

## LETTER TO THE EDITOR OPEN

AI-empowered integrative structural characterization of m<sup>6</sup>A methyltransferase complex

© The Author(s) 2022

Cell Research (2022) 32:1124–1127; https://doi.org/10.1038/s41422-022-00741-8

Dear Editor,

N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) is the most abundant and prevalent internal modification in mRNA.<sup>1</sup> In mammals, m<sup>6</sup>A exerts pivotal roles in posttranscriptional regulation and its dysregulation is implicated in various diseases including cancer.<sup>2</sup> m<sup>6</sup>A is installed by a multicomponent methyltransferase complex (MTC, also known as the m<sup>6</sup>A writer complex).<sup>3,4</sup> The mammalian MTC is composed of the core m<sup>6</sup>A methyltransferase METTL3–METTL14 complex (MTC core) and several regulatory proteins including WTAP, the adaptor responsible for METTL3–METTL14 localization and proper substrate recruitment,<sup>5</sup> and VIRMA (KIAA1429), the specificity mediator that mediates preferential m<sup>6</sup>A modification at the 3′ untranslated regions (UTRs, Fig. 1a).<sup>6</sup> Dysregulation of MTC components results in the disruptions of m<sup>6</sup>A.<sup>2</sup> Compared to the other currently identified regulators (HAKAI, ZC3H13 and RBM15), WTAP and VIRMA are reported to have greater impacts on total mRNA m<sup>6</sup>A levels upon knockdown.<sup>5,6</sup> Despite the advances in understanding the roles of individual MTC components and the structural determination of MTC core,<sup>7–9</sup> the overall molecular architecture of the m<sup>6</sup>A writer holocomplex is missing. Here, we report the cryogenic electron microscopy (cryo-EM) structure of human WTAP–VIRMA (3.1 Å) in the METTL3–METTL14–WTAP–VIRMA (M–M–W–V) complex and modeled a structure of the quaternary M–M–W–V complex based on AlphaFold2 predictions and structural restraints from intermolecular chemical crosslinking mass spectrometry (CXMS).

Utilizing a baculovirus-insect cell expression system, we obtained the METTL3–METTL14 complex (Fig. 1b) and investigated the interaction between METTL3–METTL14, WTAP, and VIRMA through co-expression and purification. WTAP was observed to form a complex with the MTC core, while VIRMA can only form a stable complex with the MTC core in the presence of WTAP (Supplementary information, Fig. S1). The methyl transfer activity of the quaternary M–M–W–V complex was significantly higher than that of the MTC core, yet a quaternary M–M–W–V complex containing a truncated VIRMA<sup>381–1486</sup> with N- and C-terminal disordered regions removed showed comparable activity to the MTC core (Supplementary information, Fig. S2). In parallel, we obtained the WTAP protein alone (residues 130–241, WTAP<sup>130–241</sup>) using the *Escherichia coli* expression system and determined the crystal structure of the WTAP (2.40 Å, Supplementary information, Table S1). Two molecules of WTAP (with the assigned density of residues 150–241, WTAP<sup>150–241</sup>) formed a symmetric parallel  $\alpha$ -helical coiled-coil homodimer through hydrophobic interactions between two identical surface patches. Specifically, the side chains of residues M158, I165, L183, L204, I208, L211, V215, M218, I222, L225, L229 and L236 from the two WTAPs formed a hydrophobic interaction network (Fig. 1c; Supplementary information, Fig. S3a).

