Structural and mechanistic insights into a mesophilic prokaryotic Argonaute

Xin Tao[1,†], Hui Ding[1,†], Shaowen Wu[2,†], Fei Wang[1], Hu Xu[1], Jie Li[1], Chao Zhai[1], Shunshun Li[1], Kai Chen[1], Shan Wu[1,*], Yang Liu[1,*], Lixin Ma[1,*]

[1] State Key Laboratory of Biocatalysis and Enzyme Engineering, Hubei Key Laboratory of Industrial Biotechnology, School of Life Sciences, Hubei University, Wuhan, Hubei, 430062, China

[2] State Key Laboratory of Swine and Poultry Breeding Industry, Agro-biological Gene Research Center, Guangdong Academy of Agricultural Sciences, Guangzhou, 510640, China

* To whom correspondence should be addressed. Tel: +86 27 50865628; Fax: +86 27 88666349; Email: malixing@hubu.edu.cn

Correspondence may also be addressed to Yang Liu. Email: 20220098@hubu.edu.cn

Correspondence may also be addressed to Shan Wu. Email: wushan91@hubu.edu.cn

[†]The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

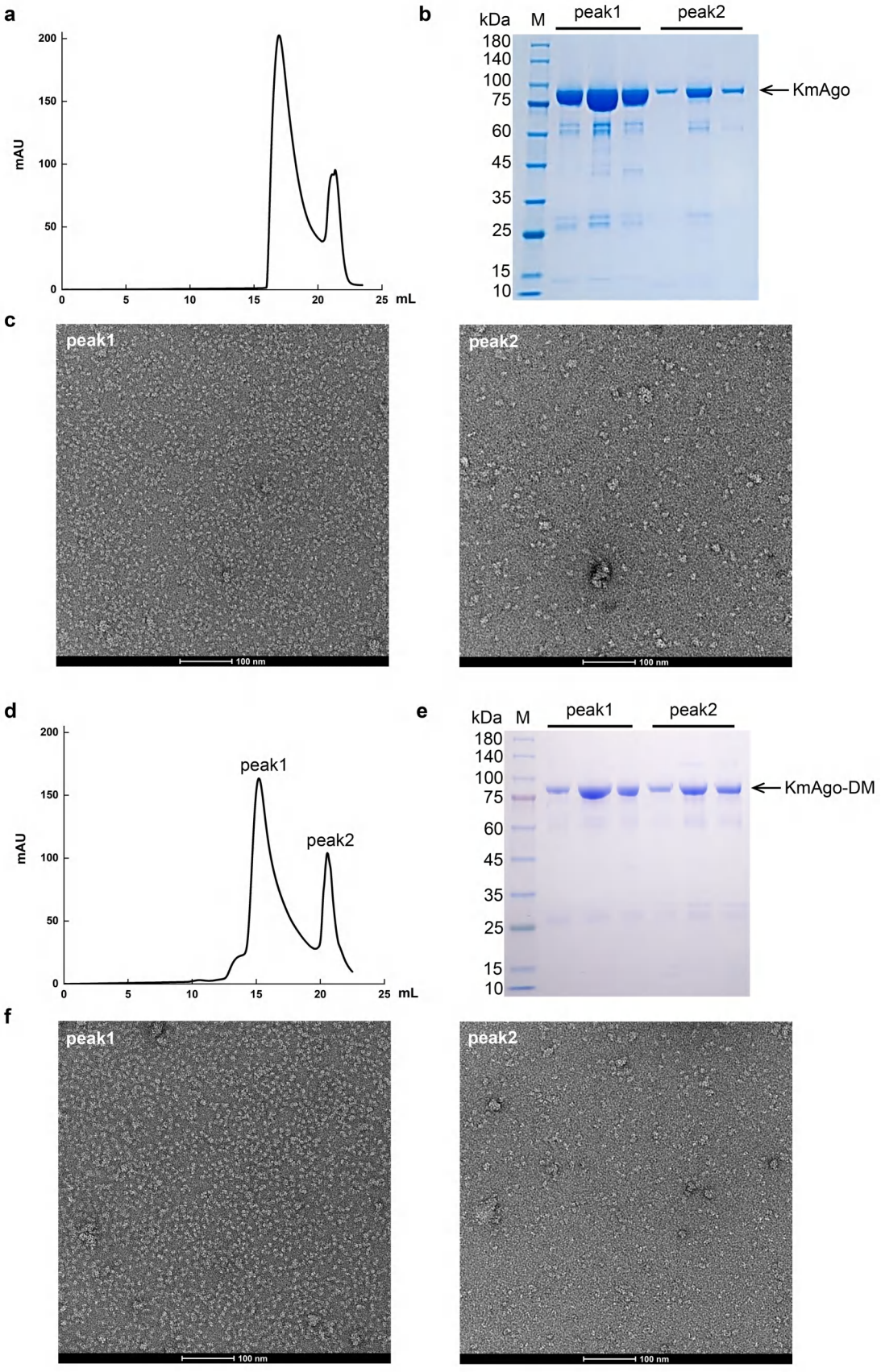
Supplementary Information

PAZ-domain

MID-domain

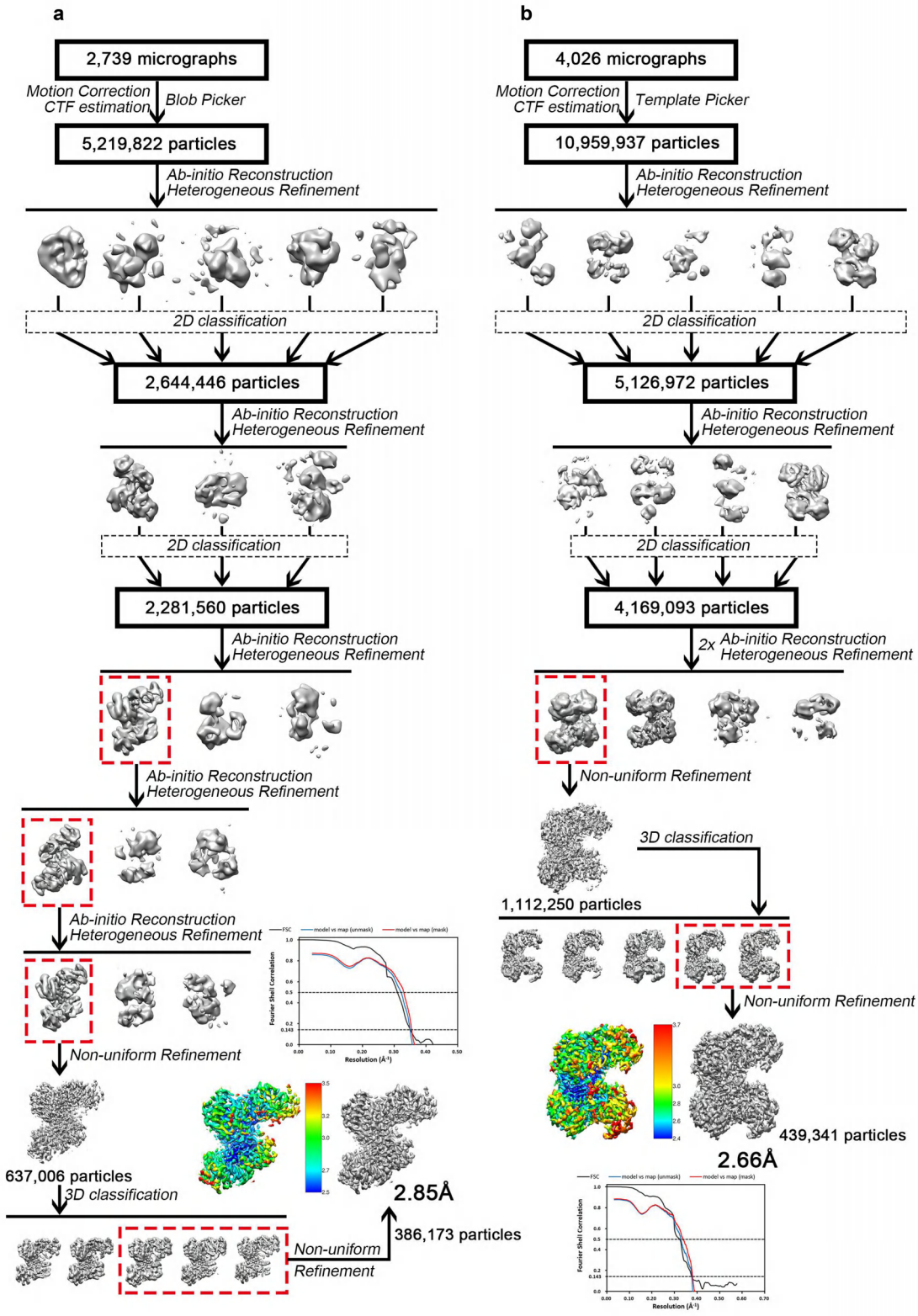
Supplementary Figure S1. Multiple sequence alignment of Ago proteins. Alignment of PAZ domain with the conserved residues are shaded and highlighted with arrows. The residues responsible for 3′-end binding in the binary complex are indicated by the orange box. Alignment of MID domains with the residues of the motif for 5′-phosphate coordination of the guide strand (X-K-Q-K) and Mn^{2+} ion coordination is shaded in blue

and highlighted with arrows. Alignment of PIWI domain with the conserved residues are shaded and highlighted with arrows. The catalytic residues are indicated by the red box.

