

Molecular basis of Gabija anti-phage supramolecular assemblies

Received: 6 August 2023

Accepted: 22 March 2024

Published online: 16 April 2024

 Check for updates

Xiao-Yuan Yang ^{1,2,3,6}, Zhangfei Shen ^{1,2,6}, Jiale Xie^{1,2,3}, Jacelyn Greenwald ⁴, Ila Marathe⁴, Qingpeng Lin^{1,2}, Wen Jun Xie ⁵, Vicki H. Wysocki⁴ & Tian-Min Fu ^{1,2,3} ✉

As one of the most prevalent anti-phage defense systems in prokaryotes, Gabija consists of a Gabija protein A (GajA) and a Gabija protein B (GajB). The assembly and function of the Gabija system remain unclear. Here we present cryo-EM structures of *Bacillus cereus* GajA and GajAB complex, revealing tetrameric and octameric assemblies, respectively. In the center of the complex, GajA assembles into a tetramer, which recruits two sets of GajB dimer at opposite sides of the complex, resulting in a 4:4 GajAB supramolecular complex for anti-phage defense. Further biochemical analysis showed that GajA alone is sufficient to cut double-stranded DNA and plasmid DNA, which can be inhibited by ATP. Unexpectedly, the GajAB displays enhanced activity for plasmid DNA, suggesting a role of substrate selection by GajB. Together, our study defines a framework for understanding anti-phage immune defense by the GajAB complex.

To mitigate phage infections, bacteria have evolved highly diverse anti-phage immune systems¹. Though some bacterial immune systems like CRISPR–Cas and cyclic-oligonucleotide-based anti-phage signaling system (CBASS) have been extensively studied^{2,3}, many newly identified systems remain unexplored^{4–6}. The study of bacterial immune system not only offers evolutionary perspectives on immune systems but also provides invaluable tools for biomedical research and disease treatment.

As a newly identified bacterial immune system, the bacterial Gabija defense system exists in at least 8.5% of sequenced genomes with two components, Gabija protein A (GajA) and Gabija protein B (GajB)^{4,7}. GajA was shown to be an endonuclease that can recognize specific DNA sequence⁸. GajB was predicted to be a UvrD-like helicase⁹. However, whether and how GajA and GajB assemble into a complex for anti-phage defense remains unclear.

In this article, we present the cryo-electron microscopy (cryo-EM) structures of GajA alone and the GajAB complex. GajA assembles into a tetramer via interactions mediated by both the ATPase domain and

the nuclease domain. We also reveal that the GajA and GajB assemble into a heteromeric octamer with four molecules of GajA and four molecules of GajB, which is critical for the anti-phage defense. Given many other supramolecular assemblies identified in bacterial immunity, we propose that supramolecular assemblies may represent a unified mechanism in bacterial immune defense.

Results

Structure of GajA

To biochemically characterize GajA, we expressed and purified GajA in *Escherichia coli* BL21(DE3) (Extended Data Fig. 1a,b). The elution volume of GajA on gel filtration indicated that GajA formed an oligomer (Extended Data Fig. 1a). To reveal the assembly of GajA, we employed cryo-EM single particle analysis to determine the structure of GajA. However, GajA had a severe orientation preference problem on grids, leading to a reconstruction with a nominal resolution of 2.9 Å but poor densities (Extended Data Fig. 1c–g). To resolve this issue, we optimized conditions for grid preparation and eventually obtained a cryo-EM

¹Department of Biological Chemistry and Pharmacology, Center for RNA Biology, The Ohio State University, Columbus, OH, USA. ²The Ohio State University Comprehensive Cancer Center, Columbus, OH, USA. ³Program of OSBP, The Ohio State University, Columbus, OH, USA. ⁴Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA. ⁵Department of Medicinal Chemistry, University of Florida, Gainesville, FL, USA. ⁶These authors contributed equally: Xiao-Yuan Yang, Zhangfei Shen. ✉e-mail: Fu.978@osu.edu

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

	Gabija A tetramer (4A), (EMD-41319), (PDB 8TKO)	Gabija AB complex 1, (4A:4B, C1 symmetry), (EMD-41321), (PDB 8TK1)	Gabija AB complex 2, (4A:4B, D2 symmetry), (EMD-41314), (PDB 8TJY)
Data collection and processing			
Magnification	81,000×	81,000×	81,000×
Voltage (kV)	300	300	300
Electron exposure (e ⁻ Å ⁻²)	50	50	50
Defocus range (μm)	1.0–2.0	0.5–2.0	0.5–2.0
Pixel size (Å)	0.899	1.12	1.12
Symmetry imposed	D2	C1	D2
Initial images (no.)	1,364	7,173	7,173
Initial particle images (no.)	849,640	6,928,153	6,928,153
Final particle images (no.)	96,633	942,091	942,091
Map resolution (Å)	3.23	2.98	2.79
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	3.0–5.0	2.8–6.0	2.6–6.0
Refinement			
Initial model used (PDB code)	AlphaFold	AlphaFold	AlphaFold
Model resolution (Å)	3.4	3.2	3.1
FSC threshold	0.5	0.5	0.5
Model resolution range (Å)	3.1–7.0	3.0–8.0	2.9–8.0
Map sharpening B factor (Å ²)	-141.5	-110.9	-121.5
Model composition			
Nonhydrogen atoms	14,816	23,424	22,692
Protein residues	1,816	2,877	2,788
Ligands	0	0	0
Nucleotide	0	0	0
B factors (Å²)			
Protein	78.80	47.07	45.75
Ligand			
Nucleotide			
Root mean square deviations			
Bond lengths (Å)	0.004	0.004	0.003
Bond angles (°)	0.614	0.938	0.469
Validation			
MolProbity score	1.71	1.40	1.40
Clashscore	5.33	3.42	3.34
Poor rotamers (%)	0.00	0.08	0.00
Ramachandran plot			
Favored (%)	93.61	96.12	96.01
Allowed (%)	6.39	3.88	3.99
Disallowed (%)	0.00	0.00	0.00

structure of GajA with a resolution of 3.2 Å by collecting a dataset using a grid with relatively thicker ice (Table 1 and Extended Data Figs. 1a and 2a–e). Although the nominal resolution is lower from reconstruction of images with thicker ice, the EM densities have been substantially improved in comparison to the 2.9 Å structure (Extended Data Figs. 1g and 2f). The cryo-EM structure of GajA revealed a symmetric tetrameric assembly with dimensions of 175 Å × 115 Å × 50 Å (Fig. 1a,b).

Each protomer of GajA is composed of an N-terminal ATPase domain that is divided into two halves by an inserted dimerization domain, and a C-terminal nuclease domain (Fig. 1c–e). The N-terminal ATPase domain is composed of an 11-stranded mixed-paralleled β-sheet, sandwiching α1 and surrounded by α2–α8 (Fig. 1d and Extended Data Fig. 3a). Structural comparison revealed that the GajA ATPase domain resembles the canonical ATP-binding cassette (ABC) ATPase proteins with a highly conserved ATP binding site^{10,11} (Extended Data Fig. 3b).

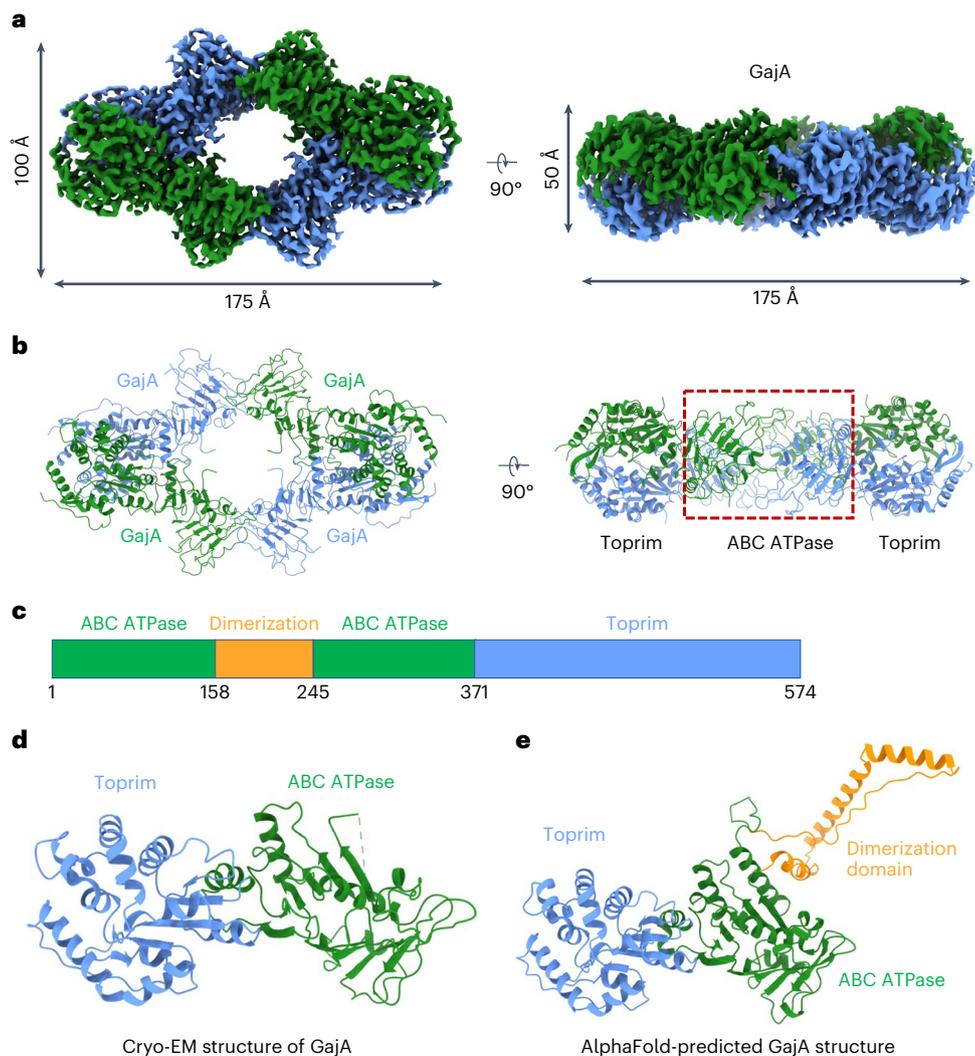


Fig. 1 | Cryo-EM structure of GajA. **a,b**, A cryo-EM density map (**a**) and ribbon diagrams (**b**) of GajA tetramer on two different planes: top view (left) and vertical rotation (right). **c**, Domain architecture of GajA. The ABC ATPase domain is

indicated in green, dimerization domain in orange and Toprim in blue. **d,e**, A ribbon diagram of a GajA protomer determined by cryo-EM reconstruction (**d**) or AlphaFold prediction (**e**) with domains colored as in **c**.

The GajA C-terminal domain folds as a topoisomerase–primase (Toprim) domain with a central four-stranded parallel β -sheet surrounded by α -helices^{12,13} (Fig. 1d and Extended Data Fig. 3c). Both the N-terminal ATPase domain and the C-terminal Toprim domain were clearly resolved in our cryo-EM structure (Fig. 1d). In contrast, the dimerization domain, predicted to consist of three α -helices by AlphaFold¹⁴, was invisible, indicating the flexibility of this domain (Fig. 1d,e). A recent cryo-EM structure of *Bacillus cereus* GajAB in complex with a phage protein Gad1 revealed that the dimerization domain of GajA not only contributed to the oligomerization of GajA but was also involved in recruiting Gad1 (ref. 15) (Extended Data Fig. 3d).