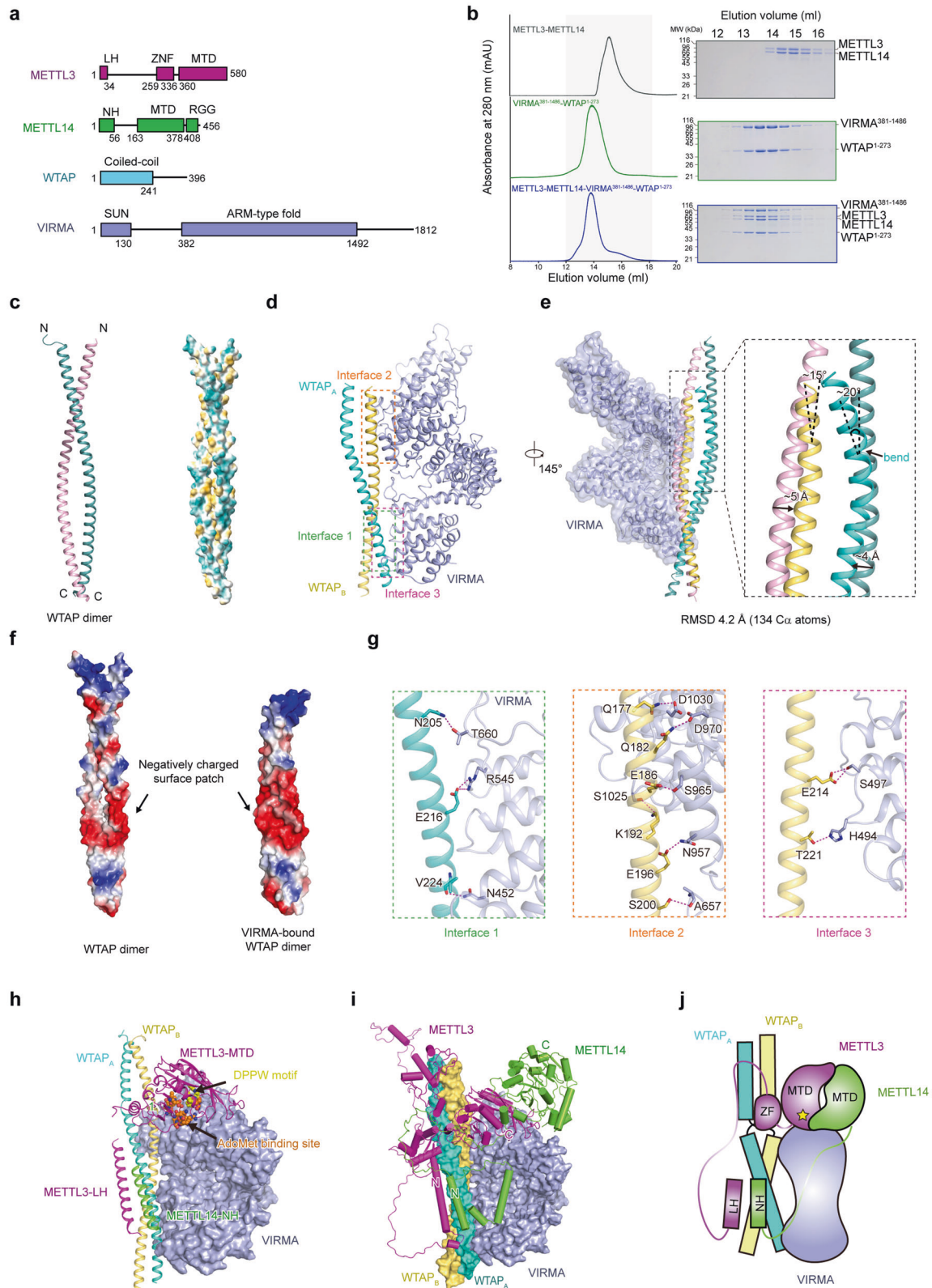
We further purified the METTL3–METTL14–WTAP<sup>1–273</sup>–VIRMA<sup>381–1486</sup> complex for gradient fixation and cryo-EM analysis

(Fig. 1b). Unexpectedly, in the cryo-EM structure, only WTAP<sup>171–237</sup> and VIRMA<sup>381–1292</sup> (with the stoichiometric ratio 2:1) and an unassigned density close to WTAP were observed (Supplementary information, Figs. S4, S5). The density for the rest of the quaternary complex was missing, suggesting high flexibility (Supplementary information, Table S2). The observed WTAP<sup>171–237</sup>–VIRMA<sup>381–1292</sup> in the quaternary complex adopted a flag-and-pole-like architecture with dimensions of  $\sim 116 \text{ \AA} \times 103 \text{ \AA} \times 80 \text{ \AA}$  (Fig. 1d). Two WTAP<sup>171–237</sup> molecules (designated WTAP<sub>A</sub> and WTAP<sub>B</sub>) formed an asymmetric homodimer (the “flag pole”) while VIRMA<sup>381–1292</sup> that contains 17 armadillo-like (ARML) repeats (designated ARML 1–17) formed a twisted  $\alpha$ -solenoid-like superhelix (the “flag”). Of the ARML repeats, each of ARML 2–4, 6, 7, 9, 12, 15, and 17 consists of three helices while the rest are two-helix units. Two long helices in between of ARML 4–5 and of ARML 7–8 mediate the turning of the VIRMA  $\alpha$ -solenoid, possibly accounting for the formation of the twisted architecture (Supplementary information, Figs. S6, S7).