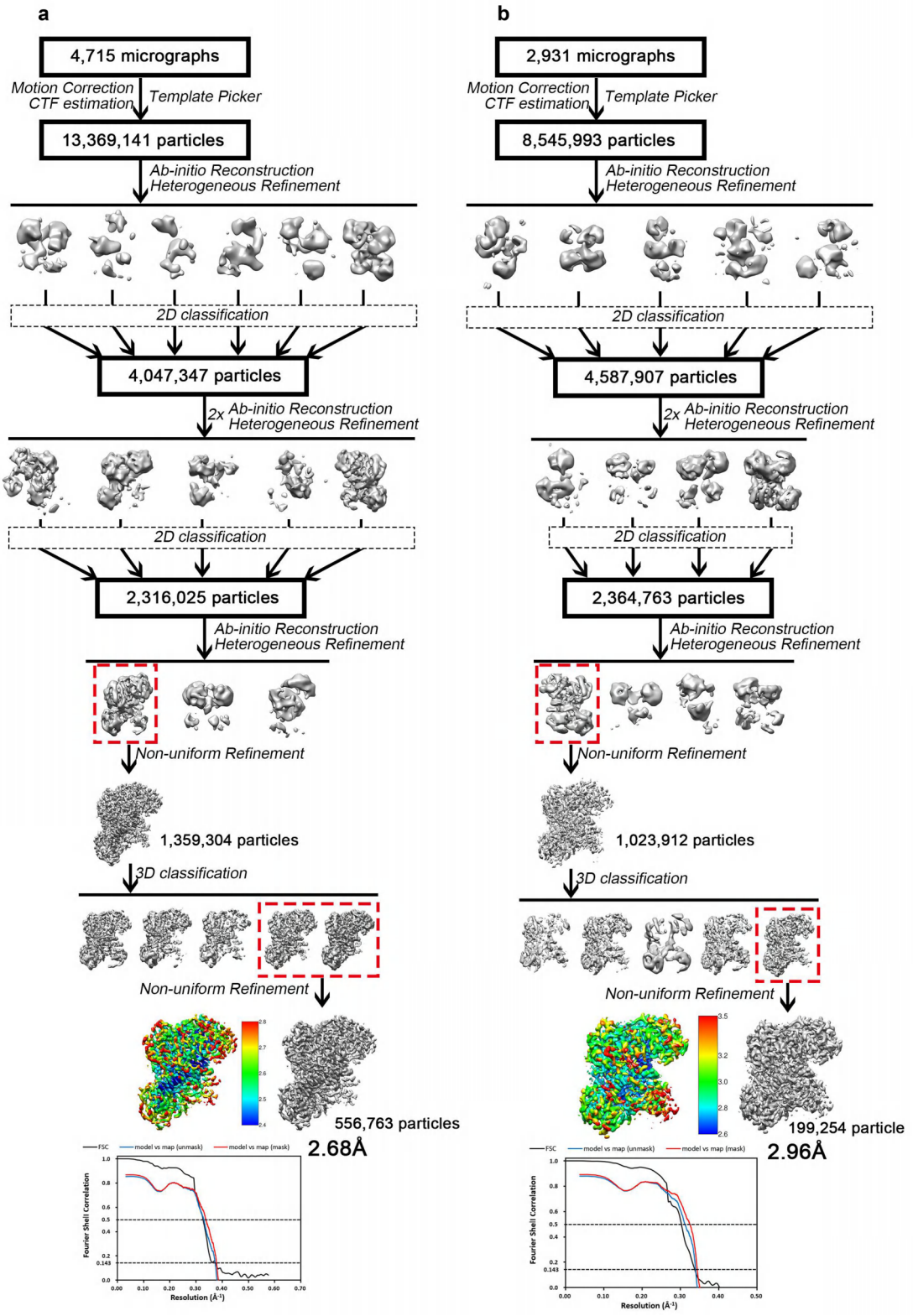


Supplementary Figure S2. Purification of KmAgo and KmAgo-DM electron microscopy samples. a-c, Size-exclusion chromatography of KmAgo on Superdex 200

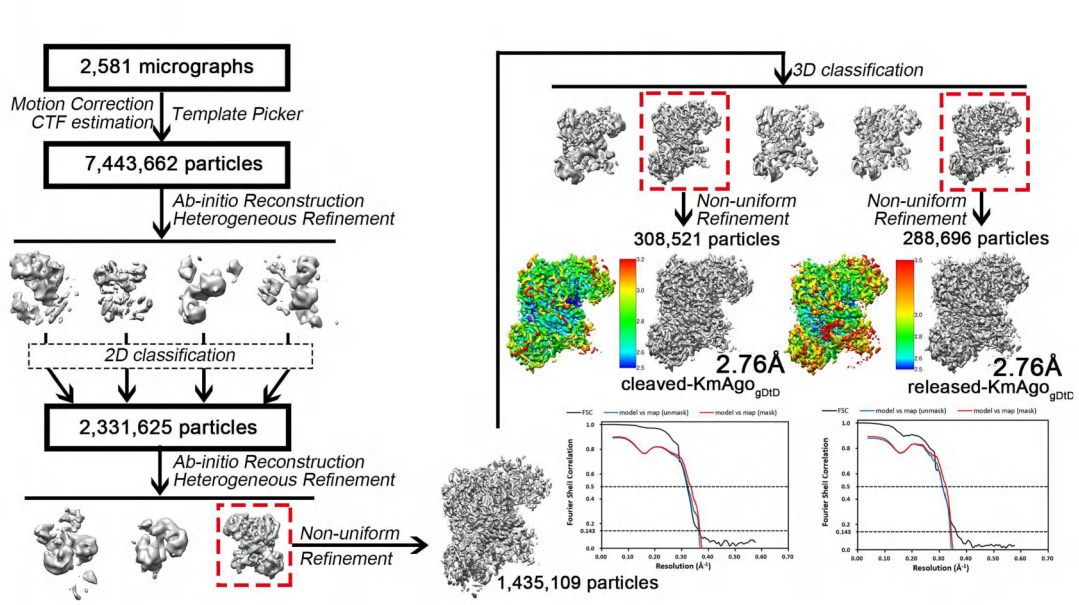
increase 10/300 GL (a), SDS-PAGE analysis (b) and negative staining of the final sample (c). **d-f,** Size-exclusion chromatography of KmAgo-DM on Superdex 200 increase 10/300 GL (d), SDS-PAGE analysis (e) and negative staining of the final sample (f).



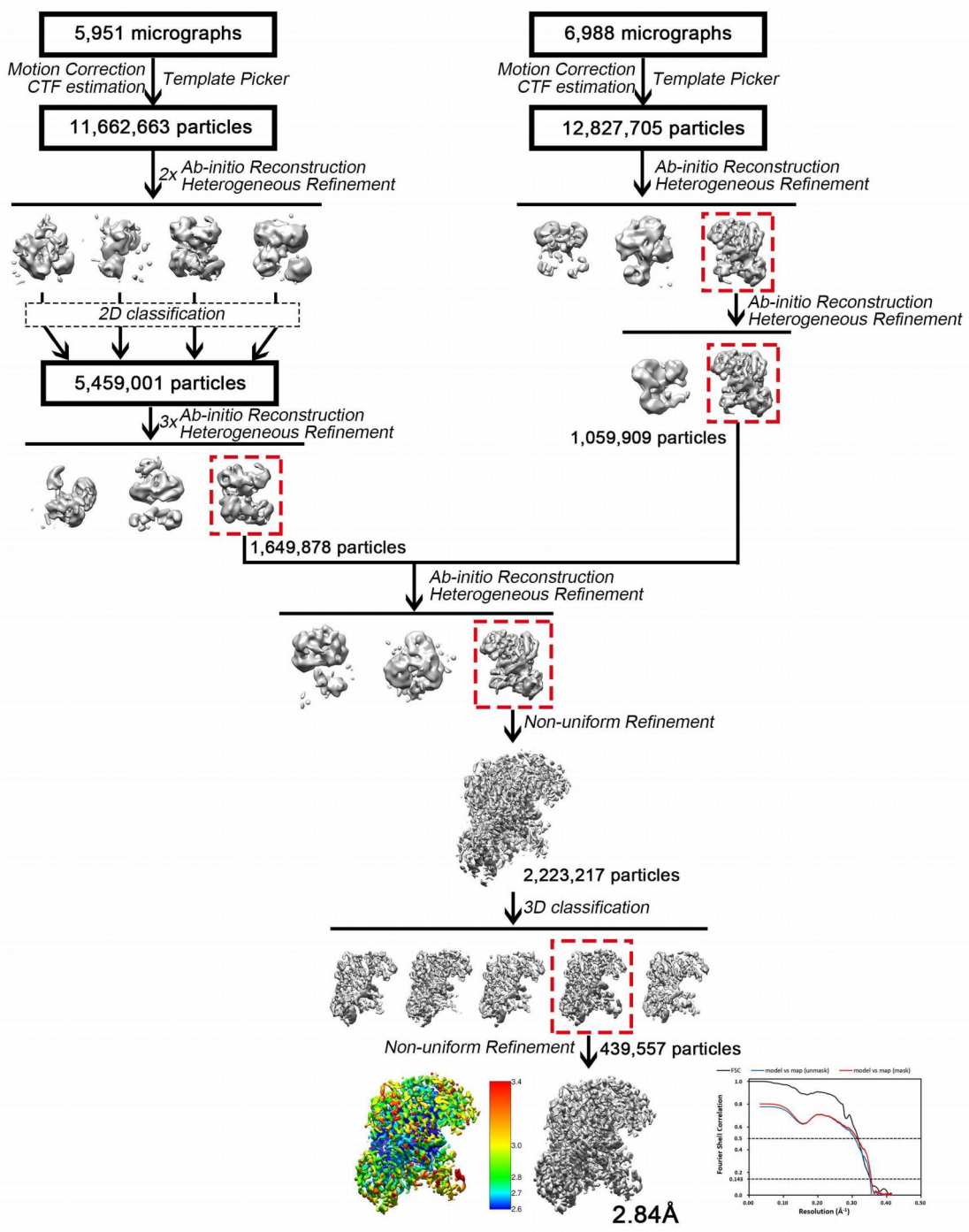
Supplementary Figure S3. Structure determination of $KmAgo_{apo}$ and $KmAgo_{gDNA}$. a, Flowchart of image processing for $KmAgo_{apo}$. **b,** Flowchart of image processing for $KmAgo_{gDNA}$.



