

Supplementary information for

**Structural insight into hierarchical DNMT3A autoinhibition and its dysregulation in disease**

Jiuwei Lu<sup>1</sup>, Emily Vig<sup>2,#</sup>, Jianbin Chen<sup>1,2,#</sup>, Kristjan H. Gretarsson<sup>3,#</sup>, Nelli Khudaverdyan<sup>1,2</sup>, Zengyu Shao<sup>1,2</sup>, Chao Lu<sup>3</sup>, Chia-en A. Chang<sup>2,4</sup>, Jikui Song<sup>1,2,\*</sup>

<sup>1</sup>Department of Biochemistry, University of California, Riverside, CA 92521, USA

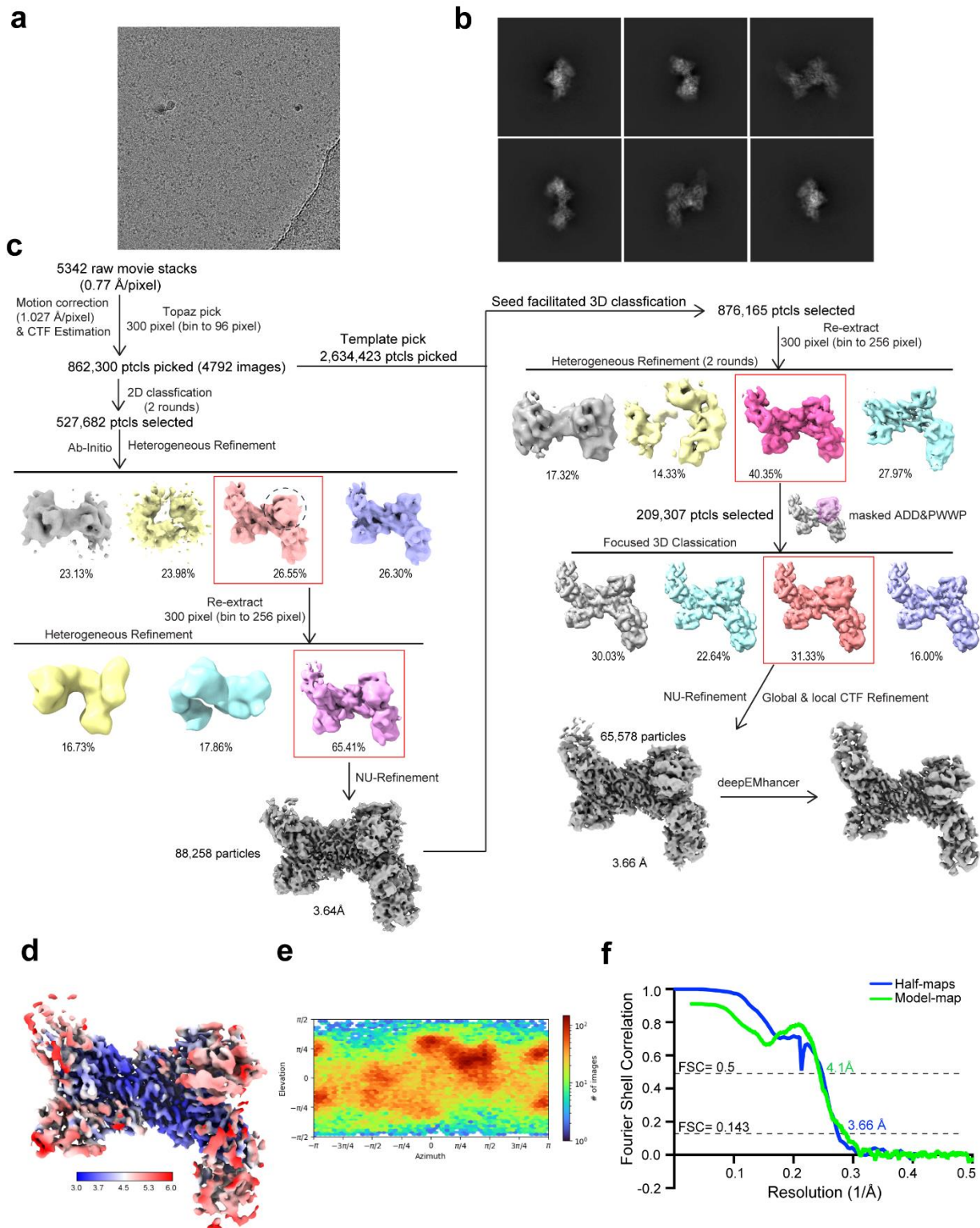
<sup>2</sup>Biochemistry and Molecular Biology Graduate Program, University of California, Riverside, CA 92521, USA

<sup>3</sup>Department of Genetics and Development and Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New York, NY 10032, USA.