The structure of the asymmetric WTAP<sup>171–237</sup> dimer in the quaternary complex is similar to the crystal structure of the symmetric WTAP<sup>150–241</sup> dimer with a root-mean-square deviation (RMSD) of 4.2 Å over 134 Ca atoms (Fig. 1e). However, compared to the WTAPs alone, VIRMA-bound WTAPs bent around the observed N-termini (around WTAP<sub>A</sub> residue 188 and WTAP<sub>B</sub> residue 185) and displayed a more compact and dense conformation with distinct intramolecular interactions (Fig. 1e). Hydrophobic interactions were found between WTAP<sub>A</sub> and WTAP<sub>B</sub> residues 180–236 through the side-chain interactions of WTAP<sub>A</sub>–WTAP<sub>B</sub> residue pairs L183<sub>A</sub>–I180<sub>B</sub> (L183 from WTAP<sub>A</sub> and I180 from WTAP<sub>B</sub>), L187<sub>A</sub>–L183<sub>B</sub>, I208<sub>A</sub>–L204<sub>B</sub>, L211<sub>A</sub>–I208<sub>B</sub>, V215<sub>A</sub>–L211<sub>B</sub>, M218<sub>A</sub>–V215<sub>B</sub>, I222<sub>A</sub>–M218<sub>B</sub>, L225<sub>A</sub>–I222<sub>B</sub>, L229<sub>A</sub>–L225<sub>B</sub>, and L236<sub>A</sub>–L229<sub>B</sub>. WTAP K191<sub>A</sub> and Q201<sub>A</sub> side chains formed hydrogen bonds with Q190<sub>B</sub> and Q201<sub>B</sub> side chains, respectively. Hydrogen bond interactions were also observed between WTAP Q201<sub>A</sub> side chain and K198<sub>B</sub> main chain and between N205<sub>A</sub> main chain and Q201<sub>B</sub> side chain. These together contributed to the formation of the compact VIRMA-bound WTAP dimer that harbors a contiguous negatively charged surface patch involved in the interaction with VIRMA (Fig. 1e, f). Some of these residues involved in WTAP<sub>A</sub>–WTAP<sub>B</sub> interactions (around residues 180–236) are also conserved in *Danio rerio* and *Drosophila melanogaster* WTAPs (Supplementary information, Fig. S3c).

Three intermolecular interaction interfaces, designated Interfaces 1, 2, and 3, could be observed between the WTAP<sup>171–237</sup> dimer and VIRMA<sup>381–1292</sup> with buried areas of  $\sim 773.9 \text{ \AA}^2$ ,  $\sim 1155.6 \text{ \AA}^2$ , and  $\sim 101 \text{ \AA}^2$ , respectively (Fig. 1d). Interface 1 is located in between WTAP<sub>A</sub> (residues 205–224) and VIRMA (residues 452–660). Specifically, the side chains of WTAP N205<sub>A</sub> and E216<sub>A</sub> and VIRMA T660 and R545 formed hydrogen bonds,

Received: 20 July 2022 Accepted: 12 October 2022  
Published online: 10 November 2022



and E216<sub>A</sub> and VIRMA R545 formed a salt bridge. V244<sub>A</sub> main chain and VIRMA N452 side chain formed a hydrogen bond. Interfaces 2 and 3 are in between of WTAP<sub>B</sub> (residues 177–200, 214–221) and VIRMA (residues 657–1030, 494–497). The side chains of WTAP Q177<sub>B</sub>, Q182<sub>B</sub>, E186<sub>B</sub>, K192<sub>B</sub>, E196<sub>B</sub>, E214<sub>B</sub>, and

T221<sub>B</sub> formed hydrogen bonds with the side chains of VIRMA D1030, D970, S965, S1025, N957, S497, and H494, respectively. WTAP S200<sub>B</sub> and E203<sub>B</sub> side chains formed hydrogen bonds with VIRMA A657 and M656 main chains, respectively (Fig. 1g; Supplementary information, Fig. S8a, b).