Assembly of GajA tetramer

The tetramerized GajA is arranged as a dimer of dimer with three types of interface, denoted as type I, type II and type III interfaces (Fig. 2a). The type I interface, located at the center of the GajA tetramer, is mediated by the first half of the ATPase domain with a buried surface area of 609 \AA^2 (Fig. 2b). Detailed analysis revealed that both hydrophobic and hydrophilic interactions are critical for the formation of type I interface (Extended Data Fig. 4a). In contrast, both hydrophobic and hydrophilic residues in the second half of the ATPase domain contribute

to the formation of type II interface that is characterized by a buried surface area of 325 \AA^2 (Fig. 2c and Extended Data Fig. 4b). Additionally, Toprim domains are arranged along a twofold axis to form the type III interface with an extensively buried surface area of 1,471 \AA^2 (Fig. 2d and Extended Data Fig. 4c). Together, interactions of these three interfaces govern the assembly of GajA tetramer.

Active site of GajA

The Toprim domains in GajA tetramer, which belong to a class of DNA endonucleases known as OLD (overcoming lysogenization defect)¹², dimerize at opposite sides of the elongated tetramer. Analysis of the two proximal Toprim domains reveals that the two active sites are about 20 \AA away from each other, suggesting that the two active sites work independently (Fig. 2e). Similar to other OLD nucleases^{12,13}, the active site of GajA is composed of a conserved DxD motif between β 3 and β 4, an invariant glutamate following β 1, and two glutamate residues on α 2 and α 7 (Fig. 2f). Collectively, these residues form a very acidic active site in GajA. Studies on BpOLD suggested a two-metal catalysis mechanism¹³, which may be shared by GajA due to the structural similarity of the active sites between GajA and BpOLD (Extended Data Fig. 4d). The conserved DxD motif and E379 may coordinate a magnesium ion

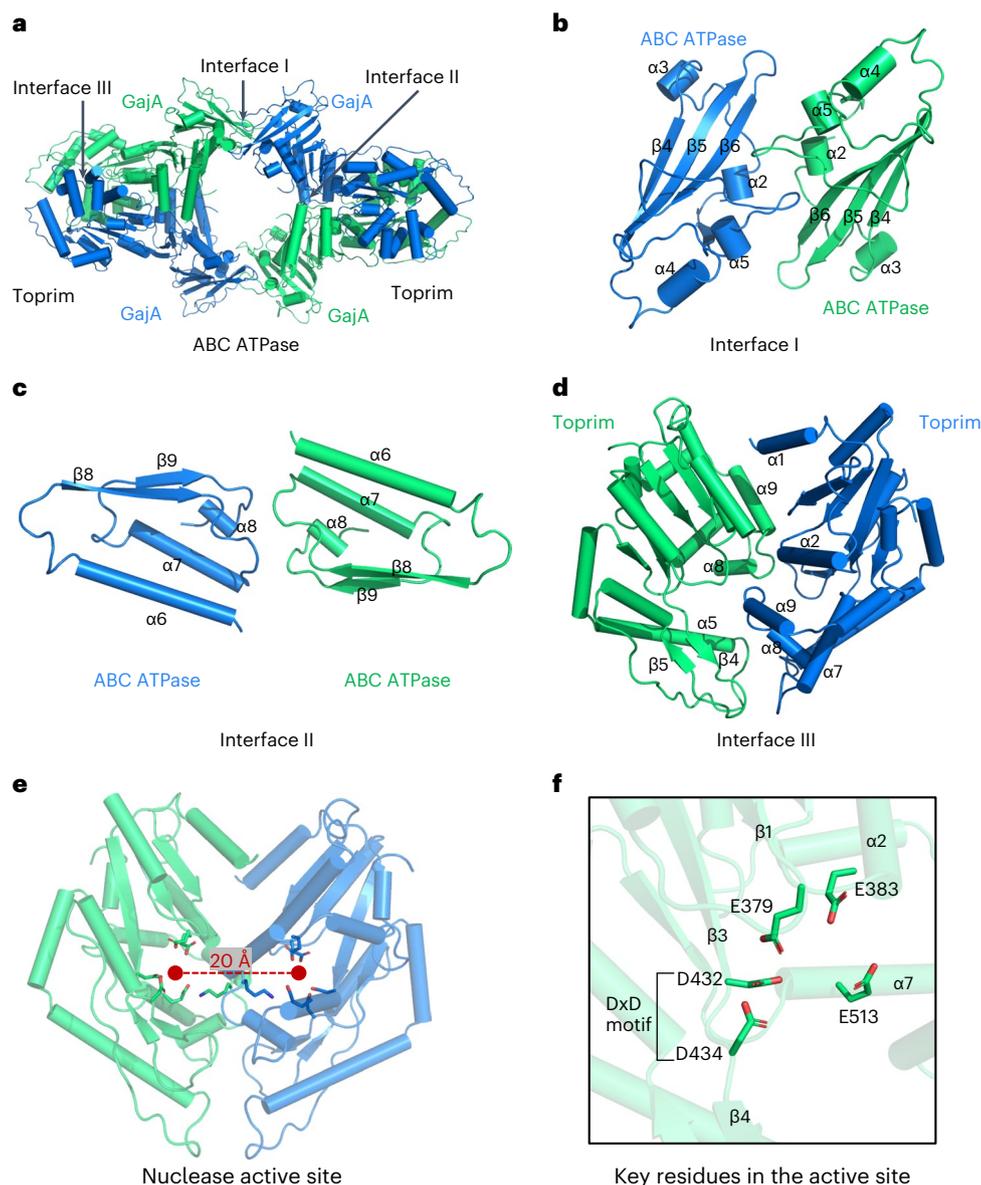


Fig. 2 | Assembly of GajA. **a**, The assembly of tetrameric GajA with three key interfaces indicated, which are denoted as interfaces I, II and III. **b**, Details of interface I mediated by the first halves of ABC ATPase domains with secondary structures indicated. **c**, Details of interface II mediated by the second halves of ABC ATPase domains with secondary structure indicated. **d**, Details of interface

III mediated by the Toprim domains with secondary structure indicated. **e**, The catalytic center of the Toprim domains. The distances between the active sites of dimeric Toprim domains are highlighted. **f**, Key residues in the catalytic center of the Toprim domain that are highlighted in sticks.

while the other magnesium ion may be coordinated by E513 and E383 (refs. 13,16) (Extended Data Fig. 4d). Consistently, the E379A mutation abolished the nuclease activity of GajA, highlighting the notable role of E379 in catalysis⁸.

To further understand how the Toprim domain specifically recognizes and cleaves double-stranded DNA (dsDNA), we generated a structural model of the Toprim domain with dsDNA using RoseTTAFold2NA¹⁷ (Extended Data Fig. 4e). In the predicted model, the cutting site on dsDNA was precisely nested in the catalytic center of the Toprim domain (Extended Data Fig. 4f,g), poised for cleavage. Positively charged residues and hydrophilic residues surrounding the catalytic center established intensive interactions with the dsDNA substrate (Extended Data Fig. 4f,g), which may partially explain the substrate specificity of the Toprim domain. Moreover, the narrow slit between two neighboring Toprim domains in the tetrameric GajA appears to be insufficient to accommodate the dsDNA substrate (Fig. 2d). As

such, we speculated that the Toprim domains would undergo large conformational changes upon DNA binding.

Structure of GajAB complex

To understand the assembly of GajAB, we further reconstituted the GajAB complex and determined a 3.0 Å cryo-EM structure of the complex (Fig. 3a,b, Table 1 and Extended Data Figs. 5a–c and 6a–c). The cryo-EM structure of GajAB reveals a 4:4 assembly of GajA and GajB with a dimension of 175 Å × 145 Å × 95 Å, contrasting with a previous assumption that GajA and GajB form a complex with variable stoichiometries⁹ (Fig. 3a,b). Consistently, gel filtration analysis and native mass spectrometry analysis further confirmed that GajA and GajB assemble into a complex with a stoichiometry of 4:4 while GajB alone exists as a monomer (Extended Data Fig. 5a–c). In the GajAB complex, the tetrameric GajA is decorated by a pair of GajB dimers at either end (Fig. 3a,b).

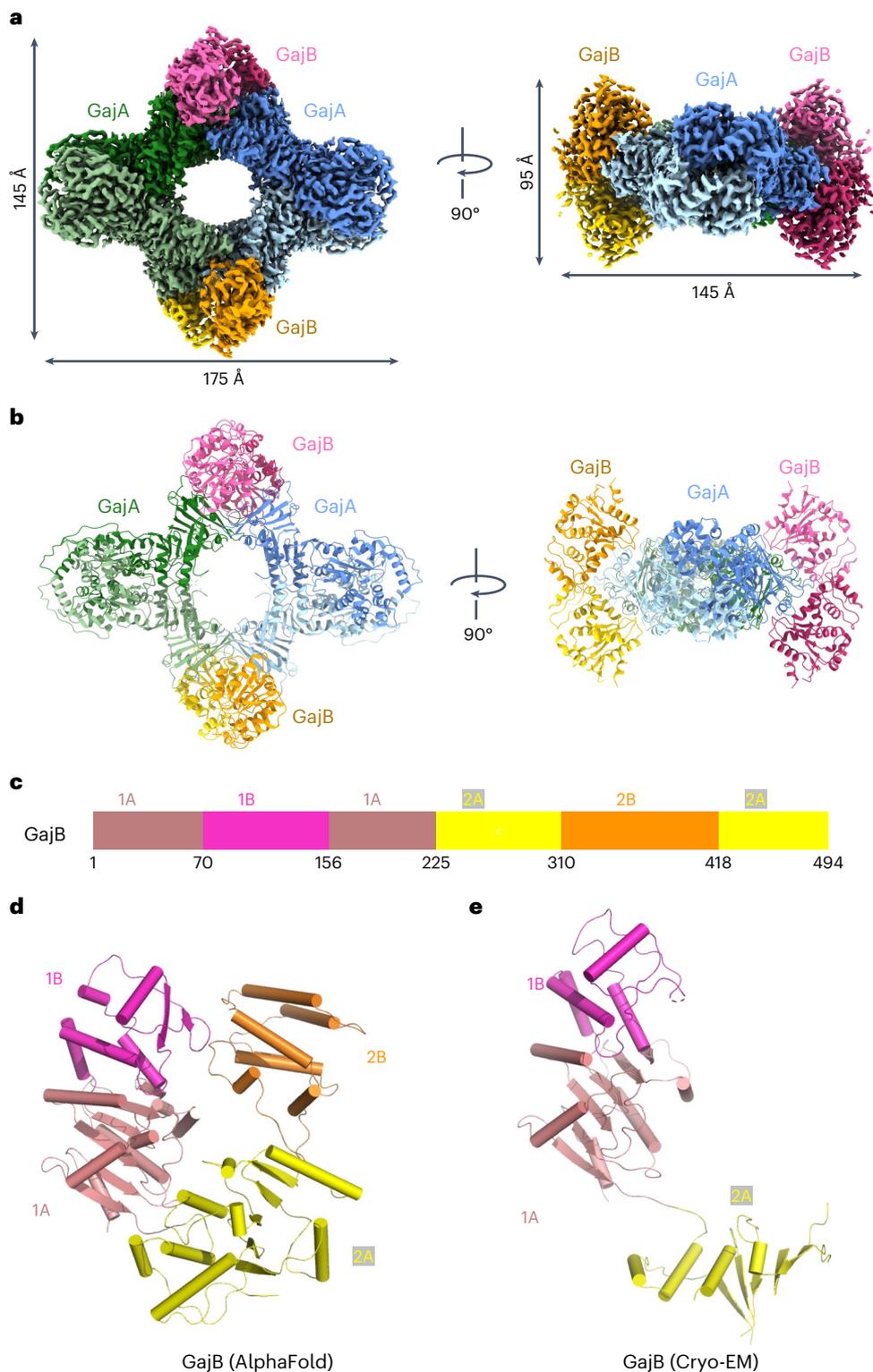


Fig. 3 | Structure of the GajAB complex. a, b, A cryo-EM density map (**a**) and ribbon diagrams (**b**) of the GajAB complex with GajA in cold colors and GajB in warm colors on two different planes: top view (left) and vertical rotation (right). **c,** Domain architecture of GajB. The 1A domain is indicated in pink, 1B domain in

magenta, 2A in yellow and 2B in orange. **d, e,** Ribbon diagram of a GajB protomer predicted by AlphaFold (**d**) or determined by cryo-EM reconstruction (**e**) with domains colored as in **c**.