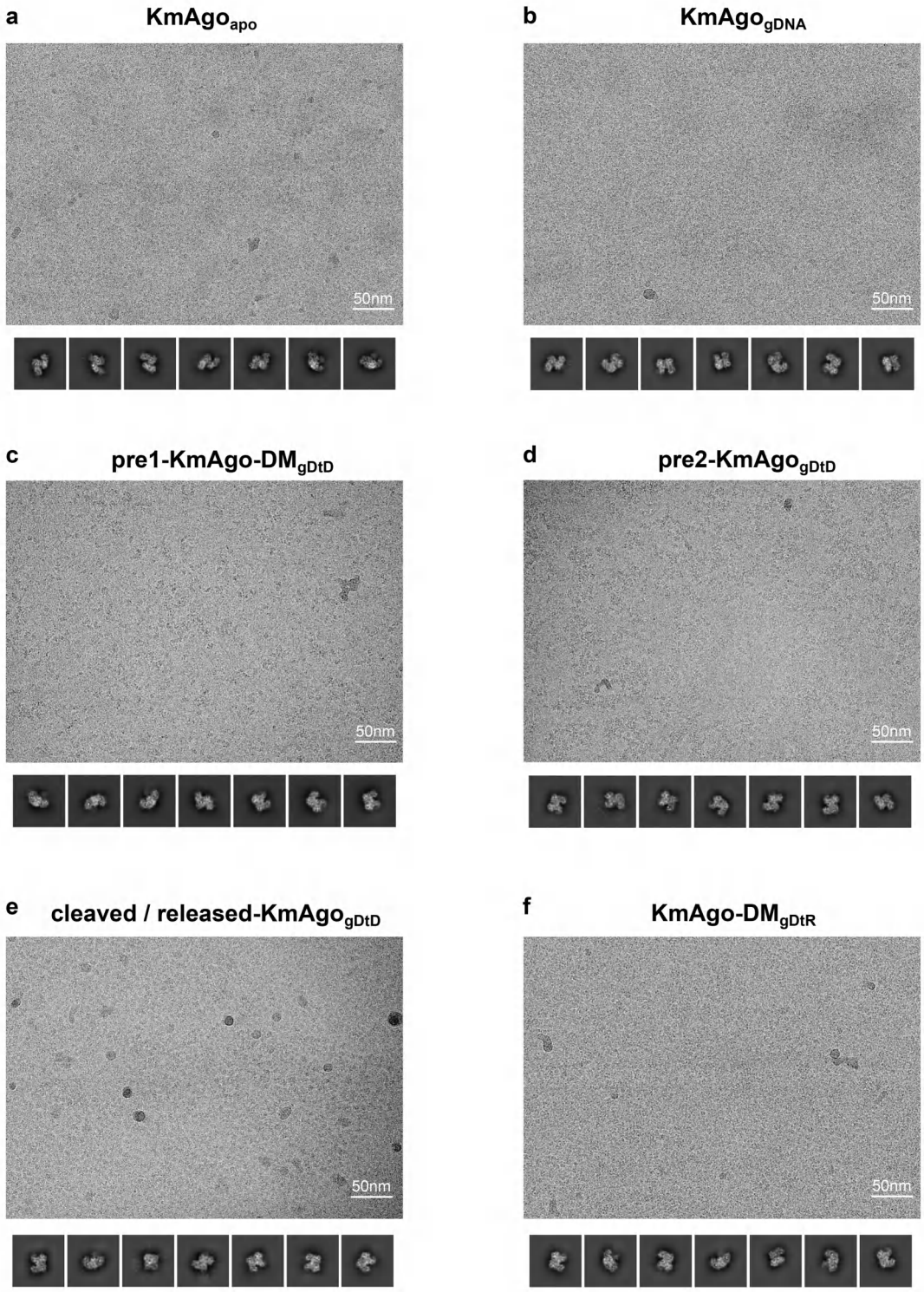
Supplementary Figure S4. Structure determination of pre1-KmAgo-DM$_{gDtD}$ and pre2-KmAgo$_{gDtD}$. a, Flowchart of image processing for pre1-KmAgo-DM$_{gDtD}$. **b,** Flowchart of image processing for pre2-KmAgo$_{gDtD}$.



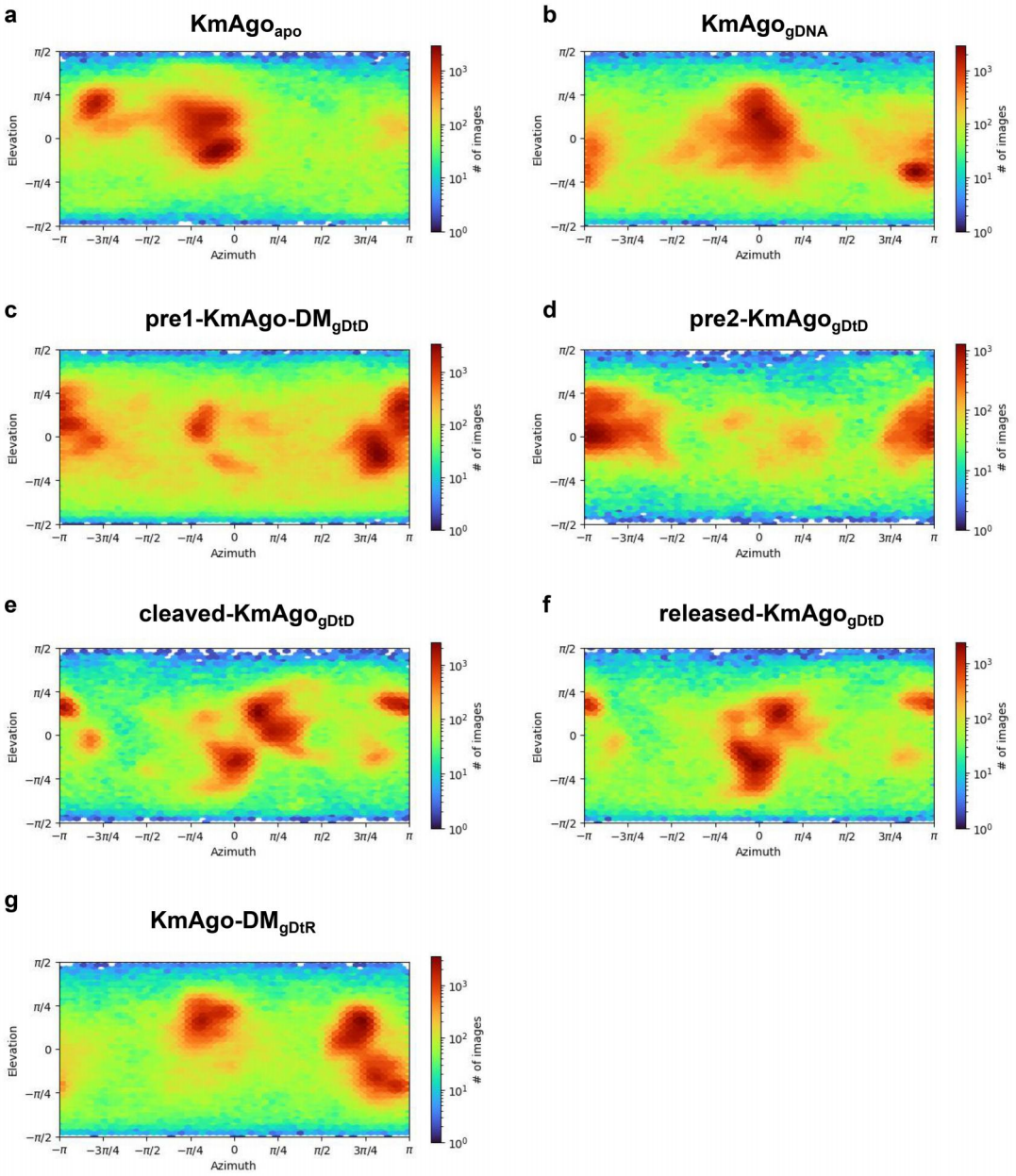
Supplementary Figure S5. Structure determination of cleaved-KmAgo$_{gDtD}$ and released-KmAgo$_{gDtD}$. Flowchart of image processing for cleaved-KmAgo$_{gDtD}$ and released-KmAgo$_{gDtD}$.