<sup>4</sup>Department of Chemistry, University of California, Riverside, CA 92521, USA

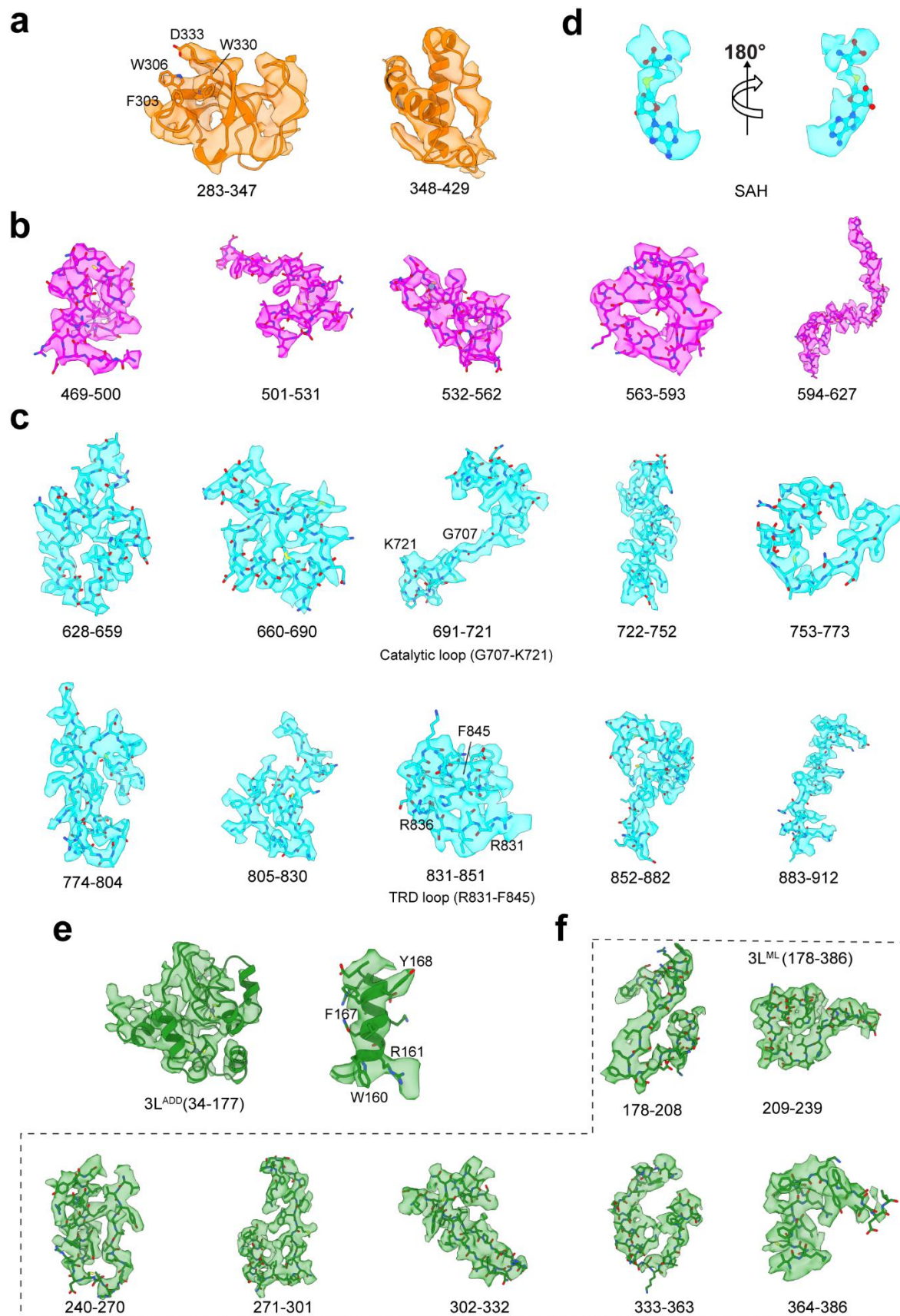
#These authors contributed to the project equally

\*Correspondence: [jikui.song@ucr.edu](mailto:jikui.song@ucr.edu)