**Fig. 1 Structure of the human m<sup>6</sup>A writer complex.** **a** Schematic diagram of the domain information of METTL3, METTL14, WTAP and VIRMA. LH leader helix, ZFD zinc finger domain, MTD MTase domain, NH N-terminus helix. **b** Gel filtration analysis of the human METTL3–METTL14–WTAP<sup>1–273</sup>–VIRMA<sup>381–1486</sup> complex. **c** The crystal structure of WTAP alone. Two molecules of WTAP form a symmetric parallel  $\alpha$ -helical coiled-coil, with the two chains colored light pink and light teal, respectively. **d** The cryo-EM structure of WTAP–VIRMA complex in the METTL3–METTL14–WTAP<sup>1–273</sup>–VIRMA<sup>381–1486</sup> quaternary complex. Three interaction interfaces can be observed between WTAPs and VIRMA. WTAP<sub>A</sub>, cyan; WTAP<sub>B</sub>, orange. **e** Comparison of the crystal structure of WTAP alone and the cryo-EM structure of VIRMA-bound WTAPs. VIRMA-bound WTAPs bend around the observed N-termini. **f** Electrostatic surfaces of WTAP alone (PDB: 7YFJ) and VIRMA-bound WTAPs (PDB: 7YG4). **g** Details of the observed interaction interfaces between WTAP and VIRMA. **h, i** Structural model of the METTL3–METTL14<sup>1–399</sup>–WTAP<sup>148–237</sup>–VIRMA<sup>342–1292</sup> complex based on CXMS measurement and AI prediction. METTL3 and METTL14 are shown in green and magenta, respectively. METTL3 LH is docked close to WTAP, and the MTC core catalytic center is docked close to VIRMA, with the DPPW motif and AdoMet binding site positioned towards a cleft between WTAP and VIRMA. **j** Schematic diagram of the modeled METTL3–METTL14<sup>1–399</sup>–WTAP<sup>148–237</sup>–VIRMA<sup>342–1292</sup> complex.

We further investigated the importance of the residues involved in the formation of Interfaces 1–3 via alanine substitutions and subsequent co-expression and purification assays (Supplementary information, Fig. S8). Specifically, Flag-tagged WTAP<sup>1–273</sup> and His-tagged VIRMA<sup>381–1486</sup> were co-expressed in the human embryonic kidney (HEK) 293F cells. The protein mixtures from the whole-cell extracts were subjected to western blot analysis against the tagged proteins and purification via anti-Flag beads. The single WTAP mutations (Q177A, Q182A, Q192A, E196A, E203A, N205A, or E216A) had little influence on WTAP–VIRMA<sup>381–1486</sup> interaction, whereas the multi-alanine substitution mutant WTAP<sup>Q177A/Q182A/K192A/E196A/E203A</sup> lost VIRMA<sup>381–1486</sup>-binding capacity (Supplementary information, Fig. S8c). The VIRMA<sup>381–1486</sup> mutants containing individual N452A, T660A, N957A, or S1025A mutations retained binding to WTAP<sup>1–273</sup>, while the single alanine substitution of VIRMA<sup>381–1292</sup> residue R545 or D970 abolished its interaction with WTAP<sup>1–273</sup> (Supplementary information, Fig. S8d).