GajB is composed of four structural domains: 1A, 1B, 2A and 2B, resembling superfamily 1A helicase proteins such as UvrD, PcrA and Rep that function to unwind and translocate DNA^{18–20} (Fig. 3c,d and Extended Data Fig. 7a). In our cryo-EM structure, domains 1A, 1B and part of 2A are visualized, while domain 2B is completely absent (Fig. 3e).

Sequence alignment revealed that GajB contains the eight sequence motifs in domains 1A and 2A of UvrD, which have been identified to be involved in ATP binding²⁰ (Extended Data Fig. 7b). These features indicated that GajB is capable of hydrolyzing ATP. Consistently, ATPase activities are detected in GajB in the presence of DNA⁹, indicating that

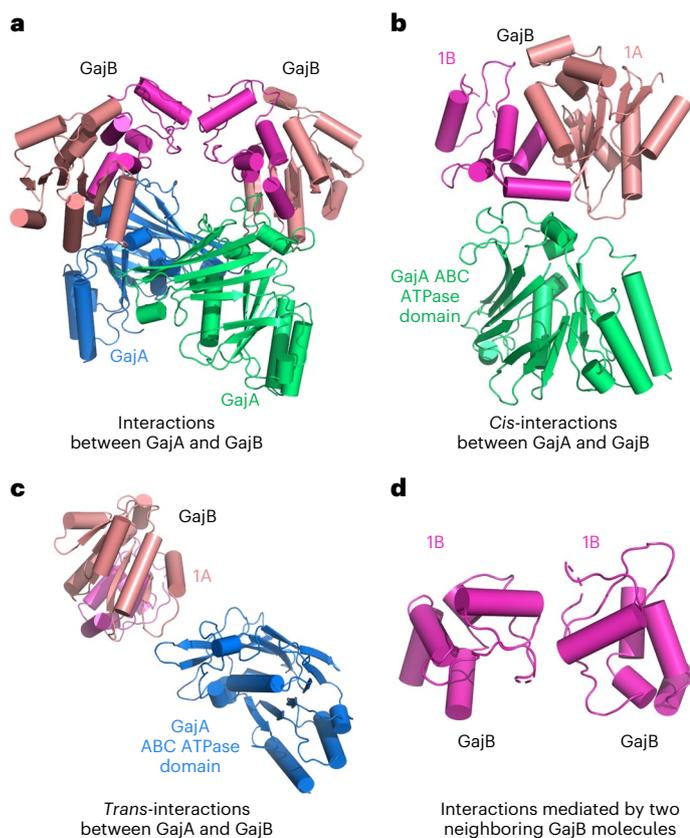


Fig. 4 | Assembly of GajAB. **a**, The assembly of GajAB with a dimeric GajA (green and blue) engaged with two GajB protomers (pink and magenta). **b**, *Cis*-interactions mediated by the GajA ATPase domain and GajB. **c**, *Trans*-interactions mediated by the GajA ATPase domain and GajB. **d**, Interactions between two neighboring GajB protomers, which are mediated by the 1B domain of GajB.

DNA substrates are required to stimulate the ATPase activity of GajB. Structural comparison to UvrD revealed that domain 2A in GajB is not well positioned to coordinate ATP (Extended Data Fig. 7c). As such, conformational changes are required for GajB to bind and hydrolyze ATP upon DNA binding. The structure of UvrD in complex with ds–ss DNA junction revealed that the single-stranded DNA (ssDNA) binds to domains 1A and 2A at their interfaces with 1B and 2B while the DNA duplex is coordinated by domains 1B and 2B²⁰. Comparisons with UvrD demonstrate that GajB contains all the key residues for coordinating ssDNA, suggesting that GajB may use a similar strategy for binding ssDNA (Extended Data Fig. 7d). In contrast, the domain 2B in GajB is much smaller than that of UvrD and lacks key residues for coordinating dsDNA, raising questions whether GajB can efficiently bind to dsDNA (Extended Data Fig. 7a,e).

Mechanism of GajAB assembly

The structure of the GajAB complex revealed that the GajA recruits a pair of GajB molecules via its ATPase domain (Fig. 4a–c). Each GajB molecule interacts with two GajA molecules (*cis* and *trans*) that form a head-to-head dimer via their ATPase domain (Fig. 4b,c). The interaction between GajB and GajA in *cis* is quite extensive with a buried surface area of 755 Å², which is dominated by interactions between the GajB 1B domain and the GajA ATPase domain (Fig. 4b and Extended Data Fig. 8a,b). Both hydrophobic and hydrophilic residues form an extensive network to dock the GajB 1B domain onto the ATPase domain of GajA, positioning the GajB 1A domain adjacent to the GajA ATPase domain with relatively weaker interactions (Extended Data Fig. 8a,b). In contrast, the interactions between GajB and GajA in *trans* are much

weaker with a buried surface area of 140 Å², which is mediated by the GajB 1A domain and the GajA ATPase domain (Fig. 4c and Extended Data Fig. 8c). Additionally, the two paired GajB molecules form relatively weak interactions with each other with a buried surface area of 380 Å², which is dominated by hydrophilic residues in the 1B domain of GajB (Fig. 4d and Extended Data Fig. 8d). Collectively, these three interfaces together with the extensive interactions among tetrameric GajA form the basis of GajAB supramolecular complex assembly.

Nuclease activities and anti-phage defense of GajAB complex

Consistent with previous studies^{8,9}, GajA has nuclease activities in the presence of Mg²⁺ and is capable of cleaving dsDNA while GajB has no nuclease activities (Fig. 5a,b). Unexpectedly, the complex of GajAB displayed similar nuclease activities toward dsDNA as GajA (Fig. 5c). The nuclease activities for both GajA and GajAB can be effectively inhibited by 0.5 mM ATP (Fig. 5a,c), indicating that ATP may serve as a critical factor in regulating the Gabija system. Contrasting a previous study⁸, we found that both GajA and GajAB are capable of cleaving plasmid pUC19 (Fig. 5d,e). Moreover, GajAB displayed higher nuclease activities toward pUC19 than GajA alone, highlighting the importance of GajAB complex assembly in effectively cleaving plasmid DNA (Fig. 5d,e). In addition, we found that GajB alone or the GajAB complex can effectively hydrolyze ATP in the presence of ssDNA, consistent with a recent study⁹. Therefore, ssDNA may activate the ATPase activity of GajB to lower the ATP level in cells for triggering the activation of GajA. Consistently, phage resistance assay showed that the GajAB complex is more effective in anti-phage defense compared to GajA or GajB alone (Fig. 5f). Moreover, we also showed that the nuclease activity of GajAB is essential for anti-phage defense (Fig. 5f). As such, the supramolecular complex assembly of GajAB is critical for anti-phage defense.

Discussion

We find that GajA and GajB assemble into a supramolecular complex for anti-phage defense (Fig. 6). Our structural analysis revealed that GajA alone forms a tetramer and further assembles into a heteromeric octamer with GajB (Fig. 6). Moreover, GajA alone has nuclease activity while GajB adopts a similar fold as UvrD that can bind DNA substrates. No obvious conformational changes have been observed in GajA upon binding to GajB. As such, we propose that GajB may function to assist GajA to better recognize its substrates and thereby promotes the activity of GajA. Consistent with this assumption, we found that the complex of GajAB displays higher nuclease activities toward plasmids than GajA alone (Fig. 5d,e). Given that the nuclease activity of GajA can be inhibited by ATP and GajB has ATPase activity in the presence of ssDNA, it is also possible that GajB may promote the activity of GajA through lowering cellular ATP level under certain conditions. Additionally, the nuclease activity of GajA is inhibited by ATP through a mechanism yet to be revealed. ATP binds to the ATPase domain and stabilizes the oligomerization of GajA. As large conformational changes are coupled with DNA binding, perhaps the presence of ATP may hinder the conformational changes of GajA and thereby prevents the DNA substrate from effectively binding to GajA.

Supramolecular assembly appears to be an emerging theme in anti-phage immune defense. More and more studies have revealed that bacterial immune systems tend to form large complexes for anti-phage defense^{21–24}. For example, RdrA and RdrB in the restriction by an adenosine deaminase acting on RNA (RADAR) system assemble into a giant assembly with a molecular weight of up to 10 MDa (refs. 21,22). Here, we present another example to show the supramolecular assembly by the Gabija system. As both the RADAR system and the Gabija system contain and oligomerize via ATPase domains, we believe that other ATPase-containing bacterial immune systems may also assemble into supramolecular complexes for anti-phage defense.