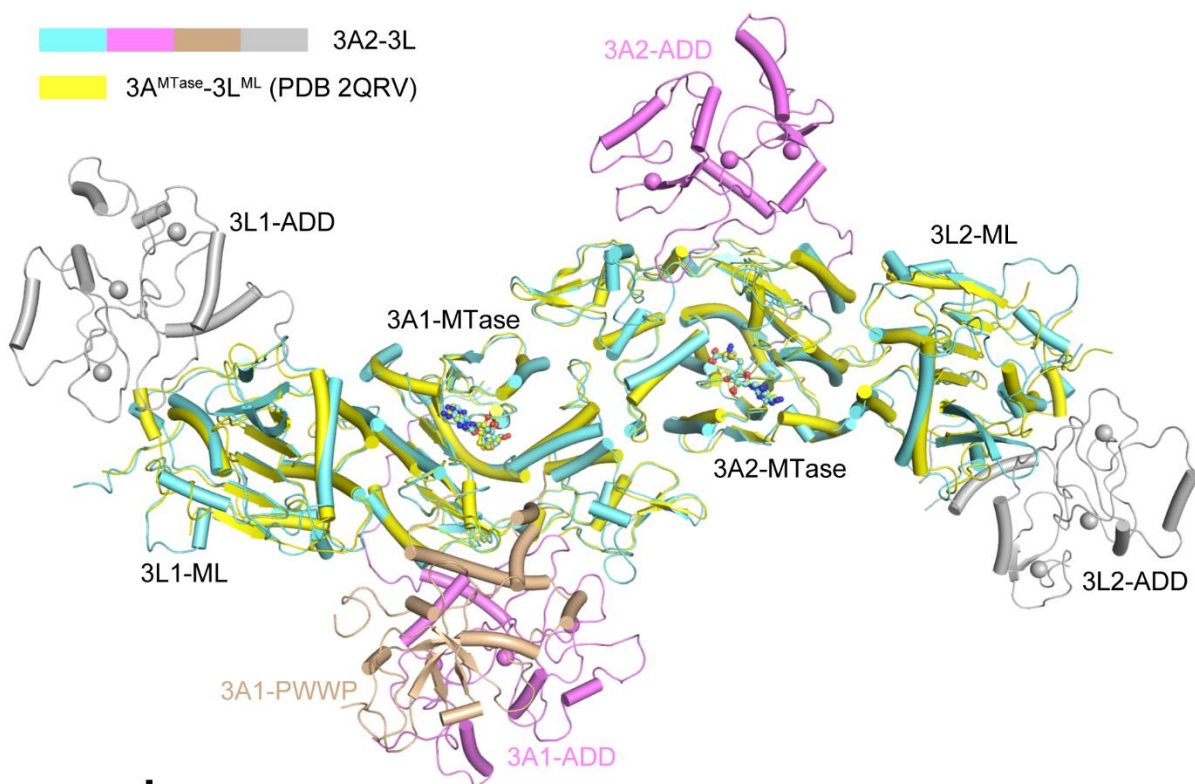
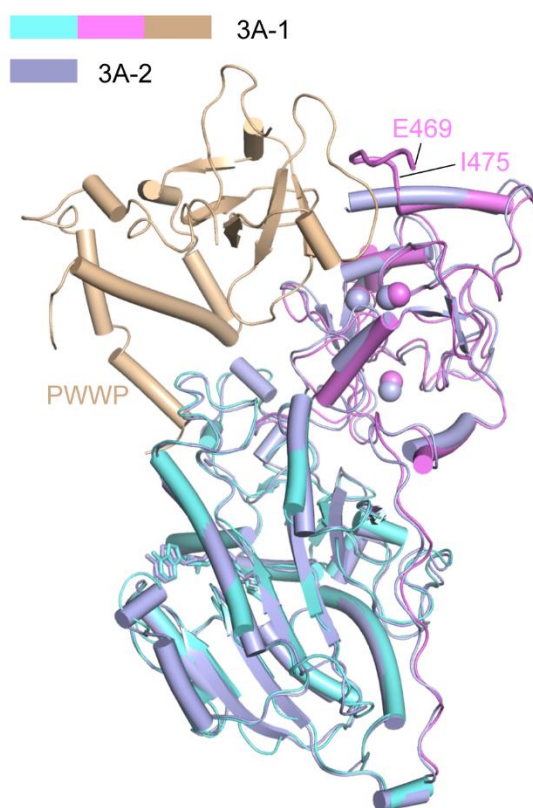
**Supplementary Figure 1. Cryo-EM structure of the DNMT3A2-DNMT3L complex.** **a** A representative micrograph of the DNMT3A2-DNMT3L complex by Cryo-EM. **b** Representative 2D classes of the DNMT3A2-DNMT3L complex from cryoSPARC. **c**

Workflow for data processing. The particle groups used for subsequent data processing were indicated by red squares. For instance, after initial heterogeneous refinement, the particles associated with the density with the feature of the PWWP domain (circled) were selected for further classification. **d-f** Local resolution (d), Angular distribution of final particles used for construction (e), and FSC curve for half maps and model-map for the structure of the DNMT3A2-DNMT3L complex (f).