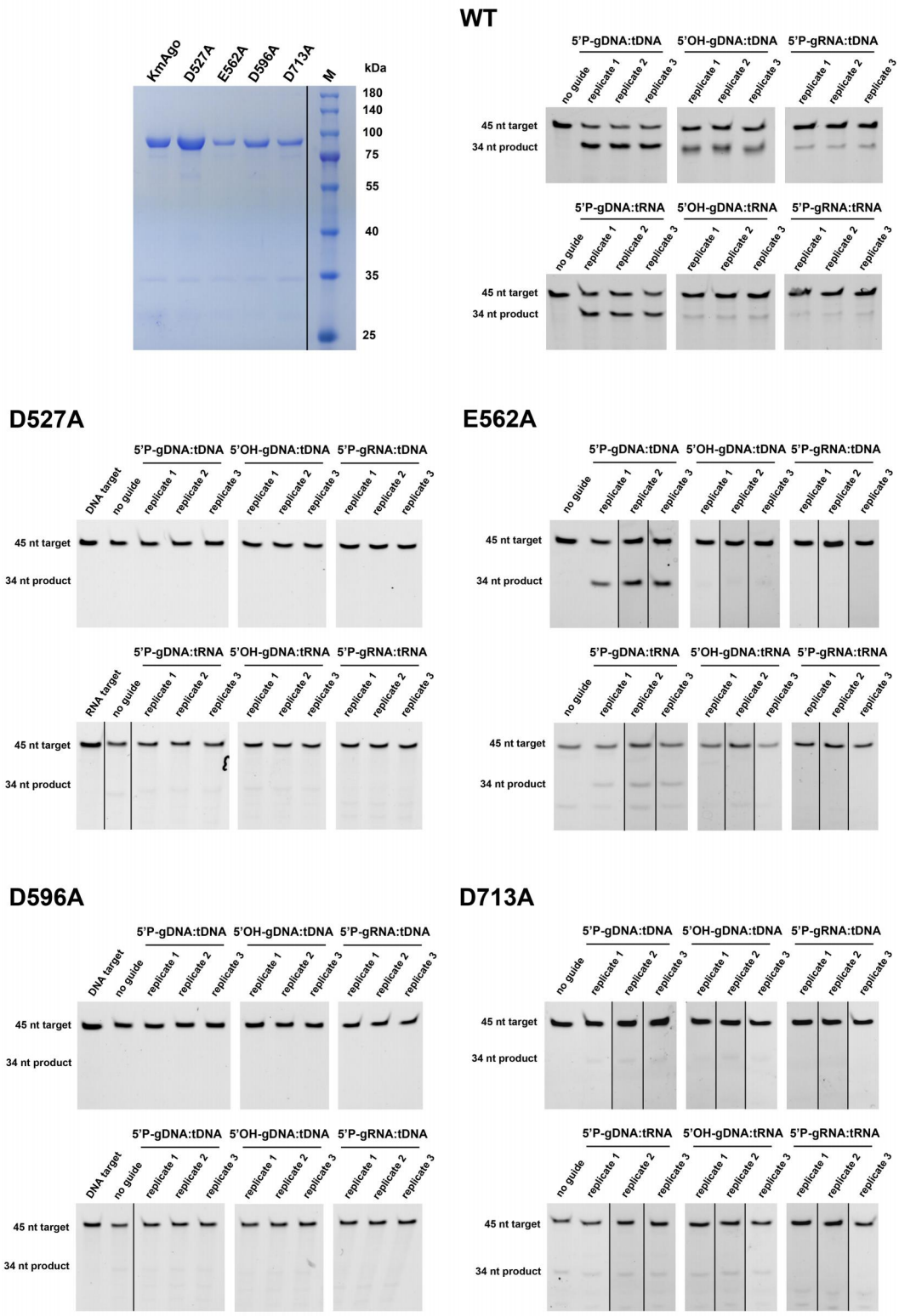
Supplementary Figure S6. Structure determination of KmAgo-DM$_{gDtR}$. Flowchart of image processing for KmAgo-DM$_{gDtR}$.



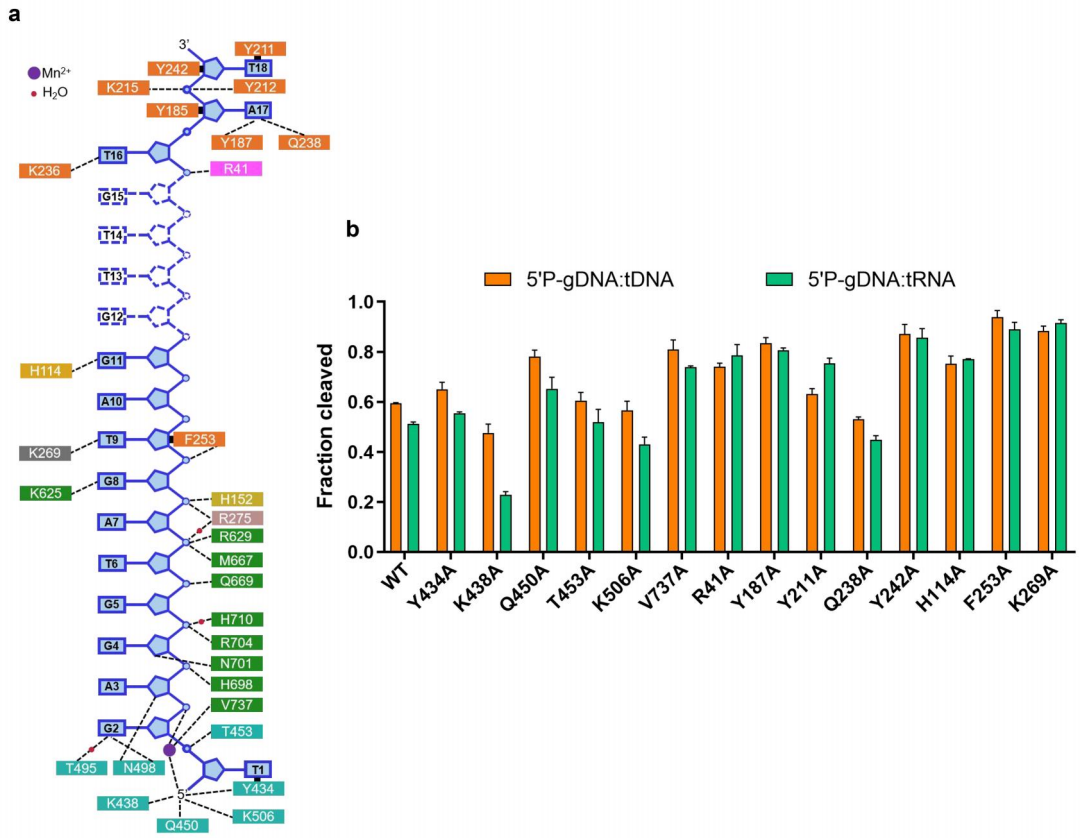
Supplementary Figure S7. Analysis of the quality of the cryo-EM datasets. Representative micrographs of the cryo-EM experiments (upper panel); Representative 2D class averages (lower panel).