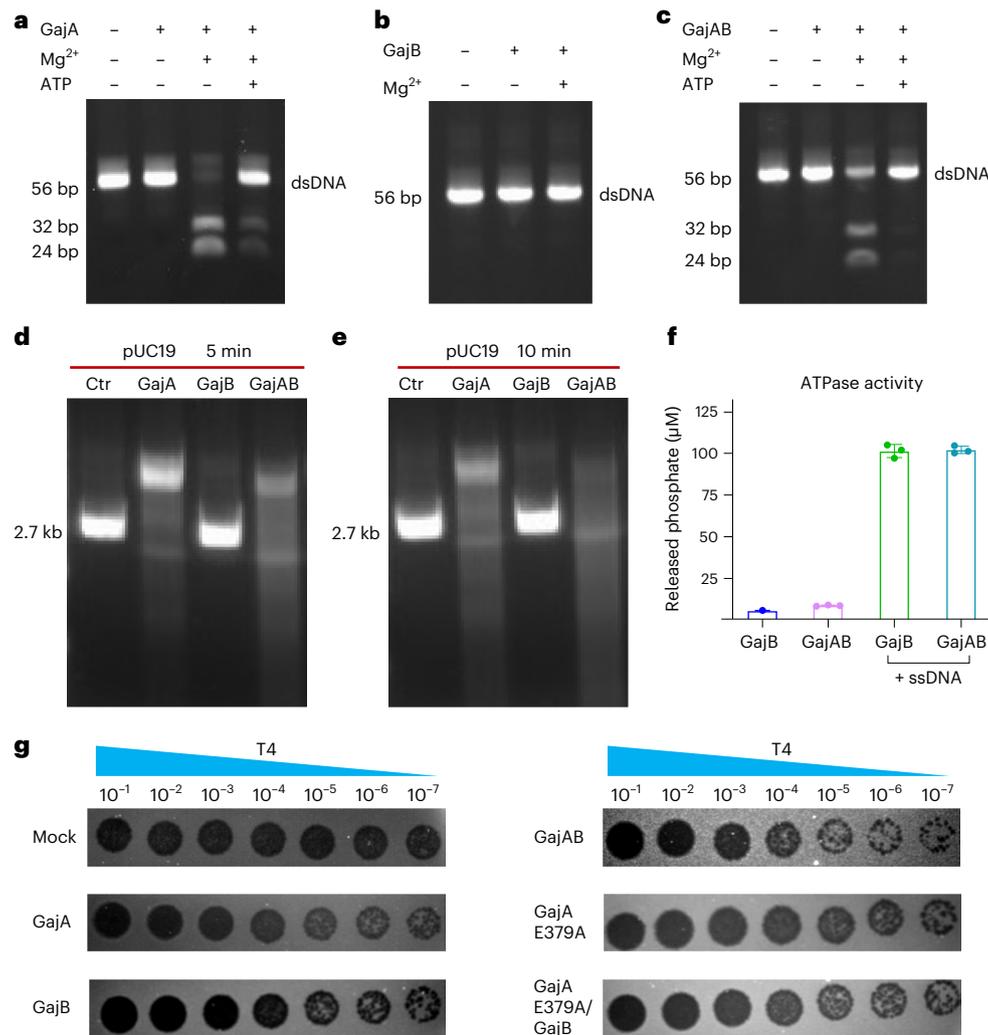


Fig. 5 | Anti-phage defense of GajAB. **a**, Nuclease assays were used to test the activity of GajA. GajA can process dsDNA in the presence of magnesium, which can be inhibited by ATP. **b**, Nuclease assays were used to test whether GajB can cleave dsDNA. dsDNA cannot be processed by GajB. **c**, Nuclease assays were employed to test the activity of GajAB. GajAB cleaves dsDNA in the presence of magnesium, which can be inhibited by ATP. **d, e**, Nuclease assays were used to test the activity of GajA, GajB and GajAB using pUC19 plasmids as substrates. pUC19 plasmids were processed by GajA, GajB and GajAB for 5 min (**d**) and 10 min (**e**) at

room temperature, respectively. GajAB displayed higher activities than GajA, underscoring the importance of GajB in promoting the catalytic activity of GajA. All the cleavage assays mentioned above (**a–e**) were repeated at least three times. **f**, GajB alone or the GajAB complex displays ATPase activity in the presence of ssDNA. The histograms correspond to the mean of three independent experiments; the error bars represent the s.d. **g**, Phage plaque assays were employed to test the anti-phage effects of GajA, GajB, GajAB and their mutants. Anti-phage defense of GajA, GajB, GajAB, GajA_E379A mutant and the complex of GajA_E379A and GajB.

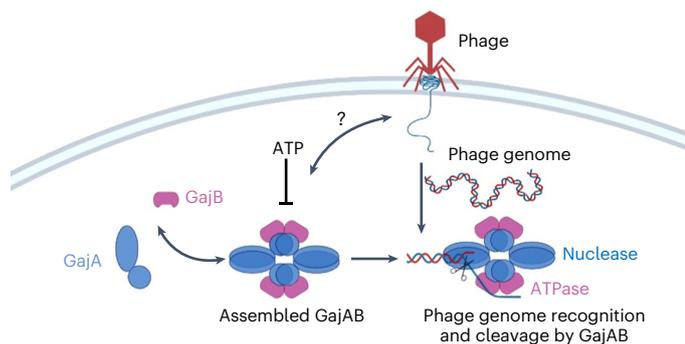


Fig. 6 | Mechanisms of GajAB assembly and function. A schematic diagram to illustrate mechanisms of GajAB assembly and function.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-024-01283-w>.

References

1. Tal, N. & Sorek, R. SnapShot: bacterial immunity. *Cell* **185**, 578–578 e571 (2022).
2. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR–Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
3. Duncan-Lowey, B. & Kranzusch, P. J. CBASS phage defense and evolution of antiviral nucleotide signaling. *Curr. Opin. Immunol.* **74**, 156–163 (2022).
4. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* <https://doi.org/10.1126/science.aar4120> (2018).
5. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).

6. Millman, A. et al. An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556–1569 e1555 (2022).
7. Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
8. Cheng, R. et al. A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res.* **49**, 5216–5229 (2021).
9. Cheng, R. et al. Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host Microbe* <https://doi.org/10.1016/j.chom.2023.06.014> (2023).
10. Liu, Y. et al. ATP-dependent DNA binding, unwinding, and resection by the Mre11/Rad50 complex. *EMBO J.* **35**, 743–758 (2016).
11. Hopfner, K. P. et al. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* **101**, 789–800 (2000).
12. Aravind, L., Leipe, D. D. & Koonin, E. V. Toprim—a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res.* **26**, 4205–4213 (1998).
13. Schiltz, C. J., Lee, A., Partlow, E. A., Hosford, C. J. & Chappie, J. S. Structural characterization of Class 2 OLD family nucleases supports a two-metal catalysis mechanism for cleavage. *Nucleic Acids Res.* **47**, 9448–9463 (2019).
14. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
15. Antine, S. P. et al. Structural basis of Gabija anti-phage defence and viral immune evasion. *Nature* <https://doi.org/10.1038/s41586-023-06855-2> (2023).
16. Schiltz, C. J., Adams, M. C. & Chappie, J. S. The full-length structure of *Thermus scotoductus* OLD defines the ATP hydrolysis properties and catalytic mechanism of Class 1 OLD family nucleases. *Nucleic Acids Res.* **48**, 2762–2776 (2020).
17. Baek, M. et al. Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods* <https://doi.org/10.1038/s41592-023-02086-5> (2023).
18. Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. & Waksman, G. Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635–647 (1997).
19. Velankar, S. S., Soutanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999).
20. Lee, J. Y. & Yang, W. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127**, 1349–1360 (2006).
21. Duncan-Lowey, B. et al. Cryo-EM structure of the RADAR supramolecular anti-phage defense complex. *Cell* **186**, 987–998 e915 (2023).
22. Gao, Y. et al. Molecular basis of RADAR anti-phage supramolecular assemblies. *Cell* **186**, 999–1012 e1020 (2023).
23. Gao, L. A. et al. Prokaryotic innate immunity through pattern recognition of conserved viral proteins. *Science* **377**, eabm4096 (2022).
24. Shen, Z., Lin, Q., Yang, X. Y., Fosuah, E. & Fu, T. M. Assembly-mediated activation of the SIR2-HerA supramolecular complex for anti-phage defense. *Mol. Cell* <https://doi.org/10.1016/j.molcel.2023.11.007> (2023).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2024

Methods

Molecular cloning, protein expression and purification

B. cereus GajA (UniProt: [J8H9C1](#)) with an N-terminal His \times 6-tag was cloned into the pET28a vector. *B. cereus* GajB (UniProt: [J8HQ06](#)) was inserted into the pETDuet-1 vector with an N-terminal His \times 6 tag. All the mutants were generated through site-directed mutagenesis.

Recombinant plasmids for protein expression were transformed into *E. coli* BL21 (DE3) cells (Thermo Fisher Scientific) that were cultured in Luria-Bertani (LB) medium containing 50 μ g ml⁻¹ kanamycin at 37 °C. When an OD₆₀₀ of 0.6–0.8 was reached, protein expression was induced at 18 °C by 0.3 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). Cells were collected after overnight induction (~16 h) and resuspended in lysis buffer (50 mM Tris–HCl pH 8.0, 500 mM NaCl and 10 mM imidazole). After sonication, the supernatant of lysate was collected through centrifugation at 30,000g, 4 °C for 50 min. The clarified lysate was loaded onto a pre-equilibrated Ni²⁺-NTA agarose column, and then the column was washed with 30 column volumes of Ni²⁺-NTA wash buffer (50 mM Tris–HCl pH 8.0, 500 mM NaCl and 10 mM imidazole). The protein was eluted in elution buffer (50 mM Tris–HCl pH 8.0, 150 mM NaCl, 250 mM imidazole and 0.4 mM tris(2-carboxyethyl) phosphine). Protein was further purified by size exclusion chromatography using gel-filtration column (Superose 6 increase 10/300 GL, Cytiva, Sigma-Aldrich) in a buffer containing 50 mM Tris–HCl pH 8.0, 150 mM NaCl and 0.4 mM tris(2-carboxyethyl)phosphine.

For the assembly of the GajAB complex, we incubated GajA and GajB with molar ratio of 1:1 on ice for 1 h followed by further purification via gel filtration.

Cryo-EM data collection

Three microliters of sample at 1.8 mg ml⁻¹ was applied to a glow-discharged Quantifoil R1.2/1.3 400 mesh gold grid (Electron Microscopy Sciences), blotted for 4 s in 100% humidity at 4 °C and plunged into liquid ethane using an FEI Vitrobot Mark IV (Thermo Fisher). All grids were screened using a Thermo Fisher Glacios microscope (OSU Center for Electron Microscopy and Analysis).

For GajA tetramer (4 GajA) in thicker ice, 1,364 micrographs were collected using a 300 kV Titan Krios microscope equipped with a K3 direct electron detector (Thermo Fisher) in counting mode with a nominal magnification of 81,000 \times , and a physical pixel size of 0.899 Å with defocus values ranging from –1.0 to –2.0 μ m. For GajA in thin ice, 6,370 images were collected using similar parameters.

For GajAB hetero-complex (4 GajA:4 GajB), 7,173 micrographs were collected using a K3 detector with physical pixel size of 1.12 Å. Each micrograph stack contains 40 frames with a total electron dose of 50 e⁻ Å⁻² s⁻¹.

Cryo-EM data processing

The detailed flowcharts for data processing of all the datasets are illustrated (Extended Data Figs. 1e, 2c and 5c). The datasets were imported into cryoSPARC (v4.1.1) implementation of patch motion correction and patch contrast transfer function estimation²⁵. Initial particle picking was done by blob picking to generate initial two-dimensional (2D) classes. Representative 2D classes were then selected as templates to pick all the particles for reconstruction.

For GajA tetramer, 849,640 particles were picked and extracted. After two rounds of 2D classification, 583,860 particles were selected and merged for further three-dimensional (3D) reconstruction and heterogeneous refinement. The best class of 96,633 particles was selected for further nonuniform refinement with D2 symmetry, resulting in a 3.23 Å map.

For GajAB complex, 6,928,153 particles were picked and extracted. After two rounds of 2D classification, 2,682,203 particles were selected for further ab initio reconstruction to generate three initial models for further refinement. The best class was selected for further 3D classification and heterogeneous refinement. The final best class of

942,091 particles was selected for nonuniform refinement with C1 symmetry and D2 symmetry, resulting in a 2.98 Å map and a 2.79 Å map, respectively.

All reported resolutions were estimated on the basis of the gold-standard Fourier shell correlation (FSC) 0.143 criterion²⁶.