**Supplementary Figure 2. Cryo-EM density maps of the DNMT3A2-DNMT3L complex.**

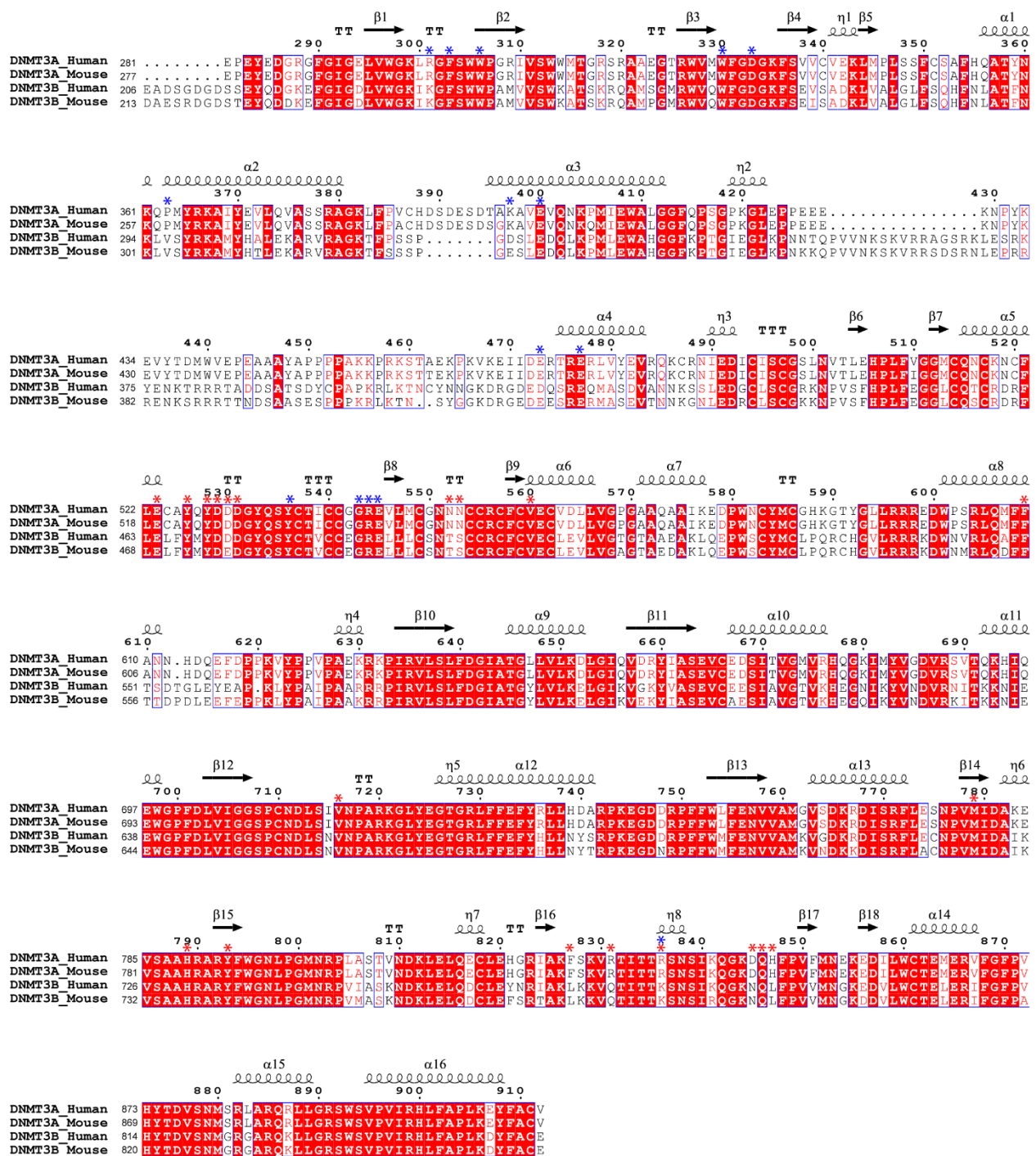
**a-d** Density maps for representative regions of the PWWP domain (a), ADD domain (b), MTase domain (c), and bound SAH (two opposite views) (d) of DNMT3A. **e,f** Density maps for representative structural elements of the ADD domain (e) and ML domain (f) of DNMT3L.

**a****b**

**Supplementary Figure 3. Structural overview of the DNMT3A2-DNMT3L complex. a**

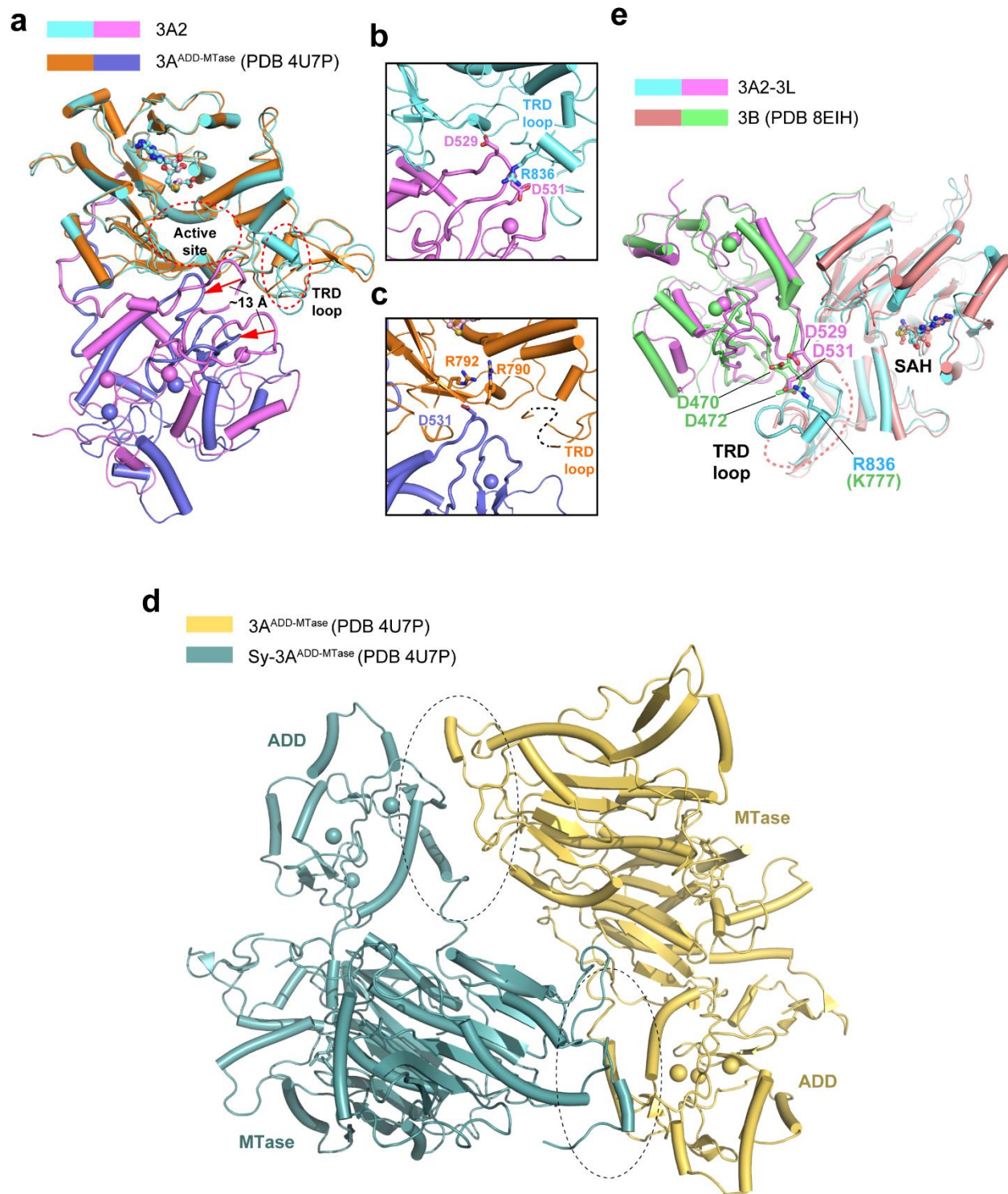
Structural overlay of the DNMT3A2-DNMT3L complex with the previously reported DNMT3A<sup>MTase</sup>-DNMT3L<sup>ML</sup> (PDB 2QRV), with the domains of individual subunits labeled.

**b** Structural overlay of the 3A-1 and 3A-2 subunits in the DNMT3A2-DNMT3L complex.



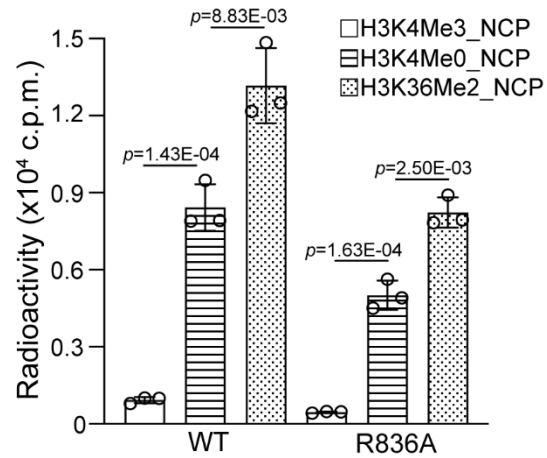
**Supplementary Figure 4. Structure-based sequence alignment of human and mouse DNMT3A and DNMT3B.** Identical or similar residues are boxed and colored in red. Completely conserved residues are shaded in red. Identical residues are colored white in red background and similar residues are colored in red. The residues involved in the PWWP-mediated interaction with the ADD and MTase domains and the binary ADD-

MTase interaction are marked by blue and red asterisks, respectively, on top. The secondary structures corresponding to DNMT3A are also indicated on top.

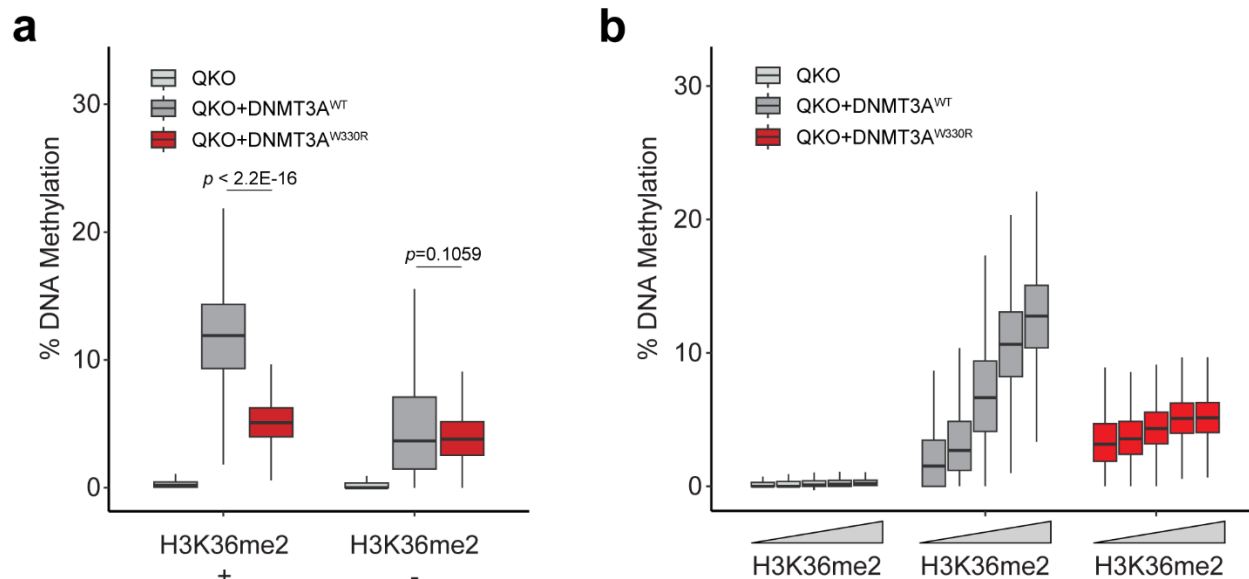


**Supplementary Figure 5. Structural comparison between the cryo-EM structure of DNMT3A2-DNMT3L and the crystal structure of DNMT3A<sup>ADD-MTase</sup>-DNMT3L<sup>ML</sup>.** a Structural overlay of the DNMT3A subunit in the DNMT3A2-DNMT3L complex and the