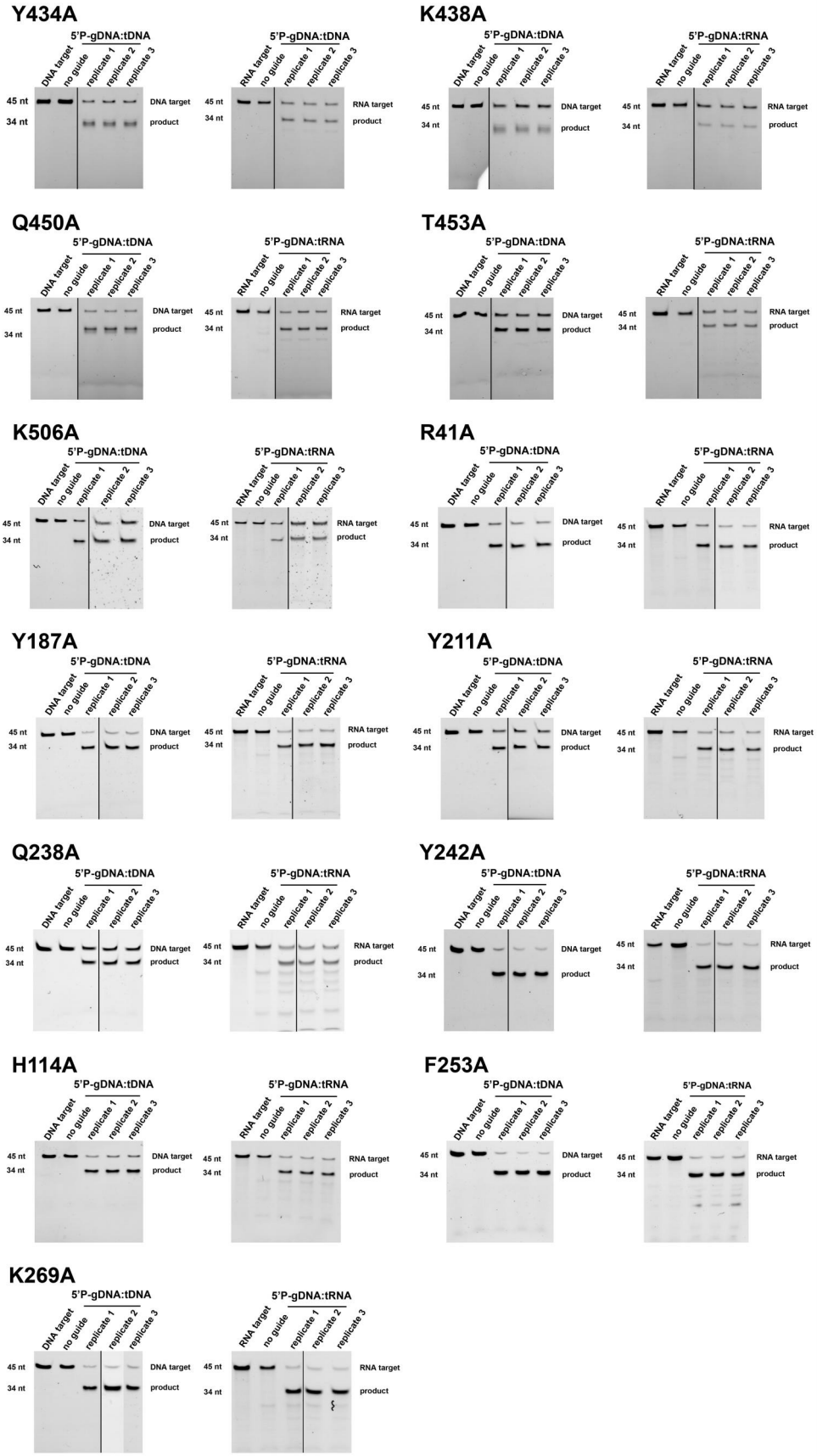
Supplementary Figure S8. Analysis of the quality of the cryo-EM datasets. Angular distributions of all the KmAgo structures as mentioned.



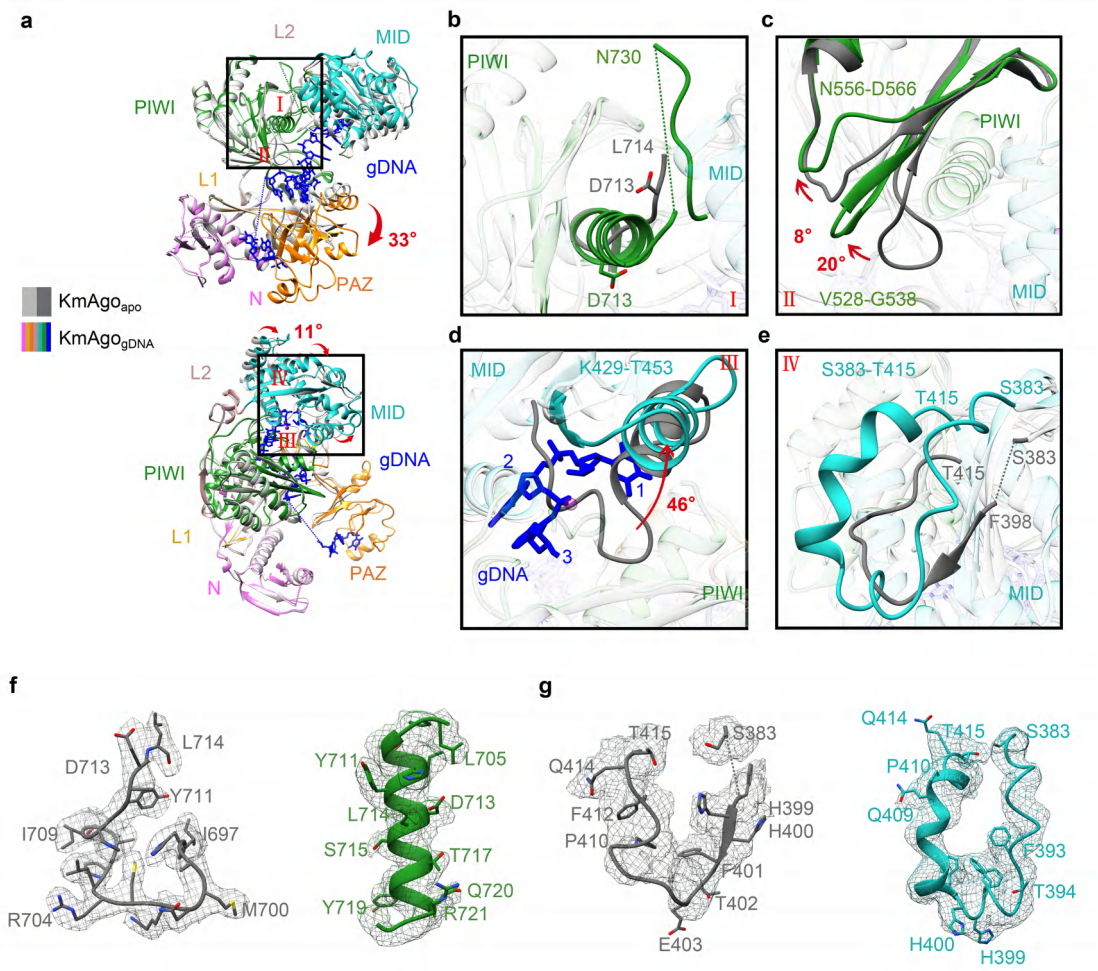
Supplementary Figure S9. The SDS-PAGE results of KmAgo and the urea-PAGE results corresponding to **Figure 1f.**



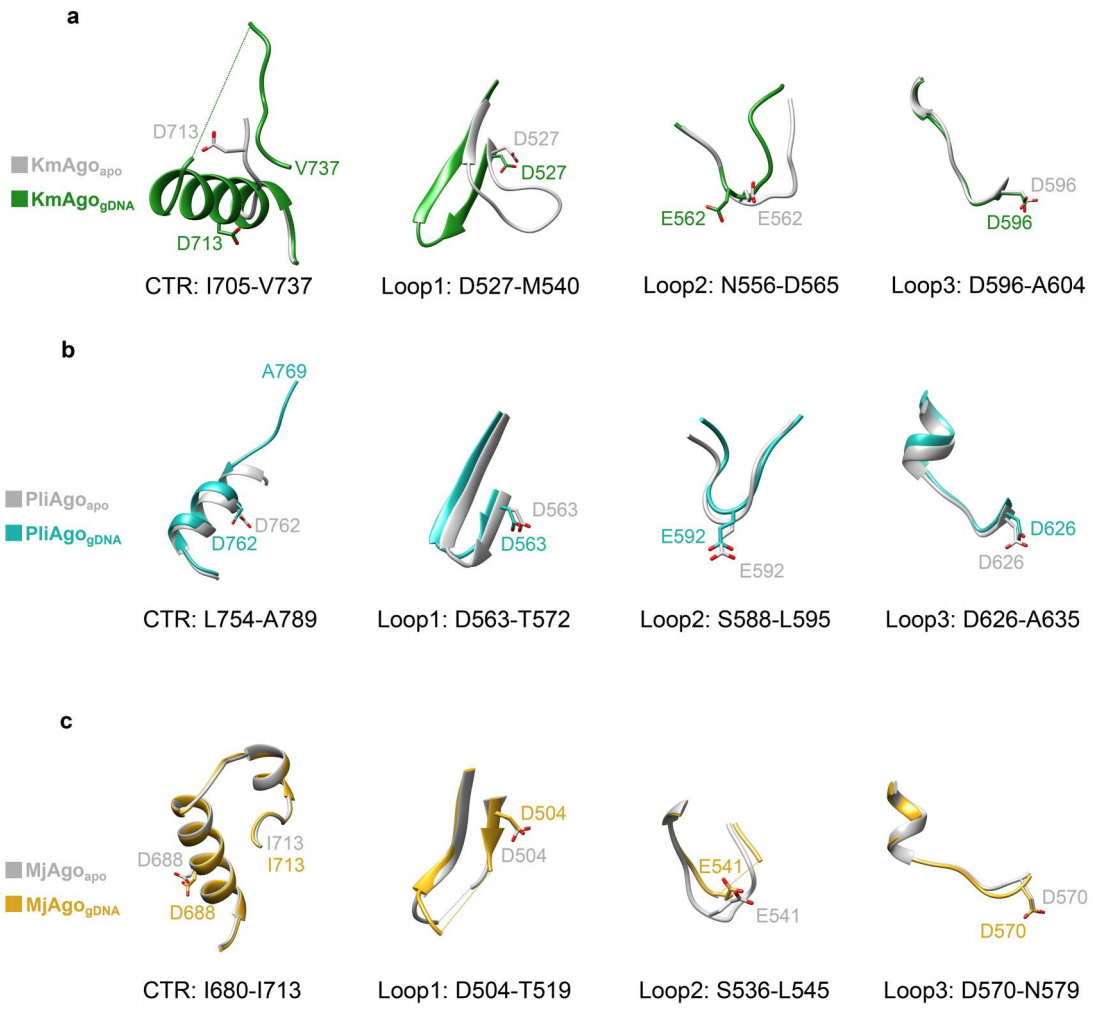
Supplementary Figure S10. 5′P-gDNA in the KmAgo$_{gDNA}$ binary complex. a, Schematic of major contacts between KmAgo and the 5′P-gDNA (blue). Residues colored by domain, as in Figure 2a. **b,** Cleavage activity of KmAgo amino acid residues responsible for guide positioning mutants. Average cleavage efficiencies from three times (technical replicate) are plotted, and error bars represent standard deviations.



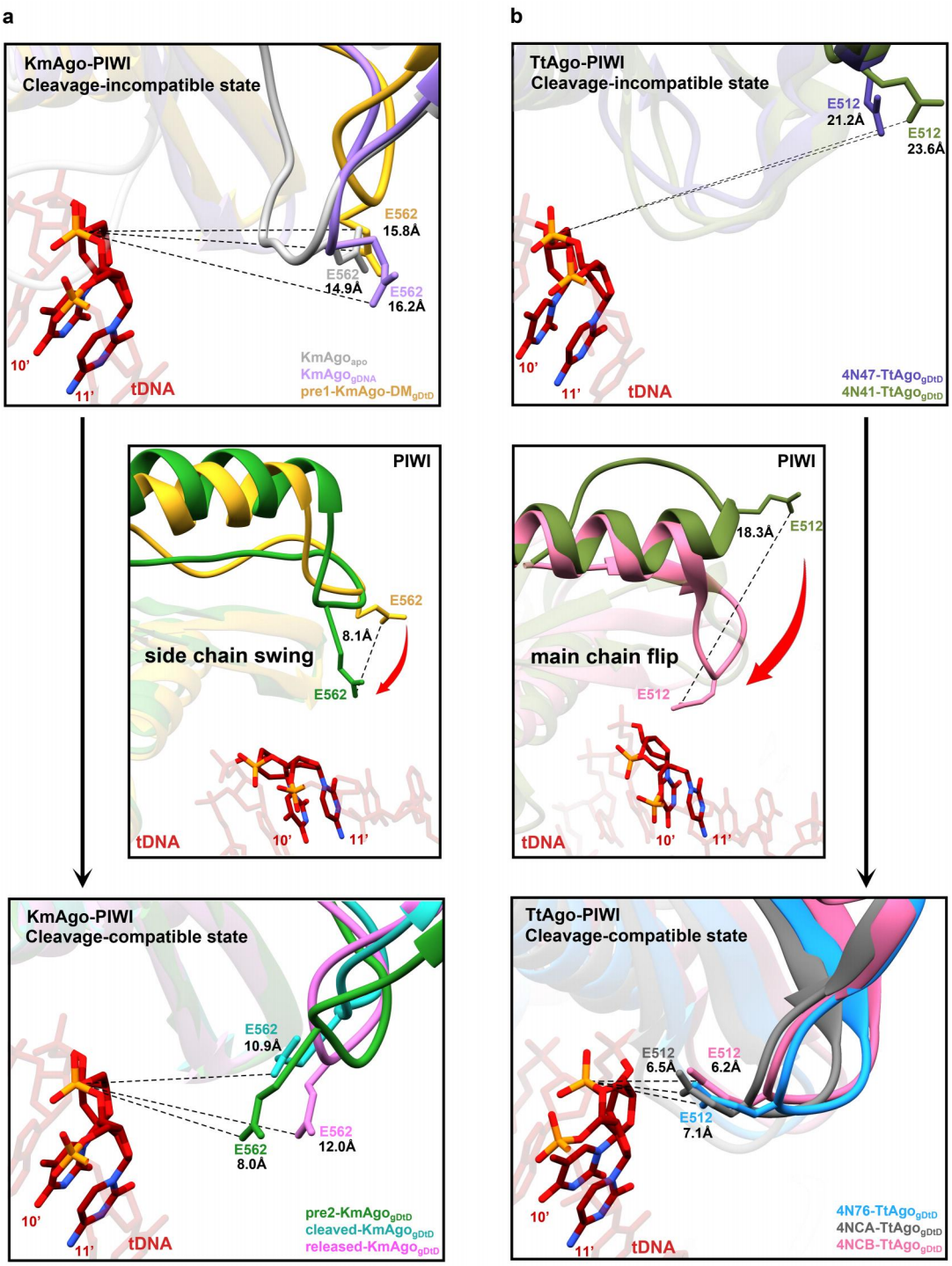
Supplementary Figure S11. The urea-PAGE results corresponding to **Supplementary Figure S10b.**



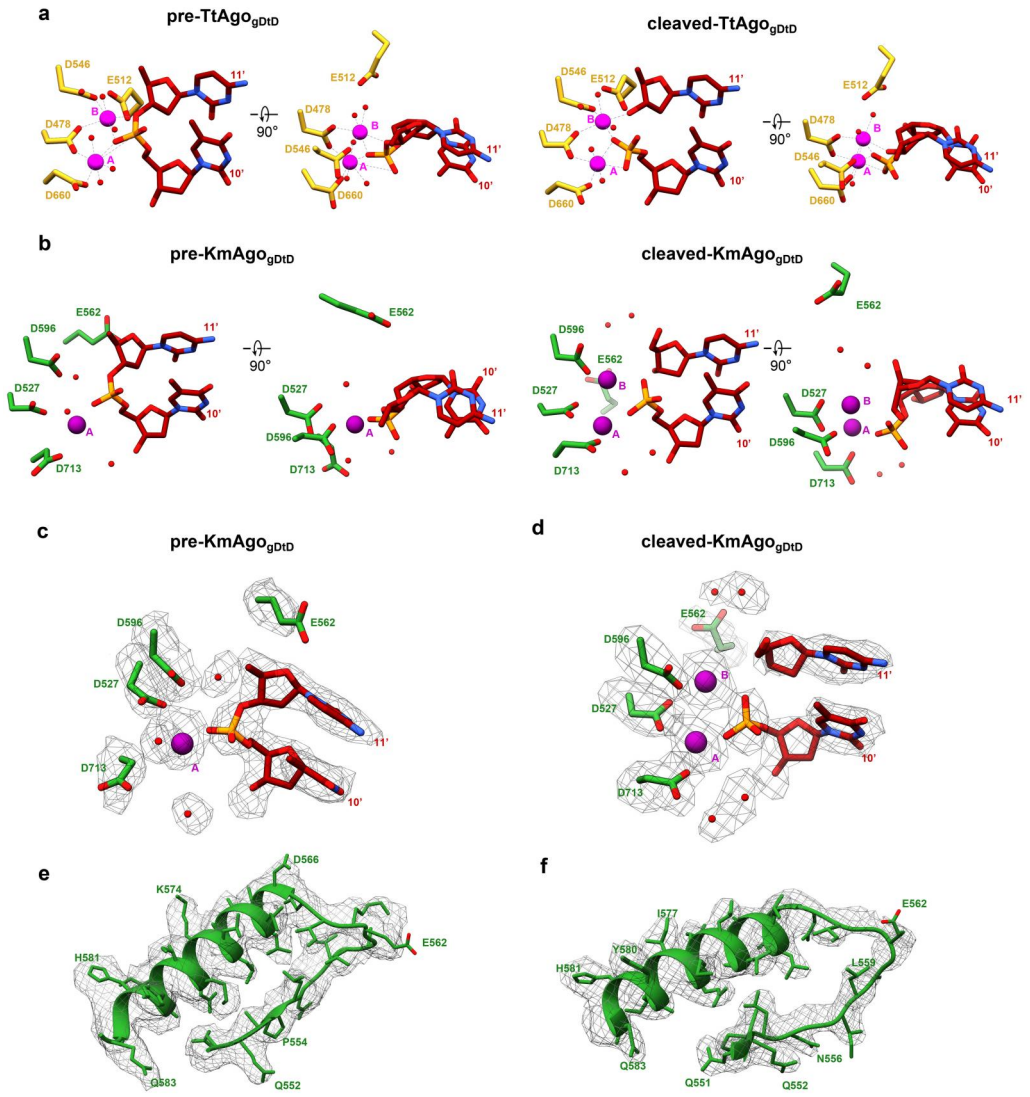
Supplementary Figure S12. Structural comparison between KmAgo$_{apo}$ (gray) and KmAgo$_{gDNA}$ binary complex (colored). a, gDNA shown as a blue stick representation. The rotational movement of the PAZ domain and MID domain upon gDNA binding is highlighted by red arrows. **b-e,** Detailed conformational changes upon gDNA binding in PIWI domain (b, c) and MID domain (d, e). **f-g,** Electron density maps correspond to the conformation changes described in panel b and panel e, respectively.



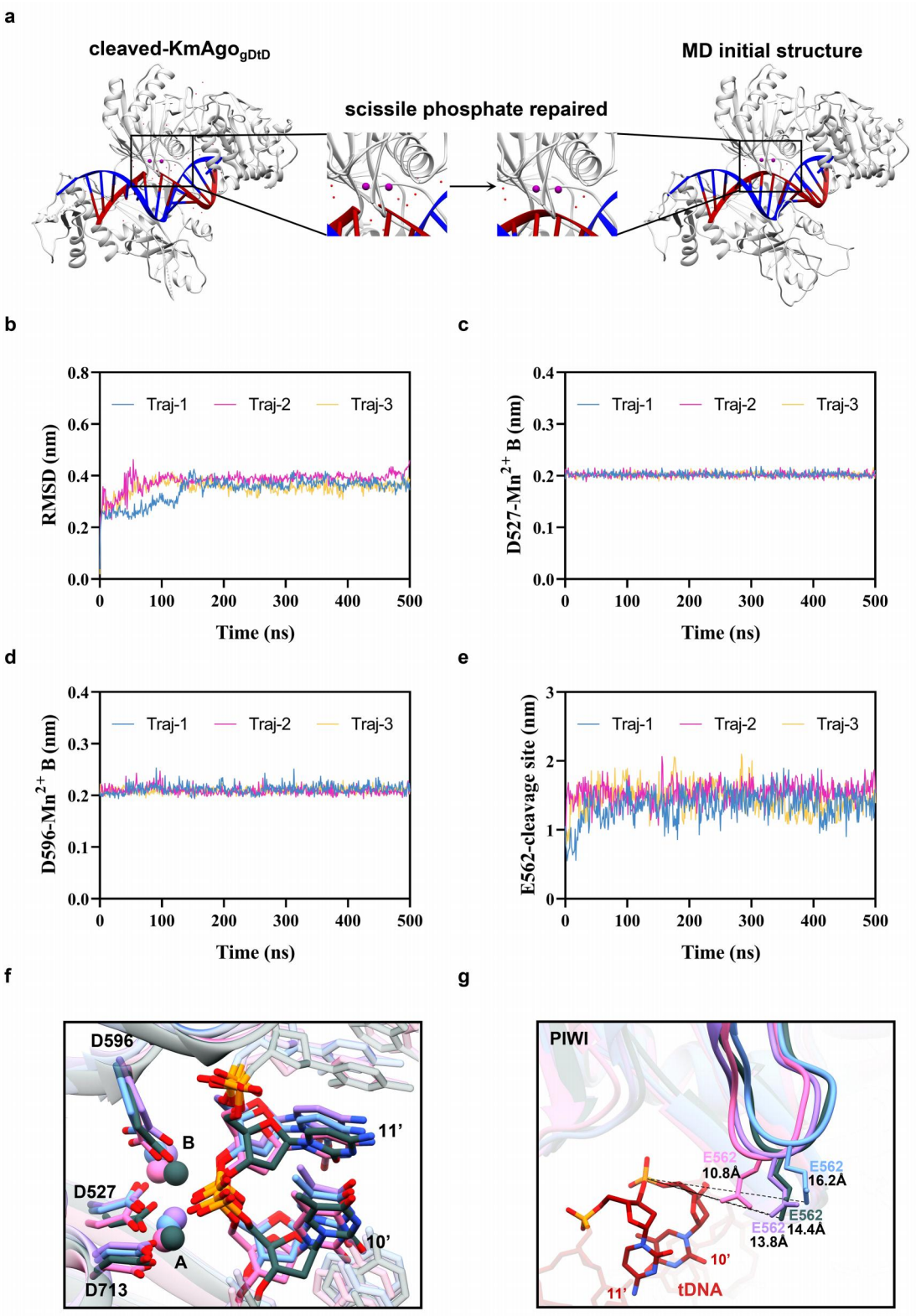
Supplementary Figure S13. Conformational changes in three loops and C-terminal region (CTR). a, Superposition of loop (1-3) and CTR of KmAgo during the transition from the free state (PDB: 8XHS) to the guide DNA-binding state (PDB: 8XHV). **b,** Superposition of loop (1-3) and CTR of PliAgo during the transition from the free state (PDB: 7R8F) to the guide DNA-binding state (PDB: 7R8H). **c,** Superposition of loop (1-3) and CTR of MjAgo during the transition from the free state (PDB: 5G5S) to the guide DNA-binding state (PDB: 5G5T).



Supplementary Figure S14. Conformational dynamics of the glutamate finger in KmAgo (a) and TtAgo (b). The distance between the glutamate finger and the cleavage site of tDNA is measured and depicted using black dashed lines. Red arrows indicate the trajectory and direction of movement.

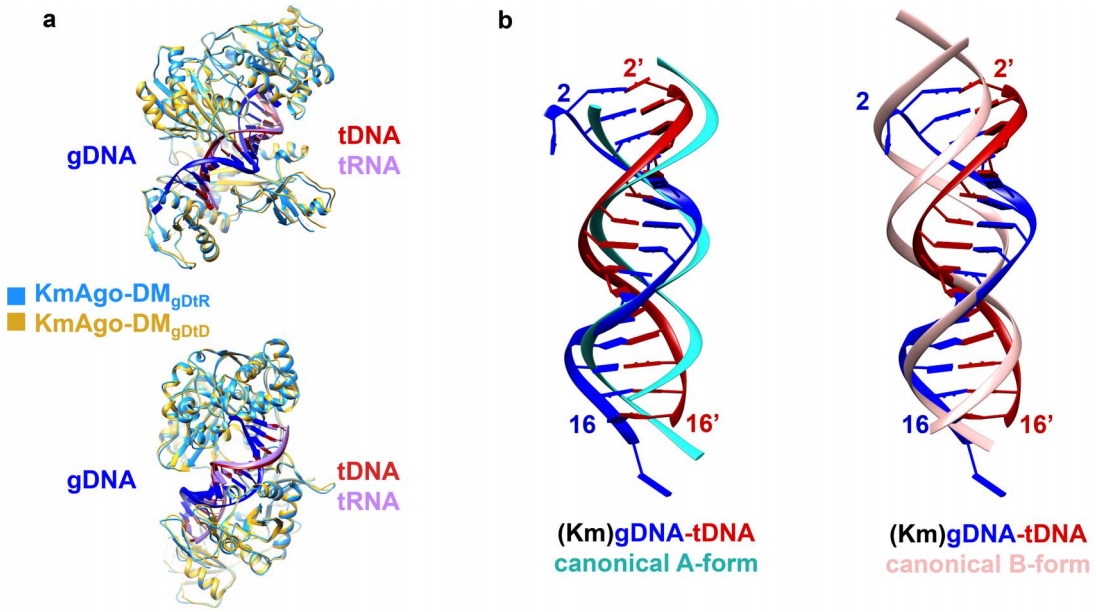


Supplementary Figure S15. Comparison of the catalytic pocket in TtAgo and KmAgo. **a,** Stereoview of the catalytic pocket in the TtAgo ternary complex, showing both intact (PDB: 4NCB) and cleaved (PDB: 4NCA) target strands. The Mg^{2+} cations, labeled "A" and "B", are depicted as magenta balls. Water molecules are depicted as red balls. The four catalytic residues, D478, D546, D660, and E512, are represented using stick. **b,** Stereoview of the catalytic pocket in the KmAgo ternary complex, showing both intact (PDB: 8XK0) and cleaved (PDB: 8XK3) target strands. The Mn^{2+} cations, labeled "A" and "B", are depicted as dark magenta balls. Water molecules are depicted as red balls. The four catalytic residues, D527, D596, D713, and E562, are represented using stick. **c-d,** The electron density map corresponding to Supplementary Figure S15b is shown as a dark gray mesh. **e-f,** The electron density map corresponding to the region near E562 in Supplementary Figure S15b is shown as a dark gray mesh.

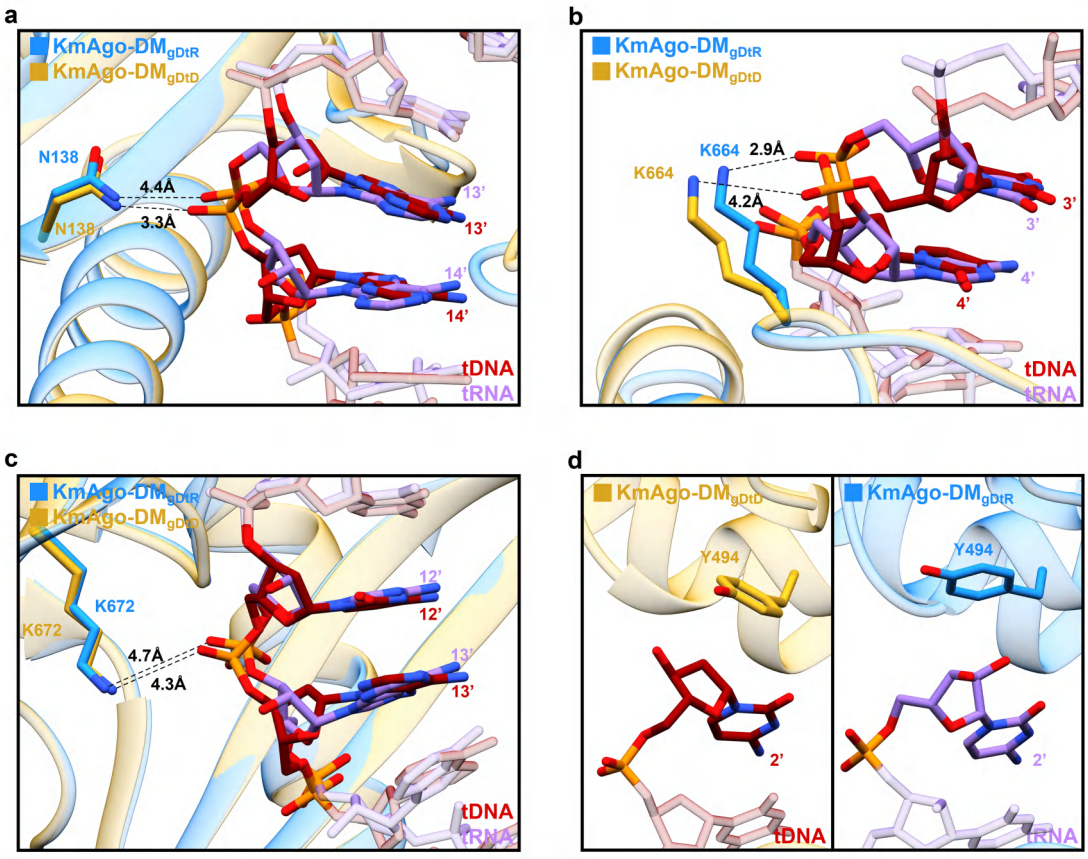


Supplementary Figure S16. MD simulations of the KmAgo ternary complexes. a, Construction of the initial structure for MD simulations. **b,** The RMSD of protein C- alpha atoms during the 500 ns simulation for three trajectories. **c-e,** Time series showing the

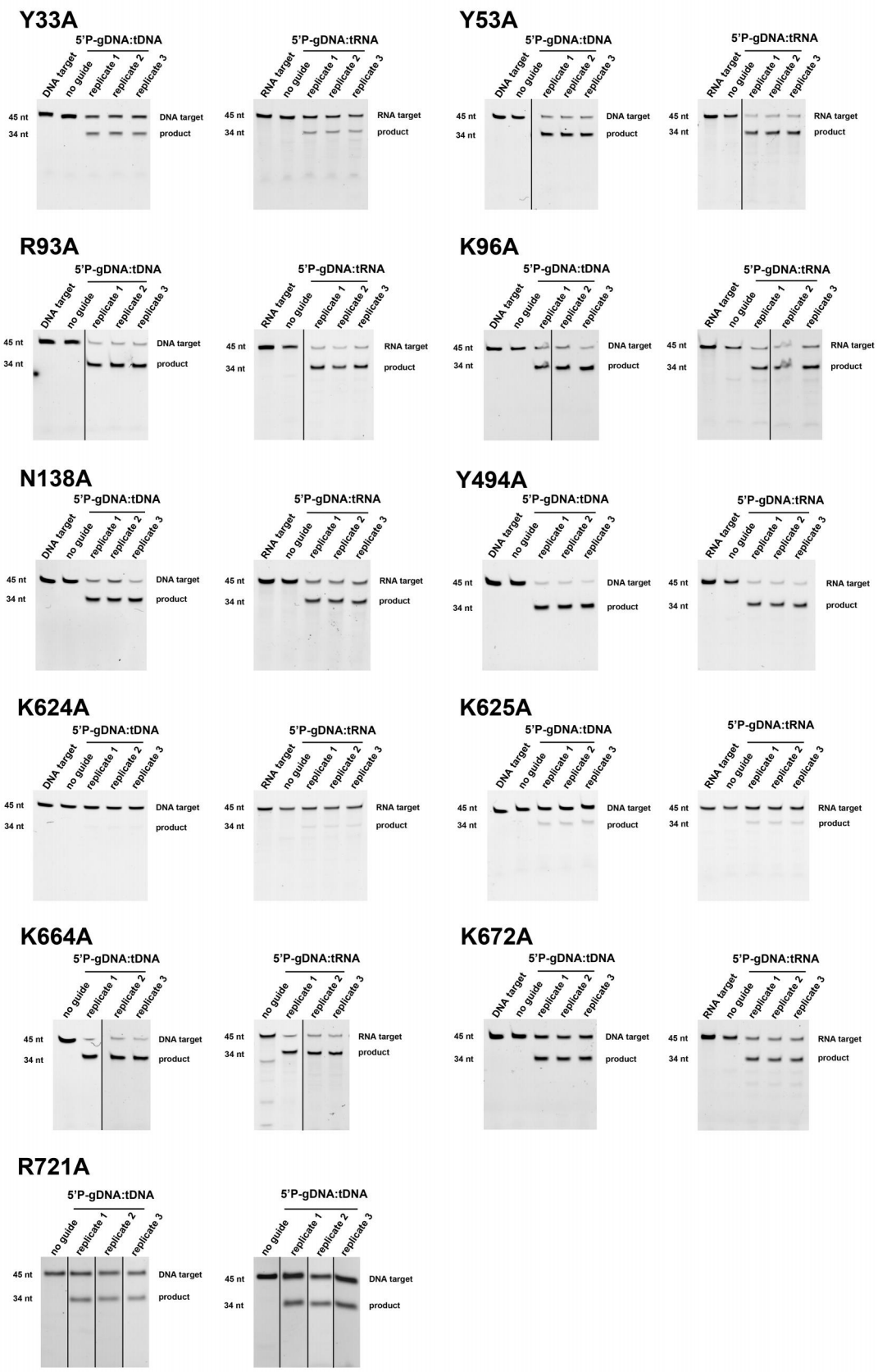
distance among residues D527, D596 to Mn^{2+} B and E562 to the cleavage site. **f-g,** Snapshots of catalytic pocket and the loop containing E562.



Supplementary Figure S17. Structural comparison of KmAgo-DM ternary complexes. **a,** Overlaid structures of KmAgo-DM ternary complexes at two different views. KmAgo-DM_{gDtD} in yellow, KmAgo-DM_{gDtR} in doggle blue. The gDNA-tDNA and gDNA-tRNA duplexes are colored as Figure 6e. **b,** Comparison of gDNA-tDNA duplex in KmAgo-DM ternary complex with canonical A-DNA (in light sea green) and canonical B-DNA (in rosy brown).



Supplementary Figure S18. Minor differences between KmAgo-DM and guide-target duplex. The color is consistent with that depicted in Supplementary Figure S17a.



Supplementary Figure S19. The urea-PAGE results corresponding to **Figure 7b**.

Supplementary Table 1. Structures of Ago proteins and their complexes.

	Argonaute Species	Structures (PDB ID)					
		Apo	Binary complex	Ternary complex			
				Target recognition	Target cleavage	Product release	Catalytic inactive mutants
Mesophilic pAgos	**KmAgo** *Kurthia massiliensis*	√ 8XHS	√ 8XHV	√ 8XK0	√ 8XK3	√ 8XK4	√ 8XJX 8XJW
	CbAgo *Clostridium butyricum*	–	–	–	–	–	√ 6QZK
	PliAgo *Pseudooceanicola lipolyticus*	√ 7R8F	√ 7R8H 7R8G	–	–	–	–
Thermophilic pAgos	**TtAgo** *Thermus thermophilus*	–	√ 3DLB 3DLH	√ 4N47 4N41 4NCB 3HVR 3HM9 3HXM 3F73	√ 4NCA 4KPY	√ 4N76	√ 3HO1 3HJF 3HK2
	AaAgo *Aquifex aeolicus*	√ 1YVU	–	√ 2F8S 2F8T	–	–	–
	AfAgo *Archaeoglobus fulgidus*	√ 1W9H	–	√ 1YTU 2BGG 2W42	–	–	–
	MpAgo *Marinitoga piezophila*	–	√ 5I4A	–	–	–	√ 5UX0
	MjAgo *Methanocaldococcus jannaschii*	√ 5G5S	√ 5G5T	–	–	–	–
	PfAgo *Pyrococcus furiosus*	√ 1U04 1Z25 1Z26	–	√ 8JPX	–	–	–
	TtdAgo *Thermococcus thioreducens*	–	–	√ 8WD8	–	–	–
	RsAgo *Rhodobacter sphaeroides*	–	–	√ 5AWH 6D8P 6D92 6D95 6D9K 6D9L 6D9A 6D8F	–	–	–
eAgos	**hAgo1** *Homo sapiens*	–	√ 4KRF	–	–	–	–
	hAgo2 *Homo sapiens*	–	√ 4F3T 4OLA	√ 4W5N 4W5O 4W5Q 4W5R 4W5T	–	–	–
	hAgo3 *Homo sapiens*	–	√ 5VM9	–	–	–	–
	KpAgo *Kluyveromyces polysporus*	–	√ 4F1N	–	–	–	–
	AtAgo *Arabidopsis thaliana*	–	√ 7SVA	–	–	–	√ 7SWQ 7SWF

Supplementary Table 2. Oligonucleotides used in the single-stranded nucleic acid cleavage assays.

Name	Sequence (5'-3')
T-tDNA	AAACGACGGCCAGTGCCAAGCTTACTA TACAACCTACTACCTCAT
U-tRNA	AAACGACGGCCAGUGCCAAGCUUACUA UACAACCUACUACCUCAU
34nt DNA product	FAM-AAACGACGGCCAGTGCCAA GCTTACTATACAACC
34nt RNA product	FAM-AAACGACGGCCAGUGCCAA GCUUACUAUACAACC
5'P-T-gDNA	5'P-TGAGGTAGTAGGTTGTAT
5'P-U-gRNA	5'P-UGAGGUAGUAGGUUGUAU
T-gDNA	TGAGGTAGTAGGTTGTAT

Supplementary Table 3. Sequence identity of Ago proteins aligned with KmAgo.

	hAgo2	KpAgo	CbAgo	MpAgo	PliAgo	RsAgo	TtAgo	PfAgo	MjAgo
Identity (%)	15.66	12.32	22.72	16.09	14.47	14.13	17.79	16.04	14.48