Model building and refinement

Two initial models of GajA and GajB were predicted by AlphaFold¹⁴ and fit into the cryo-EM maps of GajA tetramer or GajAB complex using Chimera²⁷. Manual adjustments were done using Coot to yield the final atomic model²⁸. Real-space refinement was performed to refine the model against cryo-EM density map with secondary structure and geometry restraints in PHENIX²⁹. The all-atom contacts and geometry for the final models were validated by Molprobity³⁰. All the structural figures were generated using PyMOL³¹, Chimera²⁷ and ChimeraX³².

Nuclease assays

For pUC19 plasmid, a final concentration of 400 nM protein and 1,000 ng plasmid substrate were incubated in reaction buffer (20 mM Tris–HCl pH 8.0) at 37 °C for 10 min. Samples were separated in 1% (w/v) agarose gel in 1 \times tris acetate EDTA (TAE) buffer for 30 min.

For dsDNA substrate, we annealed 5'-TTTTTTTTTTTTTTTTT TAATAACCCGGTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3' and 5'-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAACCCGGTTAT TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3' to generate a dsDNA substrate. Then, 400 nM protein and 800 nM nucleic acid substrate were incubated in reaction buffer (20 mM Tris–HCl pH 8.0) at 37 °C for 10 min. Samples were separated in 12% polyacrylamide gel electrophoresis gel in 1 \times tris borate EDTA (TBE) buffer for 1 h. In each reaction, a final concentration of 5 mM MgCl₂ or 0.5 mM ATP was added as indicated. All gels were stained by running buffer containing 0.5 ng μ l⁻¹ ethidium bromide and visualized by Imaging System (Bio-Rad).

ATPase assays

The ATPase activity was measured by an ATPase/GTPase Assay Kit (Sigma-Aldrich, #113CB04A30). GajB or GajAB in the absence or presence of ssDNA (5'-TTTTTTTTT TTTTTTTAATAACCCGGTTATTTTTT TTTTTTTTTTTTTTTTTT-3') was incubated with ATP in the buffer of 20 mM Tris pH 8.0, 75 mM KCl and 2 mM MgCl₂ at 37 °C. The reaction products were monitored by BioTek Synergy HT microplate reader at 600 nm. ATPase activities were calculated on the basis of the released phosphate.

Structural prediction

We utilized RoseTTAFold2NA¹⁷, a deep learning technique, to predict the complex between the GajA protein and dsDNA. This method relies solely on protein and DNA sequences for predicting complex structures. We used the GajA Toprim domain (residues 372–574) and a dsDNA segment (5'-GAATAACCCGGATATT-3' and 5'-AATATCCGGGT TATTC-3') as the input of the method. To predict the protein's structure, 3,361 homologous sequences of the GajA top domain were identified. RoseTTAFoldNA also offers a confidence measure for its predictions, known as predicted aligned error (PAE). PAE estimates the positional discrepancy in angstroms between predicted and actual structures, calculated for each pair of residues. In the predicted model, the average PAE between residues of the protein and residues of DNA was 16.8, indicating high confidence in the predicted complex's structure.

Plaque assays

Plaque assays were performed as previously described^{4,33}. Briefly, reconstructed plasmid was transformed into *E. coli* DE3 competent cell. A single bacterial colony was picked from a fresh LB agar plate and grown in LB broth containing antibiotic at 37 °C to an OD₆₀₀ of ~0.4. Protein expression was induced by the addition of 0.2 mM IPTG. After further growth for ~3 h, 500 μ l of the bacterial cultures was mixed with

14.5 ml of 0.5% LB top agar, and the entire samples were poured onto LB plates containing antibiotic and IPTG (0.1 mM). Plates were spotted with 4 µl of the T4 phage diluted in LB at eight tenfold dilutions, namely 10⁰ to 10⁻⁷. Lysate titer was determined using the small drop plaque assay method as previously described^{33,34}. Plates were incubated at 37 °C overnight and then imaged.

Native mass spectrometry analysis

The sample underwent online buffer exchange before mass spectrometry analysis. A mobile phase of 200 mM ammonium acetate at pH 7.0 was maintained at a flow rate of 50 µl min⁻¹. A Vanquish Duo Ultra-High-Performance LC system (Thermo Scientific) equipped with a dual pump and autosampler was used to load protein samples without further purification into a MAbPac size exclusion column (MAbPac SEC1, 2.1 × 300 mm, 300 Å, 5 µm; Thermo Scientific, 008789). One microgram of the sample was injected onto the column and was further tracked by ultraviolet absorbance. From 0 min to 16.8 min, the flow was diverted from the column to the mass spectrometer, while after this time (16.8 min to 30 min), flow was diverted to the waste. The timing of this switch was optimized to allow for buffer-exchanged protein samples to make it to the mass spectrometer, while salts from the original buffer were passed to the waste. Buffer-exchanged samples were then analyzed using a Q Exactive Ultra-High Mass Range Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific). The following settings were used for mass spectrometry analysis: the scan range was set to 1,000–20,000. The resolution was set to 12,500. The number of microscans was set to 5. The maximum injection time was set to 200 ms. The sheath gas was set to 60 psi. The auxiliary and sweep gas were set to 0 psi. Spray voltage was set to 3.85 kV in positive ion mode. The capillary temperature was set to 275 °C. S-lens RF level was set to 200 V. The in-source trapping was set to -120 V. The in-source dissociation was set to 5 V. The source d.c. offset was set to 21 V. The injection flatapole d.c. was set to 5 V. The inter-flatapole lens was set to 4 V. The bent flatapole d.c. was set to 2 V. The tapping gas pressure was set to 7. The total run time for the method was 30 min. The mass spectrum was deconvolved using Unidec³⁵.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Accession numbers for Gabija A tetramer (4A), Gabija AB complex 1 (4A:4B, C1 symmetry) and Gabija AB complex 2 (4A:4B, D2 symmetry) are as follows: coordinates of atomic models, [8TKO](#), [8TK1](#) and [8TJY](#), deposited to Protein Data Bank, and density map. [EMD-41319](#), [EMD-41321](#) and [EMD-41314](#), deposited to Electron Microscopy Data Bank. All data needed to evaluate the conclusions are present in the paper. Source data are provided with this paper.

References

- Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
- Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
- Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
- The PyMOL Molecular Graphics System, Version 3.0 (Schrödinger, 2022)
- Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
- Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages using the small drop plaque assay system. *Methods Mol. Biol.* **501**, 81–85 (2009).
- Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages by double agar overlay plaque assay. *Methods Mol. Biol.* **501**, 69–76 (2009).
- Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

Acknowledgements

Cryo-EM data of GajA were collected at OSU CEMAS with the assistance of G. Grandinetti and Y. Narui. Cryo-EM data of GajAB were collected with the assistance of A. D. Wier, T. J. Edwards, T. Fox and J. Wang at the National Cancer Institute Cryo-Electron Microscopy Center supported by grants from the NIH National Institute of General Medical Sciences (GM103310). T.-M.F. is supported by an R35 grant from National Institute of General Medical Sciences (1R35GM147465-01).

Author contributions

T.-M.F. conceived the project. X.-Y.Y., Z.S. and J.X. performed molecular cloning, biochemical purification, ATPase assay and plaque assays and determined nuclease activity. Z.S. prepared EM grids and determined the cryo-EM structures. Q.L. helped on structural reconstruction. Z.S. and X.-Y.Y. built the models. W.X. predicted the structure of the GajA–dsDNA complex. J.G. and I.M. performed native mass spectrometry analysis under the supervision of V.H.W. Z.S., X.-Y.Y. and T.-M.F. analyzed all the data together. T.-M.F. wrote the manuscript with inputs from all the authors.

Competing interests

The authors declare no competing interests.

Additional information

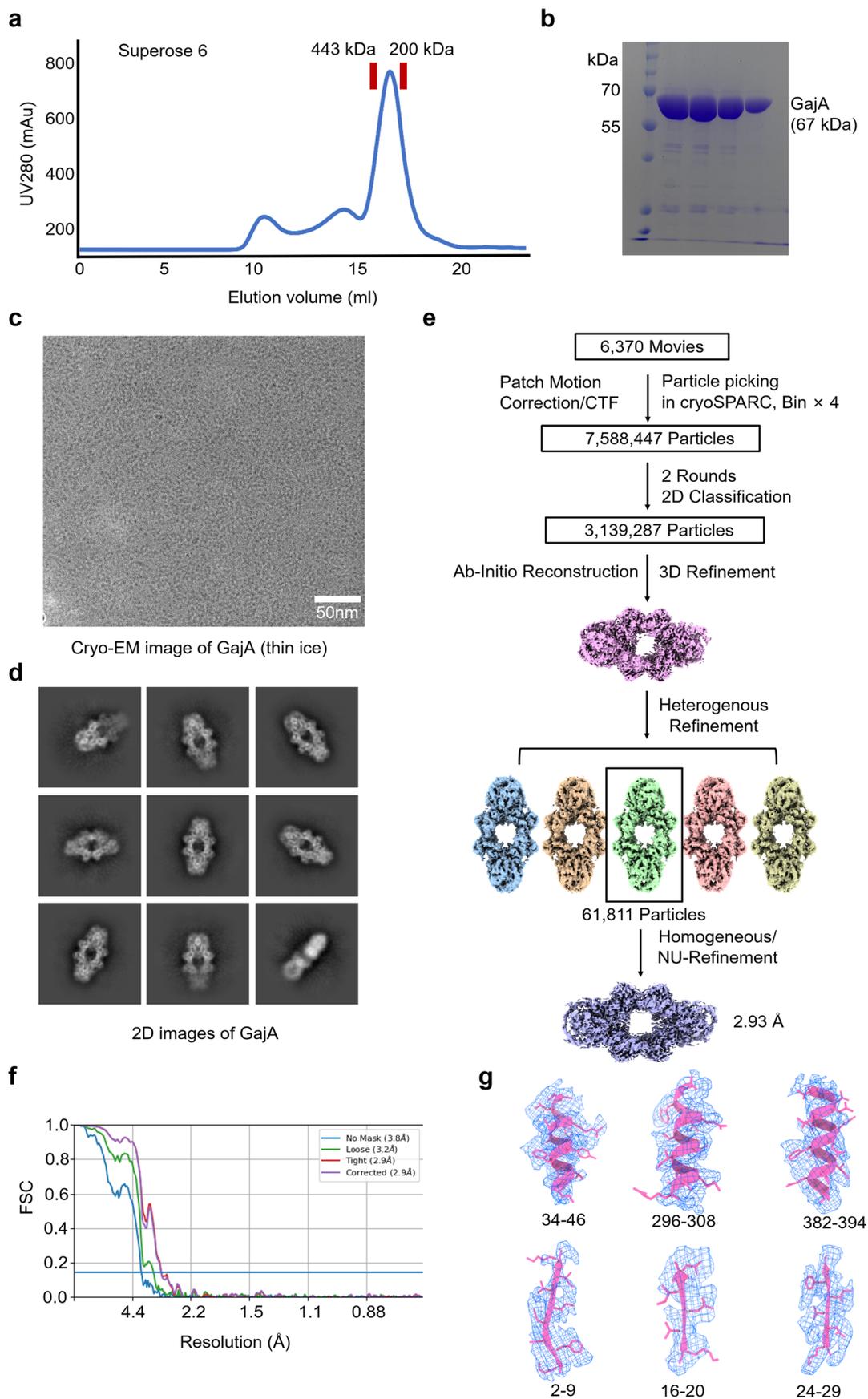
Extended data is available for this paper at <https://doi.org/10.1038/s41594-024-01283-w>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41594-024-01283-w>.