DNMT3A<sup>ADD-MTase</sup>-DNMT3L<sup>ML</sup> complex (PDB 4U7P), with a ~13-Å shift of the ADD positioning indicated by red arrows. **b,c** Close-up view of the ADD-MTase interaction in the DNMT3A2-DNMT3L complex (b) and the DNMT3A<sup>ADD-MTase</sup>-DNMT3L<sup>ML</sup> complex (PDB 4U7P) (c). Note that the ADD residue D531 interacts with TRD-loop residue R836 in (b) but residues R790 and R792 in (c). **d** Crystal packing analysis of the DNMT3A<sup>ADD-MTase</sup>-DNMT3L<sup>ML</sup> complex (3A<sup>ADD-MTase</sup>) (PDB 4U7P) and the symmetry-related complex (Sy-3A<sup>ADD-MTase</sup>), revealing that the ADD domain and its downstream linker (marked by dotted circle) are involved in crystal packing. **e** Structural overlay of between the ADD-MTase region of the 3A-1 subunit and the corresponding region in DNMT3B homotetramer (PDB 8EIH). Note that the ADD domains are positioned similarly between the two proteins. The TRD loop of DNMT3B is shown as dashed line due to disorder. The interacting residues, DNMT3A D529, D531 and R836, and the corresponding sites in DNMT3B shown in stick representation. DNMT3B K777, corresponding to DNMT3A R836, is not shown due to structural disorder.



**Supplementary Figure 6.** *In vitro* DNA methylation assay of DNMT3A2-DNMT3L, WT or R836A mutant, on the nucleosome substrates with various histone modifications. The statistical analysis used two-tailed Student's t test. Data are mean  $\pm$  s.d. (n = 3 biological repeats). Source data are provided as a Source Data file.



**Supplementary Figure 7. Genomic DNA methylation analysis of QKO mESCs with WT and W330R DNMT3A1.** **a** Box plots showing the levels of DNA methylation in CGI-excluded, H3K36me2-positive (left) or negative (right) regions in QKO reconstituted with WT or W330R DNMT3A1. Box-and whisker plots depict 25-75% in the box, whiskers are 1.5 times the interquartile range (IQR), and median is indicated. Wilcoxon rank-sum two-sided test is used for statistical analysis ( $n = 2$  biological replicates). **b** Percent of DNA methylation in WT or W330R DNMT3A1 QKO mESCs grouped by H3K36me2 CUT&RUN enrichments. Box-and whisker plots depict 25-75% in the box, whiskers are 1.5 times the interquartile range (IQR), and median is indicated ( $n = 2$  biological replicates).

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

| States                                              | DNMT3A/3L<br>(EMD-71814, PDB 9PRW)                                                       |
|-----------------------------------------------------|------------------------------------------------------------------------------------------|
| <b>Data collection and processing</b>               |                                                                                          |
| Microscope                                          | Titan Krios                                                                              |
| Camera                                              | TFS Falcon IV                                                                            |
| Magnification                                       | 130,000                                                                                  |
| Voltage (kV)                                        | 300                                                                                      |
| Defocus range ( $\mu\text{m}$ )                     | -0.8 ~ -2.5                                                                              |
| Exposure time (s)                                   | 7                                                                                        |
| Dose rate( $e^-/\text{px/s}$ )                      | 7 $e^-/\text{px/s}$                                                                      |
| Number of frames                                    | 29                                                                                       |
| Pixel size ( $\text{\AA}$ )                         | 0.7675                                                                                   |
| Micrographs (no.)                                   | 5342                                                                                     |
| Initial particles (no.)                             | 2,634,423                                                                                |
| Symmetry imposed                                    | <i>C1</i>                                                                                |
| Final particles (no.)                               | 65,578                                                                                   |
| Map resolution ( $\text{\AA}$ )                     | 3.66                                                                                     |
| FSC threshold                                       | 0.143                                                                                    |
| <b>Refinement</b>                                   |                                                                                          |
| Initial model used                                  | 4U7P,3LLR, 2PV0                                                                          |
| Model resolution ( $\text{\AA}$ )                   | 4.1                                                                                      |
| FSC threshold                                       | 0.5                                                                                      |
| Map sharpening <i>B</i> factor ( $\text{\AA}^2$ )   | deepEMhanceer (visualization)<br>uniform B-factor of -116.28 $\text{\AA}^2$ (refinement) |
| <b>Model composition</b>                            |                                                                                          |
| Non-hydrogen atoms                                  | 13,546                                                                                   |
| Protein residues                                    | 1703                                                                                     |
| Ion (zinc)                                          | 12                                                                                       |
| SAH                                                 | 2                                                                                        |
| <b><i>B</i> factors (<math>\text{\AA}^2</math>)</b> |                                                                                          |
| Protein                                             | 123.4                                                                                    |
| Ion (zinc)                                          | 205.1                                                                                    |
| SAH                                                 | 47.8                                                                                     |
| <b>R.m.s. deviations</b>                            |                                                                                          |
| Bond lengths ( $\text{\AA}$ )                       | 0.003                                                                                    |
| Bond angles ( $^\circ$ )                            | 0.661                                                                                    |
| <b>Validation</b>                                   |                                                                                          |
| MolProbity score                                    | 1.9                                                                                      |
| Clashscore                                          | 9.98                                                                                     |
| Poor rotamers (%)                                   | 0.14                                                                                     |
| <b>Ramachandran plot</b>                            |                                                                                          |
| Favored (%)                                         | 94.38                                                                                    |
| Allowed (%)                                         | 5.62                                                                                     |
| Disallowed (%)                                      | 0                                                                                        |

**Supplementary Table 2. List of primers used in this study**

| <b>Name</b>       | <b>Sequence</b>                           |
|-------------------|-------------------------------------------|
| 3A_sy_281F_VP13   | GTACATCCAAGGATCCGAACCGGAATATGAAGATGGT     |
| 3A_sy_912R_VP13   | GTCCTACAGGCGCGCCTTAAACGCAGGCAAAATATTCTTT  |
| 3L_1F_Vp13        | GTACATCCAAGGATCCATGGCGGCCATCCCAGCCCTGGA   |
| 3L_AscII_386R     | GTCCTACAGGCGCGCCTTAGGGTTGGTGTGAGGAGACT    |
| 3A_281F_pRSF_sy   | ACAGATTGGTGGATCCGAACCGGAATATGAAGATGGT     |
| 3A_427R_pRSF_sy   | CTTTACCAGACTCGAGTTATTCTTCAGGCGGTTCCAGACC  |
| 3A_476F_pRSF_sy   | ACAGATTGGTGGATCCCGTGAACGTCTGGTTTATGAA     |
| 3A_614R_pRSF_sy   | CTTTACCAGACTCGAGTTAATCGTGGTTATTTGCAAAAAA  |
| hDNMT3A_912Not1   | TAAGCATTATGCGGCCGCTTAAACGCAGGCAAAATATTC   |
| 3A_W330A_SyF      | TGGGTTATGGCGTTTGGTGTATGGTAAATTTAGC        |
| 3A_W330A_SyR      | CATCACCAAACGCCATAACCCAACGGGTGCCTT         |
| 3A_W330R_SyF      | TGGGTTATGCGTTTTGGTGTATGGTAAATTTAGC        |
| 3A_W330R_SyR      | CCATCACCAAACGCATAACCCAACGGGTGCCT          |
| 3A_D333N_syF      | TGGTTTGGTAACGGTAAATTTAGCGTTGTGTGC         |
| 3A_D333N_syR      | TAAATTTACCGTTACCAAACCACATAACCCAAC         |
| 3A_R544A_syF      | TGTGGTGGTGCTGAAGTTCTGATGTGTGGTAAT         |
| 3A_R544A_syR      | TCAGAACTTCAGCACCACCACAACAAATGGTGC         |
| 3A_E545G_syF      | GGTGGTCGTGGCGTTCTGATGTGTGGTAATAAC         |
| 3A_E545G_syR      | ACATCAGAACGCCACGACCACCACAACAAATGG         |
| 3A_R836A_syF      | ATTACCACCGCGAGCAATAGCATTAACAGGGT          |
| 3A_R836A_syR      | TGCTATTGCTCGCGGTGGTAATGGTACGAACTT         |
| 3A_D529D531AA_syF | TATCAGTATGCGGATGCCGGTTATCAGAGCTAT         |
| 3A_D529D531AA_syR | TCTGATAACCGGCATCCGCATACTGATAGGCGC         |
| DNMT3L_pRSF_178F  | AAGGAGATATACATATGATGTTTCGAAACCGTGCCTGTGT  |
| DNMT3L_pRSF_386R  | CTTTACCAGACTCGAGTTATAAAGAGGAAGTGAGTTCTGTT |

**Supplementary Table 3. Setup of molecular dynamics simulation**

| System                           | Water box (12 Å) | Water molecules | Na <sup>+</sup> Ions added | Total Number of Atoms |
|----------------------------------|------------------|-----------------|----------------------------|-----------------------|
| DNMT3A <sup>PWWP-ADD-MTase</sup> | 12               | 91759           | 25                         | 302350                |
| DNMT3A <sup>ADD-MTase</sup>      | 12               | 92490           | 28                         | 302153                |
| DNMT3A <sup>ADD-MTase-H3</sup>   | 12               | 92430           | 22                         | 302243                |