Despite that gel filtration analysis clearly indicates the presence of METTL3 and METTL14 in the METTL3–METTL14–WTAP<sup>1–273</sup>–VIRMA<sup>381–1486</sup> quaternary complex, the cryo-EM density for the MTC core is largely missing in the cryo-EM structure, indicating that the MTC core is highly dynamic. This is also supported by previous reports that both METTL3 and METTL14 contain highly flexible intrinsically disordered regions (IDRs).<sup>10</sup> Thus, we further conducted in vitro CXMS and AI-based prediction to model the structure of the quaternary M–M–W–V complex based on the solved structures. In total, we identified 227 (61 unique inter-subunit and 166 intra-subunit) bis(sulfosuccinimidyl) suberate (BS3)-crosslinked residue pairs with the false discovery rate (FDR) < 5% (Supplementary information, Fig. S9a). Notably, crosslinked pairs between WTAP coiled-coil and METTL3 N-terminal region (METTL3 residues 1–251, including METTL3–WTAP<sub>A</sub> crosslinked pairs K13–K192, K132–K192, K241–K155, and K241–K160), VIRMA and the METTL14 N-terminal helix (NH, including METTL14–VIRMA crosslinked pair K38–K399), and VIRMA and MTC core catalytic center (including METTL3–VIRMA crosslinked pairs K576–K899, K576–K978, and K578–K899; METTL14–VIRMA crosslinked pair K148–K887) were identified (Supplementary information, Fig. S9b). No crosslinked pairs were identified between the METTL3 ZFD and METTL14 RGG domains with WTAP or VIRMA. The final structural model of the quaternary complex can account for all intermolecular crosslinks between METTL3/METTL14 core domains and the two regulatory subunits, except for two crosslinks involving flexible loop residues. In the highest-ranking model, the METTL3 N-terminal leader helix (LH) region (residues 1–34) was docked in close proximity to WTAP, accounting for the unassigned cryo-EM density (Fig. 1h, i; Supplementary information, Fig. S5c). This is in line with a previous report that the METTL3 LH is crucial for WTAP–METTL3 interaction.<sup>11</sup> The rest of the METTL3 N-terminal region was also predicted to be close to WTAP. The N-terminus of METTL14 was docked close to VIRMA and potentially interacts with WTAP. Although the MTC core is likely dynamically positioned, its catalytic center was docked next to VIRMA with the catalytic center facing VIRMA ARML 9–10 (Fig. 1h, i). This model thus represents a possible resting state of the quaternary complex.

WTAP and VIRMA are important subunits of the m<sup>6</sup>A writer complex.<sup>1,3</sup> We determined the structure of WTAP–VIRMA in the M–M–W–V complex at atomic resolution and modeled the quaternary complex structure using an AI-empowered integrative approach (Fig. 1j). Consistent with the previously reported adaptor role of WTAP,<sup>3</sup> the observed WTAP dimer likely serves as a linker connecting VIRMA (and/or other MTC regulatory subunits) with the MTC core. VIRMA mediates preferential m<sup>6</sup>A mRNA methylation in 3' UTR, but the exact mechanism is unknown.<sup>6</sup> Like other ARML-containing proteins,<sup>12</sup> the observed VIRMA superhelix contains extensive solvent-accessible surfaces that can accommodate RNA substrate<sup>6</sup> and other regulator proteins.<sup>6,13</sup> Moreover, in the model the VIRMA ARML 9–10 region (containing several positively charged residues) is close to the catalytic center of METTL3, suggesting a potential VIRMA–RNA interaction. During our manuscript revision, Su and colleagues also reported the structure of the regulatory subunit m<sup>6</sup>A-METTL-associated complex (MACOM) containing WTAP–VIRMA–ZC3H13–HAKAI.<sup>14</sup> Our WTAP–VIRMA structure in the M–M–W–V complex is almost identical to the counterpart in the MACOM. Moreover, our AI-empowered integrative structural model indicates the interaction between the N-terminal helix of METTL3 and WTAP coiled-coil, and proximity between the MTC catalytic core and a VIRMA positively charged surface. As such, VIRMA N- and C-terminal disordered regions likely interact with the MTC core or substrate RNA, giving rise to the higher methyltransferase activity for the full-length VIRMA-containing quaternary complex. The two structural models provide a framework for future study on the molecular architecture and catalytic process of the m<sup>6</sup>A writer holocomplex, and offer new insights into potential therapeutic manipulation of m<sup>6</sup>A modification through the design of small molecules or peptides targeting WTAP–VIRMA interaction and modulating m<sup>6</sup>A catalytic specificity/activity.<sup>15</sup> We anticipate that future structural and biochemical characterization of the complete m<sup>6</sup>A writer complex and its substrate-bound form will unveil the specificity and regulatory details underlying mRNA m<sup>6</sup>A methylation.

Xuhui Yan<sup>1,3</sup>, Kai Pei<sup>2,3</sup>, Zeyuan Guan<sup>1,3</sup>, Feiqing Liu<sup>1</sup>, Junjun Yan<sup>1</sup>, Xiaohuan Jin<sup>1</sup>, Qiang Wang<sup>1</sup>, Mengjun Hou<sup>1</sup>, Chun Tang<sup>1,2</sup>✉ and Ping Yin<sup>1</sup>✉

<sup>1</sup>National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, Hubei, China. <sup>2</sup>Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, PKU-Tsinghua Center for Life Sciences, Center for Quantitative Biology, Peking University, Beijing, China. <sup>3</sup>These authors contributed equally: Xuhui Yan, Kai Pei, Zeyuan Guan. ✉email: Tang\_Chun@pku.edu.cn; yinping@mail.hzau.edu.cn

#### DATA AVAILABILITY

Atomic coordinates of the WTAP<sup>130–241</sup> has been deposited in the Protein Data Bank (PDB) under accession number 7YFJ. The cryo-EM density map for the human METTL3–METTL14–WTAP<sup>1–273</sup>–VIRMA<sup>381–1486</sup> complex has been deposited in EM

Database under the accession code EMD-33807. The corresponding atomic coordinates have been deposited in the PDB under accession code 7YG4. The CXMS data have been deposited to the ProteomeXchange Consortium PRIDE: PXD036144.

## REFERENCES

1. Yang, Y., Hsu, P. J., Chen, Y. S. & Yang, Y. G. *Cell Res.* **28**, 616–624 (2018).
2. Huang, H., Weng, H. & Chen, J. *Cancer Cell* **37**, 270–288 (2020).
3. Zaccara, S., Ries, R. J. & Jaffrey, S. R. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
4. Davalos, V., Blanco, S. & Esteller, M. *Cell* **174**, 498–498.e1 (2018).
5. Ping, X. L. et al. *Cell Res.* **24**, 177–189 (2014).
6. Yue, Y. et al. *Cell Discov.* **4**, 10 (2018).
7. Wang, X. et al. *Nature* **534**, 575–578 (2016).
8. Wang, P., Doxtader, K. A. & Nam, Y. *Mol. Cell* **63**, 306–317 (2016).
9. Śledź, P. & Jinek, M. *Elife* **5**, e18434 (2016).
10. Wang, Z. H. et al. *EMBO J.* **40**, e106309 (2021).
11. Schöller, E. et al. *RNA* **24**, 499–512 (2018).
12. Yan, Z. et al. *Nature* **517**, 50–55 (2015).
13. Bawankar, P. et al. *Nat. Commun.* **12**, 3778 (2021).
14. Su, S. et al. *Cell Res.* <https://doi.org/10.1038/s41422-022-00725-8> (2022).
15. Yankova, E. et al. *Nature* **593**, 597–601 (2021).

## ACKNOWLEDGEMENTS

We thank the Cryo-EM Facility of Hubei University for providing technical support during EM image acquisition and the Center for Protein Research and Public Laboratory of Electron Microscopy, Huazhong Agricultural University, for technical support, J. Cao for assistance during cryo-EM image acquisition. We thank National Center for Protein Sciences at Peking University in Beijing, China, for assistance with mass spectrometry experiments. C. Wang has improved this manuscript. This work was supported by funds from the National Key R&D Program of China (2018YFA0507700), the National Natural Science Foundation of China (31870753 and 91753132) and the Foundation of Hubei Hongshan Laboratory (2021hszd011).

## AUTHOR CONTRIBUTIONS

P.Y. and C.T. conceived the project. X.Y. and P.Y. designed all the experiments. X.Y., F.L., K.P., J.Y., X.J., and M.H. performed the experiments. X.Y. and Q.W. collected the

EM data. Z.G. determined the crystal and cryo-EM structures. K.P. and C.T. performed mass spectrometry analysis. All authors analyzed the data and contributed to manuscript preparation. X.Y., Q.W., C.T. and P.Y. wrote the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41422-022-00741-8>.

**Correspondence** and requests for materials should be addressed to Chun Tang or Ping Yin.

**Reprints and permission information** is available at <http://www.nature.com/reprints>



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022