Correspondence and requests for materials should be addressed to Tian-Min Fu.

Peer review information *Nature Structural & Molecular Biology* thanks Qian Yin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Sara Osman and Dimitris Typas were the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

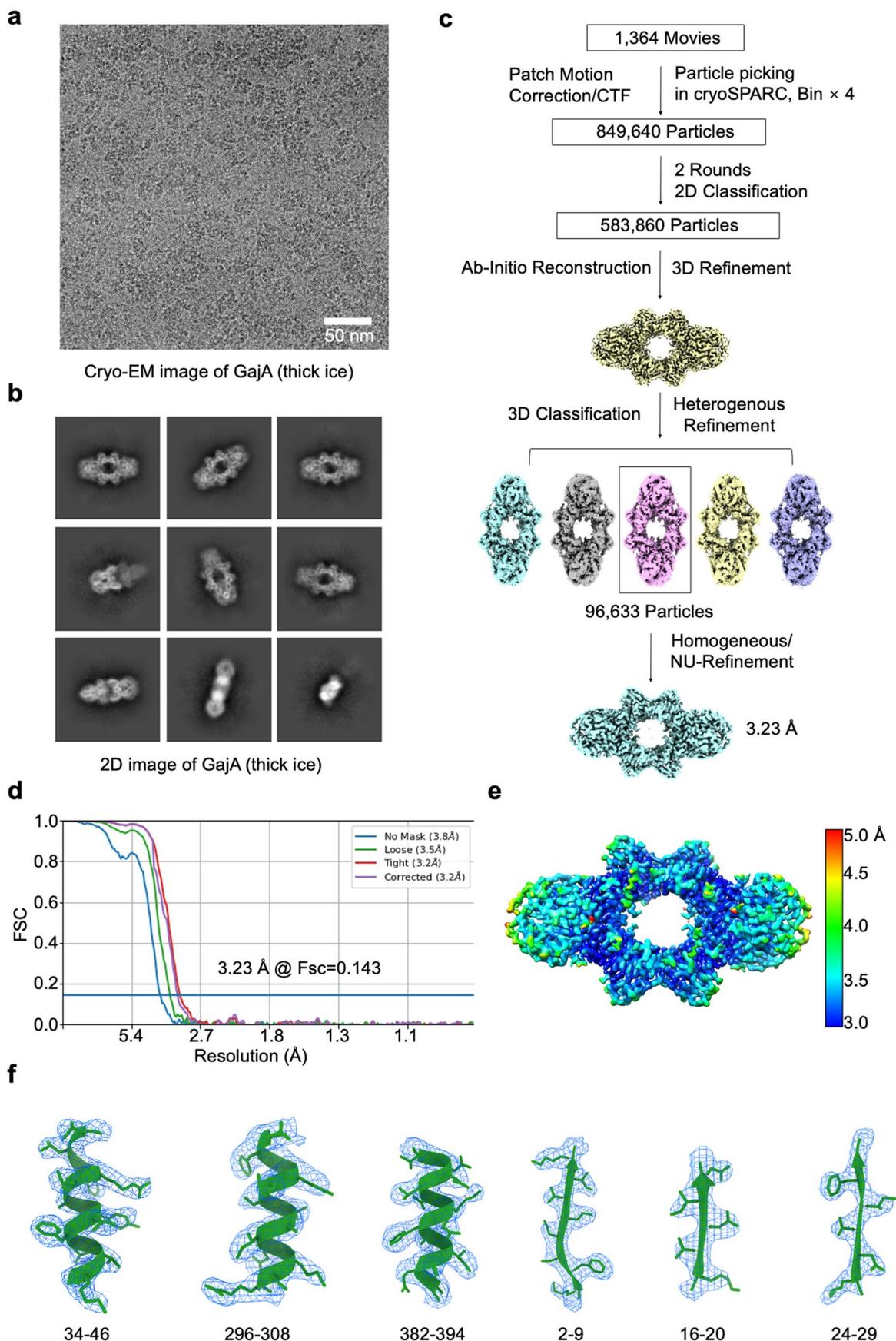
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | See next page for caption.

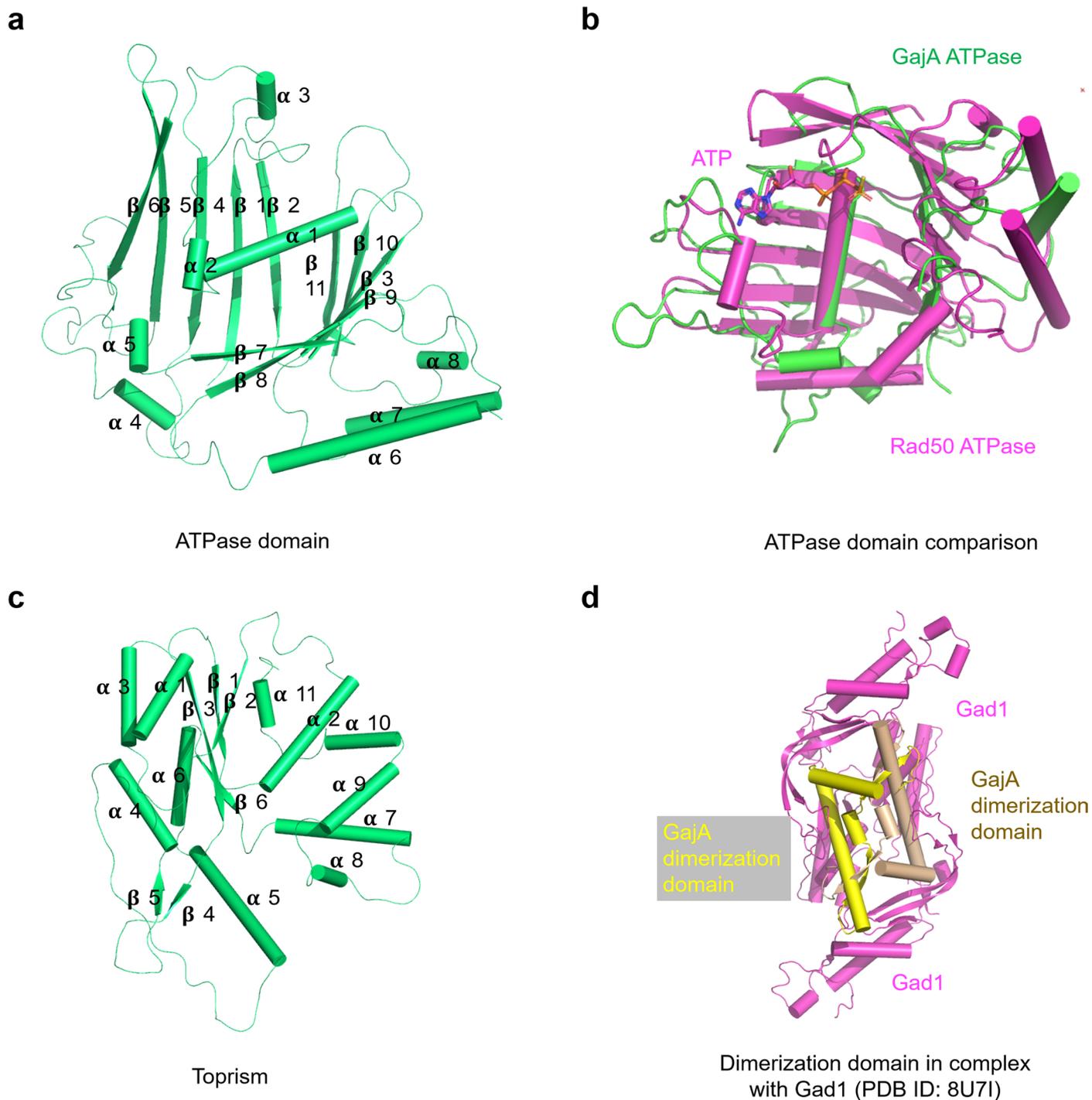
Extended Data Fig. 1 | Cryo-EM reconstruction of GajA in thin ice. a, b, Gel filtration profile (a) and SDS-PAGE gel (b) of GajA purification. The experiment was replicated at least three times. **c**, A representative cryo-EM image of GajA in thin ice. Thousands of images were collected. **d**, Representative 2D class averages of GajA calculated from thin-ice cryo-EM images. **e**, Data processing

workflow for 3D reconstruction of GajA tetramer from thin-ice cryo-EM images. **f**, The FSC curve of reconstructed GajA tetramer from thin-ice cryo-EM images. **g**, Representative cryo-EM density of GajA tetramer fitted with α -helices and β -strands. The density map was shown at a contour level of 0.03.

**Extended Data Fig. 2 | Cryo-EM reconstruction of GajA in thicker ice.**

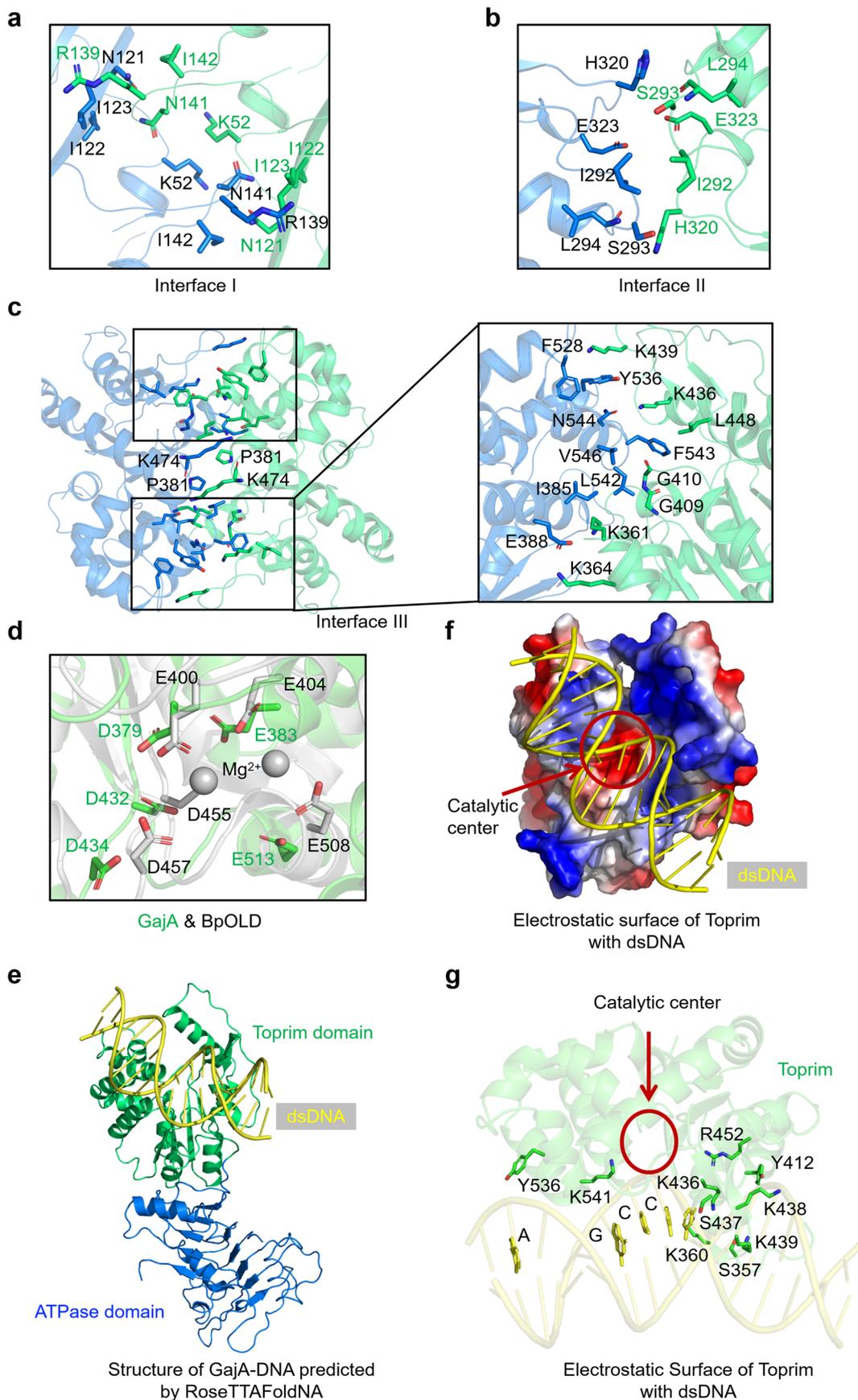
a, A representative cryo-EM image of GajA in thicker ice. Thousands of images were collected. **b**, 2D class averages of GajA calculated from thick-ice cryo-EM images. **c**, Data processing workflow for 3D reconstruction of GajA tetramer from

