schindelin et al.[90]	N/A
xvis web	Grimm et al.[91]	https://xvis.genzentrum.lmu.de
Proteome exchange submission tool version 2.7.1	Perez-Riverol et al.[92]	https://www.ebi.ac.uk/pride/markdownpage/pridesubmissiontool
Other		
25 cm x 75 μm ID, 1.6 μm C18	IonOpticks	AUR2-25075C18A-CSI
ReproSil-Pur C18-AQ 2.4 μm	Dr. Maisch	r124.aq.
RSLCnano Ultimate 3000 system	Thermo Fisher	ULTIM3000RSLCNANO

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Janosch Hennig (janosch.hennig@embl.de).

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Materials availability

All newly generated material in this study is available upon request from the lead contact.

Data and code availability

- The $MLE_{\Delta G}^{apo}$, $MLE_{\Delta G}^{ATP}$, $MLE_{\Delta G}^{ATP}$+U10, $MLE_{\Delta G}^{ATP}$+UUC, $MLE_{\Delta G}^{ATP}$+SL7, $MLE_{\Delta G}^{ADP}$+UUC and $MLE_{\Delta G}^{ADP}$+SL7 structures are available from the PDB under accession codes 8B9L, 8B9J, 8B9G, 8B9I, 8B9K, 8PJB and 8PJJ and the cryo-EM density maps are available from the EMDB under the accession codes 15935, 15933, 15931, 15932, 15934, 17703 and 17711 respectively. The mass spectrometry crosslinking data have been deposited to the ProteomeXchange Consortium via the PRIDE[92] partner repository with the dataset identifier PXD045725. The rest of the data reported in this paper will be shared by the lead contact upon request. The raw images for EMSA gels, Western blots and Immunofluorescence images was deposited at Mendeley Data with DOI: https://doi.org/10.17632/cwj8zgxvmd.1
- Fiji macro and R code have been deposited at Zenodo and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains

E.coli strains DH5α-$T1^R$, BL21 (DE3) or DH10Bac were grown in standard LB medium at 37 °C supplemented with appropriate antibiotics.

Insect Cells

Insect cells (Sf21) were maintained at 27 °C in Sf900III medium (GIBCO).

Cell culture

Cell lines used in this study are listed in key resources table. *Drosophila melanogaster* male S2 cells (subclone L2-4, Patrick Heun, Edinburgh) were cultured in Schneider's *Drosophila* medium (Gibco) supplemented with 10% fetal bovine serum (PAN-Biotech) and penicillin-streptomycin at 26°C. Cell lines were tested negative for mycoplasma.

METHOD DETAILS

Protein expression and purification

N-terminal His_6-tagged $MLE_{\Delta G}$ (1-1158), His_6-tagged MLE helicase module (257-1158) and C-terminal flag-tagged MLE-full length and the mutant constructs and were cloned, expressed, and purified as described previously[38] in either pFastBac or pHsp70-eGFP plasmids. The mutants MLE^{linker} (K253-256E), MLE^{in} (E195C, S633C) and MLE^{out} (R590A, E602A, D603A, E605A, E608A, D634A, H746A, N747A, T754A, E790A, E835A, K1027A, R1057A) were generated by overlapping PCR and site-directed mutagenesis with primers containing the respective mutations in pFastBac $MLE_{\Delta G}$ and pHsp70-MLE_{fl}-eGFP plasmids. Around 700 bp of the original *mle* cDNA sequence in pHsp70-MLE_{fl}-eGFP (wild-type and mutants, respectively) was substituted with a mutagenized, synthesized DNA sequence to generate RNAi-resistant *mle* constructs.

The proteins were expressed in SF21 insect cells using recombinant baculovirus. The cells were harvested after three days of infection and were resuspended in lysis buffer (20 mM HEPES, 500 mM NaCl, 20 mM imidazole, 0.005% NP40, 2 mM β-mercaptoethanol, pH 7.5 with protease inhibitor and 10 units of Benzonase), lysed by sonication and spun down. The supernatant was loaded onto a HisTrap HP column (GE Healthcare) and eluted with an imidazole gradient in buffer containing 20 mM HEPES, 500 mM NaCl, 500 mM imidazole and 2 mM β-mercaptoethanol, pH 7.5. His_6-tag was removed by addition of TEV protease and simultaneous dialysis into a low salt buffer (20 mM HEPES, 100 mM NaCl, 2 mM β-mercaptoethanol, pH 7.5) overnight at 4°C. The next day, the samples were applied to a second affinity chromatography using a HisTrap HP column, and the flow-through fraction was further loaded onto a HiTrap Heparin column (GE Healthcare) to remove bound RNA. The protein was eluted from the HiTrap Heparin column with a salt gradient in the elution buffer (20 mM HEPES, 1 M NaCl, 2 mM β-mercaptoethanol, pH 7.5). Fractions containing MLE were concentrated with Amicon Ultra centrifugal filters with a molecular weight cut-off of 30 kDa, and the sample was applied to size-exclusion chromatography on the HiLoad 16/60 Superdex S200 pg column (GE) in 20 mM HEPES, 150 mM NaCl, 5% Glycerol, 1 mM DTT, pH 7.5. In case of C-terminal flag-tagged MLE-full length, the cells were lysed in 20 mM HEPES, 100 mM NaCl, 0.005% NP40, 1 mM β-mercaptoethanol, pH 7.5 with protease inhibitor and 10 units of Benzonase. The lysate was incubated with the flag-beads for one hour, the beads were washed several times with the lysis buffer and the protein was eluted with the lysis buffer with 3x Flag peptide.

For NMR experiments, dsRBD1,2 domains were expressed and purified as described previously.[38] Prior to NMR experiments, the MLE helicase module was dialysed in the NMR buffer containing 20 mM potassium phosphate, 200 mM NaCl, pH 6.5. To the final NMR samples, 10% D_2O was added for the deuterium lock.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

In vitro transcription

Stem-loop 678 roX2 RNA (SL678) (5'- ACAAUAUGCAAUACAAUACAAUACAAGACAAAAAAAUGU GUCUUGGAACGCAACAUU GUACAAGUCGCAAUGCAAACUGAAGUCUUAAAAGACGUGUAAAAUGUUGCAAAUUAAGCAAAUAUAUAUGCAUAUAUGGGUAA CGUUUUACGCGCCUUAACCAGUCAAAAUACAAAAUAAAUUGGUAAAUUUCAUAUAACUAGUGAAAUGUUAUACGAAACUUAAC AAUUGCCAAAUAA-3') was prepared by *in vitro* transcription using in-house produced T7 polymerase, rNTPs, and the RNA sequence cloned into pUC19 plasmid as template. The pUC19 DNA was linearised using EcoRI. After transcription, the reaction was cleaned by phenol/chloroform extraction. The RNA was purified by denaturing 6% urea PAGE and extracted from the gel by electro-elution. The final sample was concentrated and precipitated using ethanol. RNA pellet from the precipitate was dissolved in water and the RNA samples were stored frozen at -20ºC.

Electromobility shift assays (EMSAs)

For EMSA experiments, SL7 RNA (purchased from Integrated DNA Technologies (IDT)) or in vitro transcribed SL678 RNA was labelled with Cy-5. For this, 100 pmols of RNA was incubated with 150 pmoles of pCp-Cy-5 (Jena Biosciences), 20 units of T4-RNA ligase (NEB), 1x T4-RNA ligase buffer, and 1 mM ATP overnight at 4ºC. After labelling, RNA was purified by phenol-chloroform extraction and ethanol precipitation. Before EMSA experiments, the labelled RNA was heated to 94 ºC for 2 min in water and snap-cooled on ice for 10 mins. For $SL7^{dsRNA}$, unlabelled SL7-down RNA was mixed with equimolar 5'-Cy5 labelled SL7-up RNA (both purchased from Biomers.net GmbH) and annealed by heating to 95 °C for 30 sec and then slowly cooling to 4 °C. 3'-Cy5 labelled UUC RNA was purchased from Biomers.net GmbH. EMSA reactions were carried out with indicated amounts of the protein mixed with 5 nM of Cy-5 labelled RNA substrate. In case of competition experiments with UUC RNA, the complex of 80 nM $MLE_{\Delta G}$ and 5 nM $SL7^{dsRNA}$ and in case of $SL7^{dsRNA}$, the complex of 1.2 μM $MLE_{\Delta G}$ and 5 nM UUC RNA was incubated with increasing concentrations (0.05, 0.5, 5 and 10 μM) of UUC ssRNA or $SL7^{dsRNA}$ respectively. All reactions were carried out in 20 mM HEPES, 50 mM NaCl, 1 mM DTT, 10% glycerol and 0.005% IGEPAL CA-630, pH 7.5. The reactions were incubated for 30 mins at room temperature and were loaded on a native 6% polyacrylamide gel. The gel was imaged in the Typhoon FLA 900 imager (excitation 651 nm and emission 670 nm). The experiments were performed in duplicates.

Fluorescence polarisation assays

Fluorescence polarization assays were carried out in 20 mM HEPES, 100 mM NaCl, 0.005% IGEPAL CA-630, pH 7.5 and with Fluorescein RNAs. In vitro transcribed SL678 roX2 or SL7 RNA (purchased from IDT) was labelled at the 3'-end with Fluorescein.[93] Fluorescein labelled $SL7^{dsRNA}$ (labelled at 5'-end of SL7 down) or UUC RNA were purchased from Biomers.net GmbH. An increasing protein concentration was incubated with 2.5 nM of RNA for 30 mins in Corning 384 well plate in 40 μL volume. In case of competition experiments, a complex of 12.5 nM $MLE_{\Delta G}$ and 2.5 nM $SL7^{dsRNA}$ or 300 nM $MLE_{\Delta G}$ and 2.5 nM UUC RNA was incubated with increasing concentrations of UUC or $SL7^{dsRNA}$ respectively in the presence or absence of 1 mM $AlCl_3$, 10 mM NaF and 1 mM $MgCl_2$ (transition state analogue) for 30 mins in Corning 384 well plate in 40 μL volume. The fluorescence polarisation was measured in BioTek Synergy 4 plate reader with excitation and emission of 485 and 528 nm, respectively. Data were plotted in Graph Prism v9 and were fit to the Sigmoidal 4PL equation (where X is log (concentration)) ($Y = Bottom + \frac{(Top-Bottom)}{(1+10^{((logIC_{50}-X)\times HillSlope)})}$) to calculate the IC_{50} values. The experiments were performed in triplicates.

Sample preparation for cryo-EM

Samples for cryo-EM grid preparation were prepared in the EM buffer (20 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 1 mM ADP, 2 mM $MgCl_2$, 0.005 % triton-X 100, and 0.5% glycerol ± transition state analogue (1 mM $AlCl_3$, 10 mM NaF)). RNAs used for the cryo-EM sample preparation were purchased from either Integrated DNA technologies (SL7 RNA, GUGUAAAAUGUUGCUAGC AAAUAUAUAUGC UAGUAACGUUUUACGCCCUCUUUCUUUCUU) or Biomers (UUC RNA, CCUCUUUCUUUC and U10-mer RNA). For preparation of the MLE-SL7 RNA complex, RNA was heated to 95ºC for 2 mins and snap-cooled on ice. The MLE-RNA complex was prepared by mixing 4.7 μM $MLE_{\Delta G}$ with 5.2 μM (1.1 molar excess) RNA in the EM buffer. Before applying to the grid, the samples were incubated on ice for 30 mins. EM sample for $MLE_{\Delta G}{}^{apo}$ was prepared without ADP and transition state analogue.

Cryo-EM grid preparation and imaging

Quantifoil grids were plasma cleaned with 90% Argon, and 10% Oxygen plasma for 30 sec. For $MLE_{\Delta G}{}^{apo}$ and $MLE_{\Delta G}$ samples, Quantifoil UltrAufoil grids (R 2/2, 200 mesh) and for $MLE_{\Delta G}$ in complex with RNA, Quantifoil holey carbon grids (R2/1, 200 mesh) were used. 3.5 μl of the sample was applied to the grids, and the grids were plunged into liquid ethane using Vitrobot Mark IV at 6ºC and 100% humidity. Cryo-EM data were collected on an FEI Titan Krios microscope operated at 300 kV, equipped with a K3 detector operating in counting mode and a post-column Gatan Bioquantum energy filter. For $MLE_{\Delta G}{}^{apo}$, $MLE_{\Delta G}{}^{ATP}$, $MLE_{\Delta G}{}^{ATP}$+U10, $MLE_{\Delta G}{}^{ATP}$+UUC, $MLE_{\Delta G}{}^{ADP}$+UUC and $MLE_{\Delta G}{}^{ADP}$+SL7, 5313, 5894, 19545, 8880, 7901 and 7849 movies were collected. In the case of $MLE_{\Delta G}$+SL7, datasets were collected in two microscopy sessions, with 8761 and 6742 movies recorded in each session. All datasets were recorded at a nominal magnification of 105,000x, corresponding to 0.822 Å/pixel at the specimen level. In the case of $MLE_{\Delta G}{}^{ATP}$+SL7 RNA samples, initial screening datasets indicated severe preferred orientation with only one dominant view as apparent from 2D and 3D processing. As we could not resolve this issue by adding additives in buffers and grid

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

types with different surface treatments, cryo-EM datasets for the $MLE_{\Delta G}^{ATP}$+SL7 and $MLE_{\Delta G}^{ADP}$+SL7 samples were collected with 30 degrees stage pretilt.

Cryo-EM data processing

All movies were aligned with Patch motion correction, and CTF parameters were determined using Patch CTF in Cryosparc.[84] Particles were picked using Warp.[94] Particles were extracted from the CTF corrected micrographs using coordinates imported from Warp. Extracted particles were subjected to several rounds of 2D classification, and junk and low-resolution 2D classes were removed from further analysis. Selected particles from 2D classification were used to create multiple ab initio maps. The maps that showed high-resolution features were subjected to multiple rounds of heterogeneous refinement followed by homogeneous and non-uniform refinement in Cryosparc. In case of structures in complex with SL7 RNA, after ab initio reconstruction and heterogeneous refinement in Cryosparc, the particles were further 3D classified in Relion[85] followed by homogeneous and non-uniform refinement in Cryosparc.

Model building

All cryo-EM structures were modelled by manual rigid-body fitting of the atomic models of individual domains of dsRBD2, RecA1 and RecA2 and the HA2-OB-L3 module derived from the crystal structure of MLE_{core} in complex with ssRNA in UCSF Chimera in the initial stage of model building.[8,95] The dsRNA in the $MLE_{\Delta G}^{ATP}$+SL7 and $MLE_{\Delta G}^{ADP}$ +SL7 structure was built by placing an ideal A-form helix of 19 bp into the EM-density of the RNA. The modelling was further completed by multiple iterative rounds of manual model building in Coot interrupted by real-space refinement in Phenix software suite.[80,81] The structures were analysed in UCSF Chimera, UCSF ChimeraX and PyMOL and the figures were made using UCSF ChimeraX and PyMOL.[95,96] The centre-of-mass for each subdomain was determined using the python script (https://raw.githubusercontent.com/Pymol-Scripts/Pymol-script-repo/master/center_of_mass.py).

NMR experiments

^{1}H,^{15}N HSQC NMR experiments for dsRBD1,2 were performed with 0.01 mM ^{15}N labelled dsRBD1,2 on an Avance III Bruker NMR spectrometer equipped with a cryogenic probe head at 298 K and a magnetic field strength corresponding to a Larmor frequency of 800 MHz. ^{15}N labelled dsRBD1,2 was added to an equimolar ratio of unlabelled MLE helicase module in the presence or absence of two-fold molar excess of U10 RNA and ATP transition state analogue (1 mM $AlCl_3$, 10 mM NaF, 1 mM ADP and 2 mM $MgCl_2$). Data were processed with nmrPipe/nmrDraw and were analysed in the CCPN analysis software with chemical shifts obtained from the BMRB (ID: 34326).[88,89]

NMR ATPase assays were carried out in 20 mM deuterated Tris, 100 mM NaCl, 5 mM $MgCl_2$ and ± 2 mM TCEP and 1 mM ATP, pH 7.4 with 500 nM protein and 1 μM roX2 SL678 RNA for experiments carried out in the presence of RNA. 1D NMR experiments were performed on an Avance III Bruker NMR spectrometer equipped with a cryogenic probe head and magnetic field strengths corresponding to Larmor frequencies of 600 or 700 MHz and at 303 K. For each experiment, 32768 points in the direct dimension were recorded with eight scans, and for each sample, experiments were recorded as pseudo-2D experiments with data recorded every 60 seconds for a total of 120 mins. The data were processed in NMRPipe/NMRDraw, and the resulting 1D spectra were converted to text files using pipe2txt.tcl script from NMRPipe.[89] The peaks corresponding to the H8 proton of ATP and ADP were integrated using custom scripts, and the individual peak intensities were divided by the total intensity of ATP+ADP peaks to obtain the percentage of ATP remaining or the percentage of ADP produced. The error bars were derived from the signal to noise ratio of the peaks. The data were then fit using one phase decay or one phase association equations to obtain the rate of ADP produced or ATP hydrolysed in GraphPad Prism 9.

Cross-linking mass spectrometry (CLMS)

Purified MLE_{fl}-FLAG protein was centrifuged for 15 min at 21000 g at 4°C. MLE and RNA constructs of interest were incubated in a 1:2 molar ratio (0.7 μM MLE, 1.4 μM RNA) in MLE crosslinking buffer MXB-50 (20 mM HEPES pH 7.5, 50 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT) in presence of 1 mM Adenylyl-imidodiphosphate (AMP-PNP) and RNase inhibitor RNAsin (0.5 U, Promega) for 25 min at 4°C with head-over-end rotation. Samples were incubated with 1 mM Bis(sulfosuccinimidyl)suberate BS3 for 30 min at 30°C and 950 rpm. The crosslinking reaction was quenched by addition of 50 mM Tris-HCl/pH 8.0 and incubation for 15 min at 30°C and 950 rpm.

Before tryptic digestion, 1 M urea was added to the samples to allow partial unfolding of the protein. Trypsin/Lys-C Mix, Mass Spec Grade was added in a 1:50 ratio to MLE. Tryptic digestion was carried out overnight at 37°C with 500 rpm agitation in presence of 1 mM DTT. Then 4.3 mM iodoacetamide was added and incubated for 30 min at 25°C, 500 rpm in the dark. Iodoacetamide was quenched by addition of 20 mM DTT and incubation for 10 min at 25°C and 500 rpm. Samples were acidified by the addition of 0.05% trifluoracetic acid (TFA), the pH was adjusted to 1. SDB-RPS stage tip material was equilibrated by washes with 100% acetonitrile (ACN), activation buffer (30% methanol, 0.2% TFA) and equilibration buffer (0.2% TFA). Trypsinized samples were loaded to the equilibrated stage tips and centrifuged for 15 min at 500 g. After subsequent washes with wash buffer (1% TFA in isopropanol) and equilibration buffer, peptides were eluted into low protein binding Eppendorf reaction tubes with freshly prepared elution buffer (80% acetonitrile, 1.25% ammonia). After vacuum drying of the samples at 45°C, peptides were resuspended in MS sample buffer (0.3% TFA, 2% ACN in MS grade H_2O).

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Mass spectrometry was essentially performed as described.[97] Briefly, for LC-MS/MS analysis samples were injected in an RSLCnano Ultimate 3000 system and either separated in a 15-cm analytical column (75 μm ID home-packed with ReproSil-Pur C18-AQ 2.4 μm) with a 50-min gradient from 5 to 60% acetonitrile in 0.1% formic acid or in a 25-cm analytical column (75 μm ID, 1.6 μm C18) with a 50-min gradient from 2 to 37% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into a QexactiveHF (Thermo) operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 375–1600) were acquired with resolution R=60,000 at m/z 400 (AGC target of $3x10^6$). The ten most intense peptide ions with charge states between 3 and 5 were sequentially isolated to a target value of $1x10^5$, and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250°C; ion selection threshold, 33.000 counts.

CLMS data analysis

The raw data files were first converted by the proteome discoverer 2.2 (Thermo scientific) xlinkx workflow for crosslink detection into the.mgf file format. Next, the.mgf files were analyzed by crossfinder[82,83] applying the following filter parameters for identification of cross-linking candidates: False-discovery rate (FDR) ≤ 0.05, number of fragment ions per spectrum ≥ 4, number of fragment ions per peptide ≥ 2, fractional intensity of assigned MS2 peaks ≥ 0.05, relative filter score: 95. Crosslinks were visualized using the xvis web browser for arch plots.[91]

Circular dichroism spectroscopy

The circular dichroism spectra of $MLE_{\Delta G}$ and its mutants were recorded between 260 nm and 190 nm in a 0.2 mm cuvette, at 20°C at a concentration of 1 mg/ml in 20 mM Tris pH 7.5, 150 mM NaCl, 2% glycerol and ± 2 mM TCEP. The spectra were recorded using a Jasco J-815 CD spectrometer with a 50 nm/min scan speed, digital integration time of 1 sec with a 1 nm bandwidth and 10 accumulations. The data were plotted using GraphPad Prism 9.

Size exclusion chromatography – multi-angle light scattering (SEC-MALS)

50 μl of $MLE_{\Delta G}$ or the mutants were injected onto a Superdex 200 5/150 GL gel-filtration column (Cytiva) in 20 mM Tris pH 7.5, 100 mM NaCl ± 2 mM TCEP, at a flow rate of 0.3 ml/min and at room temperature. The column was connected to a MiniDAWN MALS detector and Optilab differential refractive index detector (Wyatt Technology). Data were analyzed using the Astra 7 software (Wyatt Technology) and was plotted using GraphPad Prism 9.

Protein thermal stability measurements

The thermal stability of the proteins was determined using Prometheus NT.48 nanoDSF instrument from NanoTemper Technologies. The protein was incubated at 1 mg/ml in 20 mM Tris, 100 mM NaCl ± 2 mM TCEP, pH 7.5 for 30 mins at room temperature and was loaded into the standard capillaries. The assay measured the tryptophan fluorescence at 330 nm and 350 nm from 20 - 90ºC with a 1ºC/min increase in temperature. The data were acquired and analysed with the PR. ThermControl v2.1.2 software provided with the instrument.

Real-time fluorescence RNA helicase assay

Multiple and single turnover real-time fluorescence RNA helicase assays were carried out using a slight modification of the protocol described previously.[98] The RNA is labelled with black hole quencher-1 (BHQ-1) at the 3'-end of the 1st strand and 6-carboxyfluorescein (6-FAM) at the 5'-end of the 2nd strand. Upon annealing the two strands, the 6'-FAM fluorescence signal is quenched by close-by BHQ-1. An increase of fluorescence signal intensity upon addition of MLE reports on helicase activity. The assays were carried out in 20 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 0.005% IGEPAL CA-630, pH 7.4 and ± 2 mM TCEP and RNA probe with BHQ-1 and 6-FAM quencher-dye pair. Before the assembly of the reaction, the SL7-up RNA strand containing BHQ-1 at the 3' end (GUGUAAAAUGUUGCUAGCA-BHQ1, Biomers) and SL7-down RNA strand containing 6-FAM at the 5' end (6-FAM-UGCUAGUAACGUUUUACGCCCUCUUUCUUUCUU, Biomers) were mixed in an equimolar ratio at 25 μM concentration in 20 mM Tris, 100 mM NaCl, pH 7.4, heated to 95ºC for 1 min and then slowly cooled down to 4ºC. Reactions without protein were used as a control. The assays were carried out in triplicates in 30 μL reaction volume in Corning 384 well plates and at 300 nM protein concentration and 5 nM duplex RNA in case of single turnover reactions or 1 μM duplex in case of multiple turnover reactions. To determine the appropriate amount of competitor excess to prevent rehybridization of the BHQ-1-6-FAM RNA probe upon separation by the helicase, the multiple turnover helicase assay was carried out in 0, 10, 25 and 50-fold excess unlabelled SL7-up DNA strand. 50-fold excess of the competitor strand resulted in the helicase activity (Figure S17A); therefore, the subsequent assays were performed in 25-fold excess of the competitor strand. The fluorescent intensity of the fluorescein was measured in the BioTek Synergy 4 plate reader (excitation at 485 nm and emission at 528 nm wavelength) at 27 ºC. 2 mM ATP was added to the reaction mix after initially monitoring the reaction for 3 mins. Data were fit in GraphPad Prism 9 to one phase association equation: $Y = Y_0 + (Plateau - Y_0)(1 - e^{(-K \times X)})$ except for $MLE_{\Delta G}$ single turnover experiments which showed a rise and fall behaviour and so the data where were fit to rise and fall to steady state equation: $Y = Steadystate(1 - De^{(-K_1 \times X)} + (D - 1)e^{(-K_2 \times X)}) + Baseline$.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Generation of stably transfected cell lines

Stable S2 cell lines expressing wild type or mutant MLE_{fl} fused to C-terminal GFP were generated as described with minor modifications.[8] Briefly, 500 ng pHsp70-MLE_{fl}-eGFP wild type or mutant plasmid was co-transfected with 25 ng plasmid encoding a blasticidin or hygromycin resistance gene using the Effectene transfection reagent (Qiagen). Stable MLE_{fl}-eGFP expressing cells were selected in complete medium containing 25 ng/ml blasticidin for a duration of three weeks, followed by recovery in complete medium lacking blasticidin for another week. In the following, cell lines were cultured in complete medium lacking blasticidin. Stable cell lines expressing MLE_{fl}^{WT}-eGFP or MLE_{fl}^{linker}-eGFP, respectively, were also established in a parallel approach using hygromycin (0.3 mg/ml) for a duration of three weeks, followed by maintenance in complete medium lacking hygromycin.

In vivo RNA immunoprecipitation (RIP)

Native RNA immunoprecipitation (RIP) of MLE_{fl}-eGFP and mutant derivatives was performed as described with modifications.[8] For each replicate, 0.7 x 10^8 exponentially growing S2 cells expressing wild-type or mutant MLE_{fl}-eGFP were collected by centrifugation (220 x g, 5 min), washed once with PBS and flash frozen in liquid nitrogen. Non-transfected S2 cells served as control and were treated likewise. Cell pellets were thawed on ice and resuspended in 700 µl of cold lysis buffer (20 mM HEPES-NaOH pH 7.6, 125 mM NaCl, 0.05% SDS, 0.25% sodium deoxycholate, 0.5% NP40, 1.5 mM $MgCl_2$, 0.25 mM DTT) supplemented with 0.05 U/µl RNase-free recombinant DNase I (Roche), 0.4 U/µl RNasin (Promega) and 1x Complete EDTA-free protease inhibitor (Roche). The lysate was incubated for 20 minutes on ice with 5 seconds vortexing every 5 min and cleared by centrifugation (21,000 x g, 30 min, 4°C). 1.5% of the supernatant was kept on ice as input material for RNA extraction and Western blot, respectively. Per RIP, 30 µl GFP-Trap agarose beads (proteintech) were blocked with 2% (w/v) BSA and 0.1 mg/ml yeast tRNA (Sigma-Aldrich) in lysis buffer for 1 h at 4°C to minimize non-specific interactions. The beads were washed once in lysis buffer, mixed with the remaining supernatant and incubated at 4°C for 2 hours on a rotating wheel. Beads were washed with RIP-100, RIP-250, and RIP-100 buffer for 3 minutes each at 4°C (25 mM HEPES-NaOH pH 7.6, 0.05% NP40, 3 mM $MgCl_2$ with 100 mM NaCl and 250 mM NaCl, respectively). RNA was extracted of 75% of the bead material using Proteinase K (100 µg in lysis buffer with 0.5 % SDS; 55°C for 45 minutes), phenol-chloroform extraction and ethanol precipitation in presence of 20 µg glycogen (Roche). Input material (1.5%) was treated equally. RNA pellets were resuspended in 20 µl RNase-free water. RNA input and IP material was analyzed by RT-qPCR using the SuperScript III First-Strand Synthesis System (Thermo Fisher) and Fast SYBR Green master mix (Applied Biosystems) with primers specific for roX2, RpS29, 7SK and GAPDH (key resources table). roX2 and RpS29 RNA enrichment of MLE_{fl}-eGFP and its mutants was calculated as IP/Input and normalized to unbound 7SK or GAPDH RNA. Two-tailed t-tests were used for statistical analysis. Western blot analysis of 1.5% input and 25% bead material was performed with antibodies against GFP and Lamin and visualized using the LI-COR Odyssey system and Image Studio software.

RNAi interference and immunocytochemistry

RNA interference of target genes *mle* and *gst* was essentially performed as described.[99] Double-stranded RNA fragments (dsRNA) were generated using the MEGAscript T7 transcription kit (Thermo Fisher) using PCR-amplified DNA templates (for PCR primers, see key resources table). RNA was precipitated using lithium chloride according to the manufacturer's instructions. RNA pellets were resuspended in RNase-free water and annealed to dsRNA by incubation at 85°C for 10 min followed by slowly cooling down to 20°C.

For RNAi treatment, 1.5 x 10^6 MLE_{fl}-eGFP (wild type and mutant) expressing S2 cells and non-transfected S2 cells, respectively, were seeded in 6-well plates and supplemented with 10 micrograms of GST or MLE dsRNA. Cells were incubated with dsRNA for 7 days at 26°C. RNAi efficiency was controlled by Western blot analysis of 1 x 10^6 cells with primary antibodies against MLE and Lamin and visualized using the LI-COR Odyssey system and Image Studio software.

Immunostaining with mouse anti-GFP, rat anti-MLE and rabbit anti-MSL2 primary antibodies was performed according to.[100] RNAi-treated cells were settled and fixed with PBS/3.7% paraformaldehyde for 10 min at room temperature, permeabilized with PBS/0.25% Triton X-100 for 6 min on ice and blocked with 3% BSA/PBS for 1h at room temperature. Cells on coverslips were incubated overnight at 4°C with primary antibodies diluted in PBS/5% NGS. Following two washes with PBS/0.1% Triton X-100, fluorophore-coupled secondary antibodies donkey anti-mouse-Alexa488, donkey anti-rat-Cy3 and donkey anti-rabbit-Alexa647 were diluted in PBS/5% NGS and incubated for 1 h at room temperature. DNA was counterstained with DAPI. After PBS/0.1% Triton X-100 and PBS washes, cells were mounted in VECTASHIELD (Vector Laboratories).

Fluorescence microscopy

Fluorescence images were recorded at the core facility bioimaging of the Biomedical Center with a Leica Thunder Imager 3D Live Cell TIRF based on a DMi8 stand, equipped with a Leica DFC9000 GT sCMOS camera with 2048x2048 pixels and a Leica LED5 fluorescence excitation source with individually switchable LEDs for specific excitation. DAPI, GFP/Alexa Fluor 488 (MLE-GFP), Cy3 (endogenous and recombinant MLE) and Alexa647 (MSL2) signals were recorded with a Quad-Band filter cube and an additional emission filter wheel in the emission beam path to avoid tunnel crosstalk. Image stacks of 7 planes with a step size of 1 µm were recorded with a HC PL APO 100x/1.47 oil CORR TIRF objective at a pixel size of 65 nm.

Image processing, montage assembly and quantification was done in Fiji.[90] Images shown are maximum intensity projections of z-stacks. Images were resized by a factor of 4 without interpolation followed by gaussian filtering with a radius of 2 pixels. Representative cells are shown for each cell line and RNAi condition.

Please cite this article in press as: Jagtap et al., Structural basis of RNA-induced autoregulation of the DExH-type RNA helicase maleless, Molecular Cell (2023), https://doi.org/10.1016/j.molcel.2023.10.026

Mean nuclear fluorescence quantification was done on maximum intensity projections of raw images. The macro code is available on request. Briefly: nuclei were segmented on median filtered DAPI and MSL2 sum images using Otsu dark auto thresholding upon rolling ball background subtraction. Clumped objects were separated with a Watershed algorithm. Objects (nuclei) were included in a size range of 12 – 60 sq μm and a circularity of 0.6 – 1.00. Objects at image edges were excluded. Nuclear mean intensity was measured in GFP, MLE and MSL2 channels. Cells exhibiting mean GFP signal below 500 gray values (2^8.966) were categorized as GFP^{low}, above 500 GV as GFP^{high}. Strongly overexpressing cells (mean GFP level > 2000 gray values (2^10.966)) were excluded from the analysis. The population medians of the MSL2 mean nuclear signal in MLE-GFP^{high} and MLE-GFP^{low} cells of at least three independent biological replicates were calculated, their ratios were log_2-transformed and plotted using R-Studio. The R-code is available on request. Two-tailed t-tests were used for statistical analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data in the curves shown in Figures 4E, 4G, and 6B represent mean ± standard deviation derived from experiments performed in n=3 technical replicates. Error bars shown in the affinities in Figures 4G and 6A, initial velocities shown in Figure 6B and fraction of duplex unwound shown in Figures 6C and 6D represent standard error of fitting the experimental curves obtained from experiments performed in n=3 technical replicates. Results in Figures 7B and 7D represent the mean of n=6 and n=3 biological replicates, respectively. The number of cells analyzed in each biological replicate is given in Table S3. Results in Figures 7E and S18C represent the mean ± standard deviation of n=3 biological replicates. Results in Figures 7F and S18E represent the mean ± standard deviation of n=4 biological replicates. Statistical significance of data shown in Figures 7B–7F was calculated by a two-tailed paired t-test using Excel. P-values are denoted as: ** $p<0.01$; * $p<0.05$; ns = not significant.