thick-ice cryo-EM images. **d**, The FSC curve of reconstructed GajA tetramer from thick-ice cryo-EM images. **e**, Local resolution of reconstructed GajA tetramer from thick-ice cryo-EM images. **f**, Cryo-EM density of GajA tetramer fit with α -helices and β -strands. The density map is shown at contour levels of 0.03.



Extended Data Fig. 3 | Architecture of GajA. **a**, Ribbon diagram of GajA N-terminal ATPase domain with secondary structures indicated. **b**, Overlaid structures of GajA N-terminal ATPase domain (green) and Rad50 ATPase domain

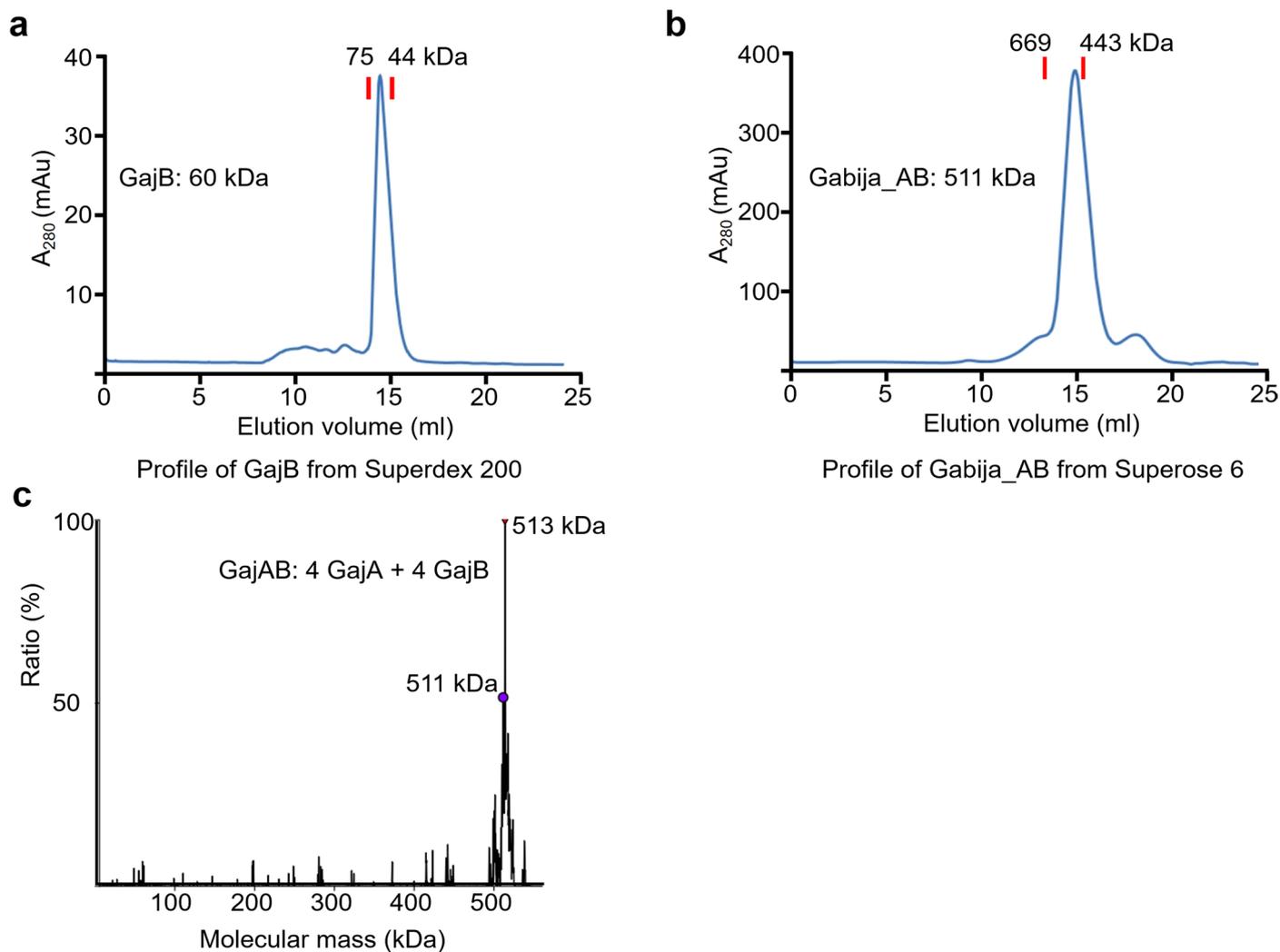
(PDB ID *SDNY*, magenta). **c**, Ribbon diagram of GajA C-terminal Toprim domain with secondary structures indicated. **d**, The dimerization domain of GajA in complex with a phage protein Gad1 (magenta).



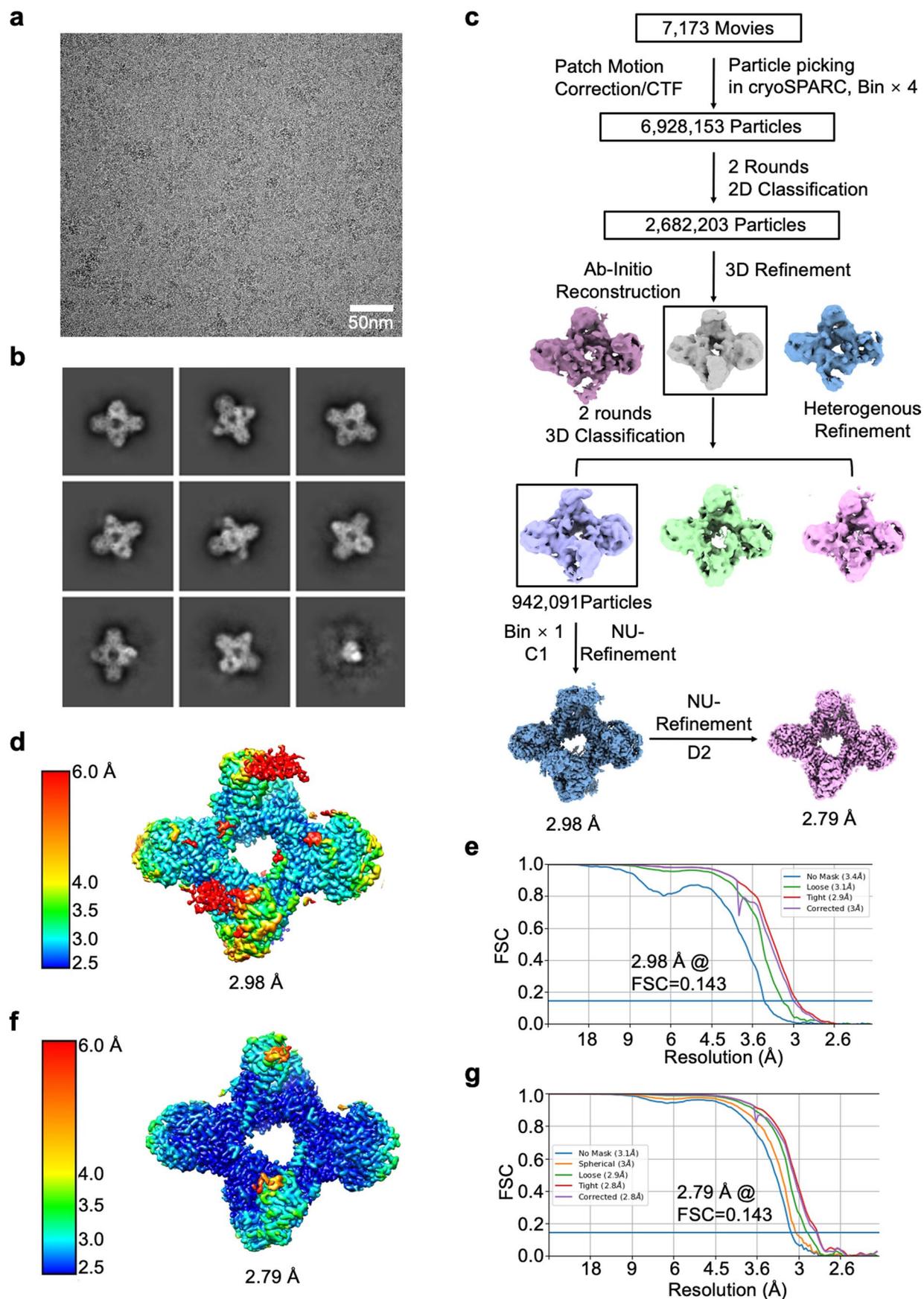
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Interfaces in GajA tetramer. **a–c**, Enlarged views of interface I (a), interface II (b), and interface III (c) in GajA tetramer. Key residues on the interfaces were highlighted in sticks. **d**, Superimposed structures of the active sites from GajA (green) and BpOLD (PDB ID [6NK8](#), gray). **e**, Structure of GajA in complex with dsDNA (Yellow) that was predicted by RoseTTAFoldNA.

f, Electrostatic surface representation of GajA with dsDNA. The catalytic center of GajA is highlighted by a red circle. Negatively charged residues surrounding the catalytic center of GajA coordinate dsDNA. **g**, Key residues involved in coordinating dsDNA are highlighted in sticks.

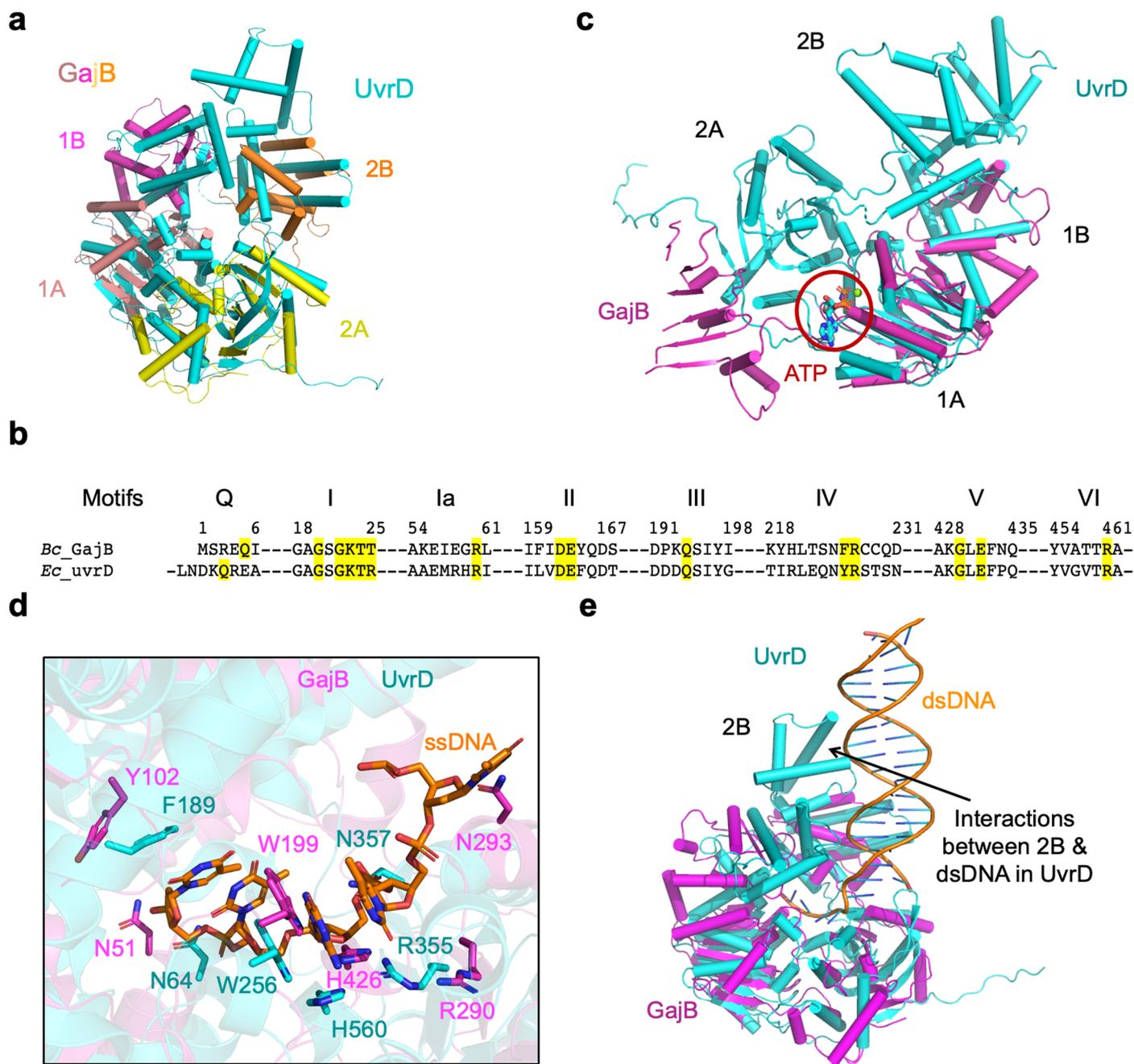


Extended Data Fig. 5 | Oligomerization state of GajB and GajAB. **a**, Gel filtration profile of GajB indicates that GajB alone assembles as a monomer. **b**, Gel filtration profile of GajAB indicates that GajAB assembles as a tetramer of heterodimer. **c**, Native mass spectrometry analysis revealed that there are four copies of GajA and four copies of GajB in the GajAB complex.



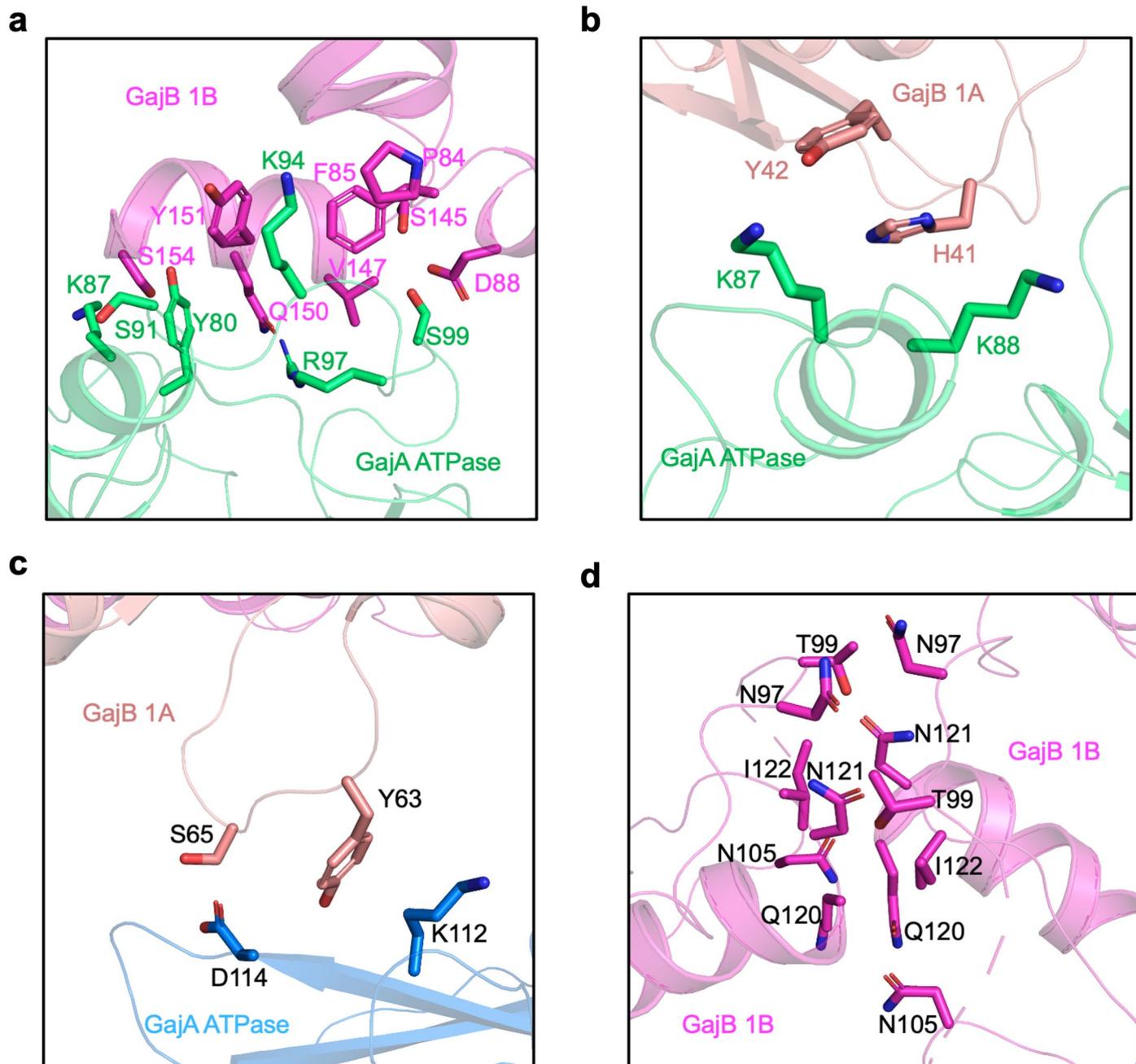
Extended Data Fig. 6 | Cryo-EM reconstruction of GajAB. **a**, A representative cryo-EM image of GajAB complex. Thousands of images were collected. **b**, 2D class averages of GajAB complex. **c**, Data processing workflow for 3D

reconstruction of GajAB complex. **d**, **e**, Local resolution (**d**) and FSC curve (**e**) of reconstructed GajAB complex without symmetry setting. **f**, **g**, Local resolution (**f**) and FSC curve (**g**) of reconstructed GajAB complex with D2 symmetry setting.



Extended Data Fig. 7 | Structural comparison of GajB and UvrD. **a**, Overlaid structures of GajB (magenta, pink, yellow, and orange, AlphaFold predicted structure) and UvrD (PDB ID 2IS2, cyan). **b**, Sequence alignment of ATP binding motifs between GajB and UvrD. **c**, Overlaid structures of GajB (magenta) and UvrD (cyan) showed that domain 2A of GajB is not well positioned to coordinate

ATP. **d**, Expanded view of key residues involved in coordinating ssDNA from GajB (magenta, AlphaFold predicted structure) and UvrD (cyan). **e**, Overlaid structures of GajB (magenta, AlphaFold predicted structure) and UvrD (cyan) revealed that domain 2B in GajB lacks key motifs for coordinating dsDNA.



Extended Data Fig. 8 | Interfaces in GajAB. **a**, Key residues mediating interactions between GajB domain 1B (magenta) and GajA ATPase domain (green). **b**, Key residues mediating *cis*-interactions between GajB domain

1A (pink) and GajA ATPase domain (green). **c**, Key residues mediating *trans*-interactions between GajB 1A (pink) and GajA ATPase (blue). **d**, Key residues mediating interactions of two neighboring GajB protomers.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection EPU v2.11.1 for cryo-EM data collection.

Data analysis We used cryoSPARC v4.1.1, Coot 0.9.8.7, PHENIX 1.20-4487, PyMOL 2.5.3, Molprobity 4.02-528, UCSF Chimera 1.15 and UCSF ChimeraX 1.5 for cryo-EM data analysis. We used RoseTTAFold2NA to predict the structure of the GajA and DNA complex.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Accession numbers for Gabija A tetramer (4A), Gabija AB complex 1 (4A:4B, C1 symmetry), and Gabija AB complex 2 (4A:4B, D2 symmetry) are as follows: (coordinates of atomic models: 8TK0, 8TK1, and 8TJY deposited to Protein Data Bank), and (density map: EMDB-41319, EMDB-41321, EMDB-41314, deposited to Electron Microscopy Data Bank). All data needed to evaluate the conclusions are present in the paper.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No any sample sizes or statistical analyses were pre-determined. In the cryo-EM datasets, each datasets contains thousands of images, from which millions of particles were selected for final reconstruction. For our biochemical experiments, no information was derived from a population based sampling.
Data exclusions	In cryo-EM data processing, bad images with CTF outliers (ice contamination or crushed) were removed form further analysis. We also discarded "junk" particles that could not be classified into useful 3D refinemet during 2D or 3D classifications. Those procedures are widely used practice in the cryo-EM field. No other data were excluded from analysis.
Replication	All biochemical experiments were performed independently at least three times with the similar result.
Randomization	For cryo-EM maps determination, FSCs were calculated using two independent havles of the datasets, into which the particles were randomly allocated. Randomization is less relevant to other in vitro biochemical analysis. In vitro biochemical experiments were conducted in a highly controlled environment, where variables such as temperature, pH, concentration, and other experimental conditions can be precisely regulated. This controlled setting reduces the potential for external factors to influence the outcome of the experiment, mitigating the need for randomization to balance out confounding variables.
Blinding	Blinding was not necessary for cryo-EM analysis as well as in vitro biological experiments, including cleavage assay, ATPase measurement, Plaque assay and Native Mass Spectrometry Analysis. All the in vitro biological experiments involve standardized protocols and procedures that are followed systematically by laboratory personnel. These protocols are designed to minimize variability and ensure consistency in experimental conditions, reducing the likelihood of bias introduced by differences in experimental handling or interpretation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging