Article

Structures and activation mechanism of the Gabija anti-phage system

Jing Li[1,2,3,4,9], Rui Cheng[5,9], Zhiming Wang[1,2,3,4,9], Wuliu Yuan[1,2,3,4], Jun Xiao[1,2,3,4], Xinyuan Zhao[5], Xinran Du[6], Shiyu Xia[7], Lianrong Wang[3,4,8], Bin Zhu[5 ✉] & Longfei Wang[1,2,3,4 ✉]

Prokaryotes have evolved intricate innate immune systems against phage infection[1–7]. Gabija is a highly widespread prokaryotic defence system that consists of two components, GajA and GajB[8]. GajA functions as a DNA endonuclease that is inactive in the presence of ATP[9]. Here, to explore how the Gabija system is activated for anti-phage defence, we report its cryo-electron microscopy structures in five states, including apo GajA, GajA in complex with DNA, GajA bound by ATP, apo GajA–GajB, and GajA–GajB in complex with ATP and Mg^{2+}. GajA is a rhombus-shaped tetramer with its ATPase domain clustered at the centre and the topoisomerase–primase (Toprim) domain located peripherally. ATP binding at the ATPase domain stabilizes the insertion region within the ATPase domain, keeping the Toprim domain in a closed state. Upon ATP depletion by phages, the Toprim domain opens to bind and cleave the DNA substrate. GajB, which docks on GajA, is activated by the cleaved DNA, ultimately leading to prokaryotic cell death. Our study presents a mechanistic landscape of Gabija activation.

Gabija is one of the most widespread prokaryotic anti-phage defence systems[1–7]. It is present in approximately 15% of all bacterial and archaeal genomes that have been analysed[8,10,11]. The Gabija system is encoded by two genes, *gajA* and *gajB*. GajA is a member of the overcoming lysogenization defect (OLD) family of nucleases, which contain an N-terminal ATP-binding cassette (ABC) ATPase domain and a C-terminal Toprim catalytic domain[12–14]. Similar to other OLD nucleases, GajA is a DNA-nicking endonuclease that utilizes a two-metal catalytic mechanism for DNA cleavage through conserved E and DxD motifs[9,15–19]. However, unlike other OLD family nucleases, GajA is inhibited by ATP[9,20], suggesting that it has a unique regulatory mechanism that is tailored for anti-phage defence. GajB is predicted to be a UvrD/PcrA/Rep-like helicase[8], which are often found in tandem with OLD nucleases[19]. GajB does not bind to DNA or exhibit helicase activity[21,22]—rather, it senses single-stranded DNA and hydrolyses ATP[21]. Both GajA and GajB are required to exert strong resistance against phage by degrading phage and bacterial genomic DNA upon phage infection[9,21].

To elucidate the mechanism of the Gabija system, we determined its cryo-electron microscopy (cryo-EM) structures in five distinct states. These structures include apo GajA, DNA-bound GajA, ATP-bound GajA, apo GajA–GajB, and ATP and Mg^{2+}-bound GajA–GajB. GajA assembles as a tetramer with its ATPase domain clustered at the centre and the Toprim domain located peripherally. ATP and Mg^{2+} bind to the putative ATPase active site of GajA and introduce conformational changes of the insertion region within the ATPase domain, keeping the Toprim domain in a closed conformation. In the absence of ATP and Mg^{2+}, GajA binds to DNA in a bipartite manner through the Toprim domains. GajB docks

on the GajA tetramer to form a supramolecular GajA–GajB complex. DNA cleaved by GajA activates the ATPase activity of GajB and leads ultimately to cell death. Our findings provide mechanistic insights into the activation of Gabija, an abundant prokaryotic immune system.

GajA tetramer

To elucidate how Gabija defends against phage infection, we first cloned and purified GajA from *Bacillus cereus* VD045 (Extended Data Fig. 1a–c). GajA contains an N-terminal ABC ATPase domain and a C-terminal Toprim domain (Fig. 1a). Unlike previously reported OLD family members that are monomers or dimers[19,20], GajA eluted as a tetramer by size-exclusion chromatography (Extended Data Fig. 1b). The purified GajA tetramer was functional and could perform a double-strand cleavage of a 955-bp known substrate[9] generated by the introduction of a near-palindromic recognition sequence into pUC19-955 plasmid and PCR amplification with flanking sequences (Extended Data Fig. 1d,e); this is hereafter referred to as pUC19-955 DNA.

To validate the tetrameric assembly of GajA and interrogate its function, we determined the cryo-EM structure of GajA (Extended Data Fig. 1f–i). The final cryo-EM map of GajA had a resolution of 2.66 Å and was sufficient to build an near-atomic model (Extended Data Fig. 1f,g). The overall structure of GajA has the shape of a rhombus and can be viewed as a dimer of dimers divided by the shorter diagonal (Fig. 1b,c). The 'head' is designated as the Toprim–ATPase dimer, and the 'shoulder' is the ATPase dimer alone (Fig. 1c). Two GajA protomers form extensive interactions at the dimer interface of the long diagonal to bring
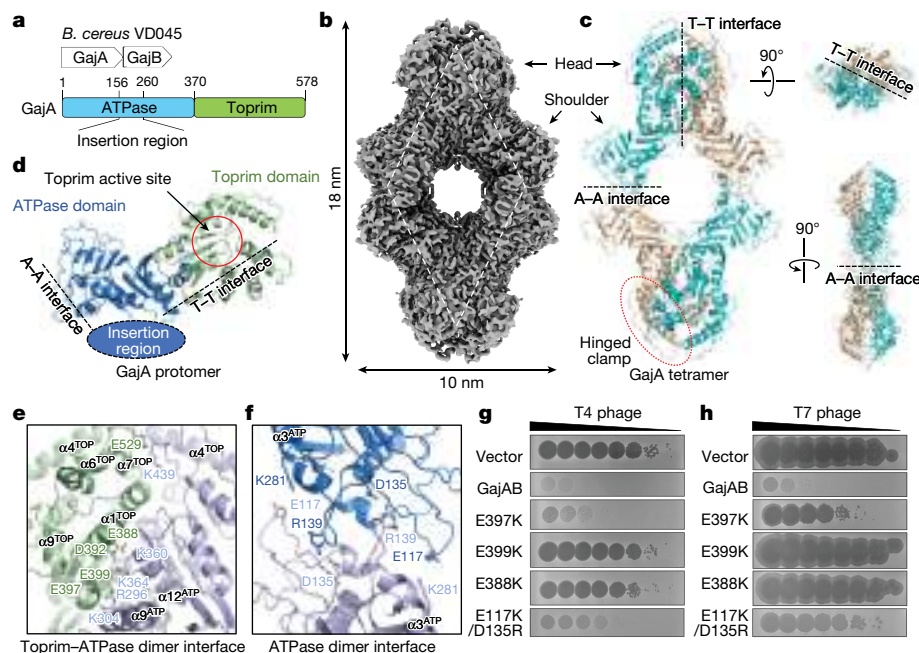
Fig. 1 | Cryo-EM structure of GajA tetramer. a, Schematic of the Gabija operon and domain organization of *B. cereus* GajA. An insertion region is present in the ABC ATPase domain. **b**, The cryo-EM density map of GajA tetramer. **c**, Ribbon diagram of a GajA tetramer in three orthogonal views. GajA protomers are coloured in cyan and wheat. **d**, Structure of a GajA protomer. Dimer interfaces are indicated by dotted lines, the Toprim active site is indicated by a circle. **e**, Enlarged view of the Toprim–ATPase dimer interface. The Toprim–ATPase dimer interface contains charge–charge interactions mediated by K304/E397, R296/E399, K360/D392, K364/E388 and K439/E529. **f**, Enlarged view of the ATPase dimer interface. The ATPase dimer interface contains charge–charge interactions mediated by E117/K281 and D135/R139. **g,h**, Plaque assay with tenfold serial dilutions using T4 (**g**) and T7 (**h**) phages to infect *E. coli* B expressing the wild-type Gabija system (GajAB) and indicated mutants of GajA. Images are representative of three biological replicates.

two Toprim domains together (Fig. 1c,d). Residues 436–500 from two Toprim domains constitute two hinged clamps (Fig. 1c), which are common in topoisomerases[23]. In addition, the GajA Toprim domain has a conserved two-metal active site (Extended Data Figs. 1l,m and 2a). We previously demonstrated that mutation of the active site E379A abolishes the DNA endonuclease activity of GajA[9].

The N-terminal ABC ATPase domain (1–370) mediates interaction at both dimer interfaces, whereas the C-terminal Toprim domain is only present in the head region (Fig. 1c,d). An insertion (156–260) with predicted α-helices is present in the ABC ATPase domain of GajA but is not resolved in the cryo-EM map, probably owing to flexibility. A similar insertion was reported to facilitate the dimerization of *Thermus scotoductus* OLD, but this dimerization was not a productive conformation for ATP hydrolysis[20], suggesting that this insertion region may have other functions.

To test whether the GajA tetramer is required for its biological function, we carried out a plaque assay using *Escherichia coli* cells expressing the Gabija operon that consists of *gajA* and *gajB*. We first examined the Toprim–ATPase and ATPase dimer interfaces. We found extensive salt bridges at the Toprim–ATPase interface (Fig. 1e and Extended Data Fig. 1j), mediated by five negatively charged residues (E388, D392, E397, E399 and E529) and five positively charged residues (R296, K304, K360, K364 and K439). Conversely, fewer interactions were found at the ATPase dimer interface (Fig. 1f and Extended Data Fig. 1k), mediated by only four residues (E117, D135, R139 and K281). Charge-reversal mutations of these key residues compromise Gabija-mediated resistance to infection by phages T4 and T7 (Fig. 1g,h and Extended Data Fig. 3b). Notably, mutations at the Toprim–ATPase dimer interface (E388K, E399K, E397K/E399K and E388K/D392R/E529K) completely abolish the anti-phage activity of Gabija and a mutation at the ATPase dimer interface (E117K/D135R) partially impairs this anti-phage activity. Furthermore, charge-reversal mutants of GajA remain soluble and can be purified (Extended Data Fig. 3c), but no longer form a tetramer (Extended Data Fig. 3d,e) and lose endonuclease activity (Extended Data Fig. 4f,g). Together, our results suggest that GajA functions as a tetramer.

DNA binding and cleavage by GajA

OLD family nucleases can efficiently process DNA as monomers or dimers[19,20], which raises the question of why GajA functions as a tetramer. To answer this question, we set out to determine the structure of GajA with its DNA substrate in order to understand how GajA tetramer binds to DNA.

We first screened for the minimal DNA substrate containing a previously characterized near-palindromic sequence[9] (Extended Data Fig. 3a–c), and identified a 21-bp DNA substrate that can still bind to and be cleaved by GajA (Fig. 2a–c). We incubated GajA with the 21-bp dsDNA substrate in the presence of Ca^{2+} and performed cryo-EM imaging (Extended Data Fig. 3d–g). Since we previously demonstrated that Ca^{2+} inhibits the nuclease activity of GajA but allows its binding to dsDNA[9], we expected to capture a stable GajA–DNA complex. Two-dimensional class averages of the GajA–DNA sample were similar to the GajA tetramer but had an expanded head region (Extended Data Fig. 3d), indicating DNA-induced conformational changes. The final 3D reconstruction yielded an electron microscopy density map at 3.1 Å resolution (Extended Data Fig. 3d,e), with extra densities corresponding to DNA located at both ends of the GajA tetramer, sandwiched by the Toprim domains (Fig. 2d,e). DNA binding does not change the overall architecture of GajA tetramer but induces marked conformational changes at the Toprim domains (Fig. 2d,e).

To better define GajA–DNA interactions, we performed focused refinement on one of the two GajA dimers (Extended Data Fig. 3h–k). This procedure improved the map resolution of the head region to 2.9 Å (Extended Data Fig. 3k–m), which enabled us to trace all bases of the 21-bp dsDNA. The 21-bp DNA is threaded into the Toprim dimer,
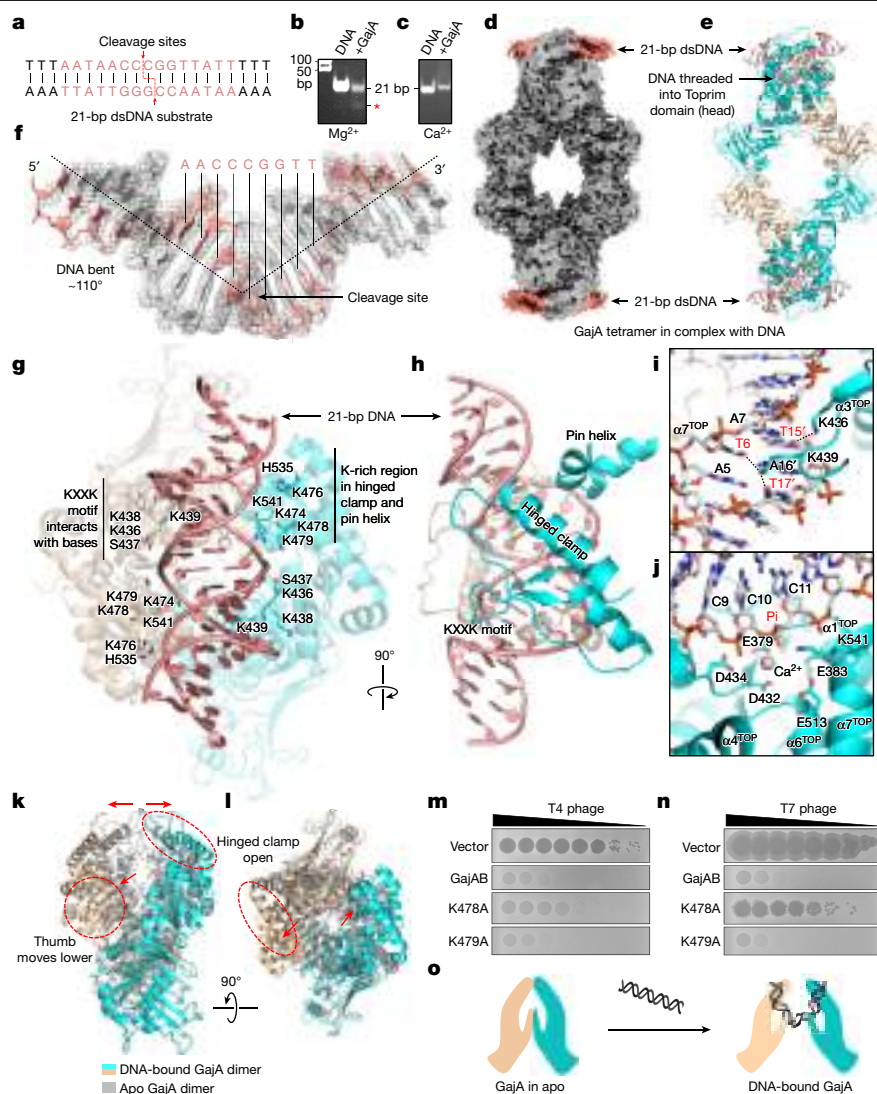
Fig. 2 | Cryo-EM structure of GajA in complex with DNA. a, The 21-bp DNA substrate sequence and cleavage sites of GajA on a 21-bp dsDNA substrate. Recognition sequences are highlighted in red. **b**, Cleavage of 21-bp dsDNA by GajA in the presence of 1 mM Mg^{2+}. **c**, Binding of 21-bp dsDNA by GajA in the presence of 5 mM Ca^{2+}. Assays in **b**,**c** were repeated three times with similar results. **d**, Cryo-EM density map of GajA in complex with two 21-bp dsDNAs. The 21-bp dsDNAs are coloured in salmon; GajA tetramer is coloured in grey. **e**, Structural overview of GajA in complex with two 21-bp dsDNAs. **f**, Enlarged view of the bound DNA. Nucleotides in one DNA strand are coloured in salmon. **g**,**h**, Enlarged view of the interface between GajA (cyan and wheat) and bound DNA (salmon) (**g**) and an orthogonal view (**h**). **i**, Enlarged view of interface between KXXK motif (cyan) and bound DNA. **j**, Enlarged view of catalytic residues at the Toprim active site (cyan) that can interact with bound DNA via Ca^{2+} (grey). **k**,**l**, Overlaid ribbon diagram, showing the conformational changes from apo GajA (grey) to DNA-bound GajA (coloured) (**k**) and an orthogonal view (**l**). **m**,**n**, Plaque assay with tenfold serial dilutions using T4 (**m**) and T7 (**n**) phages to infect *E. coli* B expressing the wild-type Gabija system and indicated mutants of GajA. Images are representative of three biological replicates. **o**, Schematic diagram of GajA binding to DNA.

embraced by the two hinged clamps (Fig. 2e–g). This embracing of DNA by GajA is similar to that of a T-segment by topoisomerase II[23,24]. Notably, the 21-bp DNA is bent about 110° upon binding to GajA (Fig. 2f), similar to the G-segment bound to topoisomerase II[25], suggesting that DNA bendability may be required for substrate recognition in OLD family nucleases.

Interactions between DNA and the Toprim domain are quite extensive through three interfaces (Fig. 2g–i and Extended Data Fig. 4a). The first DNA–GajA interface is the KXXK (436–439) motif that protrudes into the minor groove of substrate DNA, with K436 and K439 forming hydrogen bonds with the bases of T6 on one DNA strand, and T15' and T17' on the other DNA strand (Fig. 2g–i). The second DNA-binding interface involves a lysine-rich region (458–503) with lysine residues (K474, K476, K478 and K479) located in the hinged clamp that interact with the phosphate backbone of the DNA (Fig. 2g,h). The third interface consists of a pin helix (533–550) that inserts into the major groove of DNA to facilitate DNA bending (Fig. 2h). These DNA-binding interfaces collectively bind and distort DNA such that the Toprim active site is aligned perfectly to the DNA cleavage site between C10 and C11. We found only one Ca^{2+} at the Toprim active site, coordinated by E379, D434 and D432, as well as the phosphate between C10 and C11 (Fig. 2j). Since OLD family nucleases utilize a two-metal catalytic mechanism for DNA cleavage, it is likely that Ca^{2+} inhibits the cleavage activity of GajA by occupying only one of the two metal sites.

By overlaying DNA-bound GajA dimer with apo GajA, we found obvious conformational changes of the Toprim domain of GajA upon DNA binding, including the opening of the hinged clamps to accommodate the DNA and lowering of the 'thumb' that contains the Toprim active site (Fig. 2k,l). To validate the interactions between GajA and DNA, we mutated key residues of GajA at the three DNA-binding interfaces and

carried out DNA cleavage and binding assays (Supplementary Table 1). Expectedly, mutations of residues from all three interfaces (K438A, K439A, K474A, K476A, K478A and K539A) result in significantly reduced DNA binding and DNA cleavage activity (Extended Data Fig. 4b–h and Supplementary Table 1). Mutation of K541 at the active site does not affect DNA binding but abolishes DNA cleavage. Of note, although K436 is not part of the active site, it abolishes DNA cleavage without affecting DNA binding. Further DNA cleavage assays using pUC19 plasmid DNA revealed that although K436A does not cleave the 955-bp substrate, it nicks the pUC19 plasmid (Extended Data Fig. 4f), indicating a shift of the substrate specificity of GajA. Thus, all three interfaces are important for DNA binding and K436 may contribute to substrate recognition.

We then tested whether DNA binding is critical for Gabija's anti-phage defence, and carried out plaque assays using selected mutations at the DNA binding interfaces. As expected, K478A, a mutant that significantly reduces DNA binding and abolishes DNA cleavage (Extended Data Fig. 4b,g and Supplementary Table 1) can greatly reduce the anti-phage activity of Gabija in the plaque assay (Fig. 2m,n). K479A shows weak cleavage activity but still has strong anti-phage defence, suggesting that weak cleavage activity of GajA is still sufficient for the anti-phage defence (Fig. 2m,n and Extended Data Fig. 4b,g). Similarly, although K436A of GajA does not cleave the 955-bp DNA substrate, it can still nick the pUC19 plasmid and exhibit strong anti-phage activity in the plaque assay (Fig. 2i).

Together, our results revealed three surprising discoveries: (1) GajA tetramer binds to DNA in a bipartite manner through dimerized Toprim domains at the two ends; (2) DNA is bent approximately 110° upon binding to GajA; and (3) the Toprim dimer undergoes conformational changes, including opening of the hinged clamps and lowering the thumb region, allowing DNA to go through (Fig. 2o).

Apo GajA–GajB

Bacterial defence systems utilize nucleases as the executer of cell death by degrading host and phage DNA, therefore it is likely that GajA functions similarly. However, GajB is still essential for anti-phage defence[21], suggesting that there is functional coupling between GajA and GajB. GajB is homologous to the UvrD helicases (Fig. 3a and Extended Data Fig. 5a–c) that are often associated with OLD family nucleases[19,20] (like GajA). Therefore, it is likely that GajA and GajB function as a complex. To test this hypothesis, we added purified GajB to GajA and performed cryo-EM imaging (Extended Data Figs. 6 and 7a–c). Notably, GajB has two isoforms, GajB and GajB-S; in the current study, we focused on the more stable GajB[21]. Two-dimensional class averages of the cryo-EM sample revealed several GajA–GajB complexes with the same GajA tetramer but different numbers of GajB bound to the shoulder region of the tetramer, indicating stepwise binding of GajB to GajA (Fig. 3b). Out of all GajA–GajB complexes, we were able to obtain high-resolution structures of the 4:1 and 4:4 GajA–GajB complexes (Fig. 3c,d and Extended Data Fig. 6). In both complexes, GajB is mounted on the extruding region (residues 76–104) between strand β4 and β5 in the ATPase domain of GajA (Fig. 3e). Similar to UvrD, the structure of GajB can be divided into four subdomains: 1A, 1B, 2A and 2B. Among the four domains, the 1B domain forms extensive interactions with GajA (Fig. 3f). In UvrD, 1A and 2A are involved in ATP hydrolysis, which drives conformational changes to translocate single-stranded DNA[26]. By overlaying GajB and UvrD, we observed two major differences (Extended Data Fig. 7d–i): (1) Although 1A and 1B stay the same, the conformations of 2A and 2B are very different (both apo and DNA-bound states); and (2) the 2B domain of GajB is significantly shorter than that of UvrD. The 2B domain mediates the majority of the interactions between UrvD and dsDNA. These differences suggest an alternative function of GajB. Indeed, we demonstrated that in contrast to other UvrD-like helicases[21], GajB alone does not bind or unwind DNA.

In the 4:4 GajA–GajB complex, 4 GajB molecules dock onto the shoulder regions of the GajA tetramer to form a supercomplex of approximately 500 kDa (Fig. 3g), such that there are two GajB molecules on one side of the GajA tetramer to create a GajB dimer (Fig. 3h). GajB is expressed at a lower level than GajA[21], suggesting that the molar ratio between GajA and GajB affects the function of GajA–GajB. To validate whether the stoichiometry of GajA and GajB is important for the Gabija system, we perturbed the molar ratio of GajA and GajB and performed the plaque assay (Extended Data Fig. 7m). When GajA and GajB are expressed using separate promoters (referred to as GajA+B, Extended Data Fig. 7j), GajB is expressed at a much higher level than when both are expressed under a single promoter (Extended Data Fig. 7k,l) and Gabija completely loses anti-phage activity (Extended Data Fig. 7m). When over-expressing GajB in the presence of GajA–GajB (referred to as GajAB+B), we also observed a complete loss of anti-phage activity (Extended Data Fig. 7m). Thus, we confirmed that the molar ratio of GajA–GajB is important for the anti-phage defence. To investigate whether this ratio preference is due to the interference between two neighbouring GajA-bound GajB, we overlaid the 4:1 GajA–GajB complex onto the 4:4 GajA–GajB complex, finding only subtle conformational changes of GajB (Extended Data Fig. 7n). However, when DNA-bound UvrD is superimposed onto the 4:4 GajA–GajB complex anchored at 1A and 1B subdomains, there are clashes between UvrD and the neighbouring GajB protomer (Extended Data Fig. 7d), suggesting a potential clash between two GajB protomers if DNA induces conformational changes in GajB.

To test whether the GajA–GajB complex is crucial for the anti-phage activity of the Gabija system, we mutated key residues that mediate the GajA–GajB interactions and carried out the plaque assay (Fig 3i,j and Extended Data Fig. 6i,j). A single mutation of D93A and double mutations of H46A/K160A, D93A/K160A and Q155A/Y347A in GajB all resulted in the reduction of anti-phage activity (Fig. 3k,l). In particular, the Q155A/Y347A mutation almost entirely abolishes the anti-phage defence against T4 and T7. To confirm that Q155A/Y347A indeed disrupts the binding of GajA and GajB, we purified the GajB Q155A/Y347A mutant (Extended Data Fig. 7o–q), and carried out both native PAGE (Extended Data Fig. 7r) and pull-down assays using GajA as the bait (Extended Data Fig. 7s,t). In both experiments, we observed much weaker binding between GajA and GajB(Q155A/Y347A) compared with GajA and wild-type GajB. Therefore, the supramolecular assembly of GajA–GajB is critical for anti-phage defence via the Gabija system.

Pre-catalytic GajA–GajB

Similar to other abortive infection systems[27,28] such as cyclic oligonucleotide-based anti-phage signalling system[29,30] (CBASS) and antiviral ATPases/NTPases of the STAND[31,32] (AVAST), GajA–GajB is not active in normal cellular conditions. GajA is inactive in the presence of ATP[9] (Extended Data Fig. 9a), but how GajA–GajB remains in a pre-catalytic inactive state is unclear. To answer this question, we obtained the structure of GajA–GajB in complex with ATP and Mg^{2+} at 3 Å resolution (Extended Data Fig. 8a–g). The overall structure of ATP-bound GajA–GajB is almost identical to the apo GajA–GajB structure, except for the additional density covering the central cavity of the GajA tetramer (Fig. 4a). After local refinement of the GajA region, we confirmed that the additional density belongs to the insertion region and were able to trace the entire insertion region for model building (Fig. 4b,e and Extended Data Fig. 9b,c). Further examination revealed an ATP-binding site in the ATPase domain of GajA, and the Mg^{2+} next to the ATP (Fig. 4b).

To further interpret how ATP binds to the ATPase domain, we analysed the ATP-binding site in the ATPase domain. ATP binding involves the phosphate-binding loop (P-loop) (residues 29–36, GMNDIGKT) and H320, which interact directly with the tri-phosphate (Fig. 4c). A Mg^{2+} is also present at the ATPase active site and coordinated by the phosphates and three residues, T36, E288 and E289 (Fig. 4d). ATP
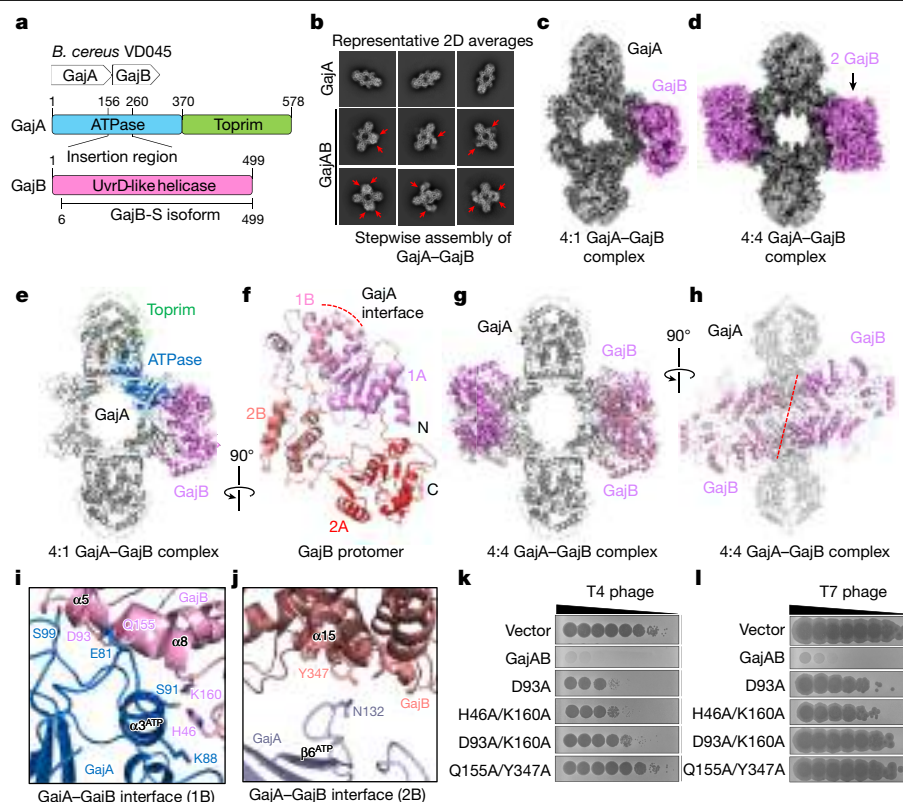
Fig. 3 | Cryo-EM structure of 4:1 and 4:4 GajA–GajB complexes. a, Schematic of the Gabija operon and domain organization of *B. cereus* GajA and GajB. GajA has an N-terminal ABC ATPase domain and C-terminal Toprim domain. An insertion region is present in the ABC ATPase domain. GajB has an isoform termed GajB-S, in which the first five amino acids are absent. **b**, Two-dimensional class averages of different GajA–GajB complexes. GajB is indicated by red arrows. **c,d**, Cryo-EM density map of 4:1 (**c**) and 4:4 (**d**) GajA–GajB complexes. GajA is shown in grey and GajB is shown in magenta. **e**, Ribbon diagram of a 4:1 GajA–GajB complex. One GajA protomer is coloured with the Toprim domain in pale green and the ATPase domain in blue, and GajB is shown in magenta. **f**, Structure of GajA-bound GajB divided into four domains: 1A (magenta), 1B (pink), 2A (red) and 2B (salmon). **g,h**, Ribbon diagram of a 4:4 GajA–GajB complex with GajA in grey and GajB in magenta and pink (**g**) and an orthogonal view (**h**). **i,j**, Enlarged view of GajA–GajB interfaces at 1B (**i**) and 2B (**j**) domains of GajB. **k,l**, Plaque assay with tenfold serial dilutions using T4 (**k**) and T7 (**l**) phages to infect *E. coli* B expressing the wild-type Gabija system and indicated mutants of GajB. Images are representative of three biological replicates.

density is absent in the cryo-EM structure of GajA with ATP (Extended Data Fig. 8h–k), indicating that Mg^{2+} is required for ATP binding at the ATPase site. The insertion region coordinates ATP through a G-rich loop (G-loop) (residues 258–260, GDG) that also interacts with the tri-phosphates (Fig. 4c,d and Extended Data Fig. 9d,e). Y254 of the insertion region and the adenine group of ATP form π-stacks. D259 of the insertion region coordinates the Mg^{2+}. Thus, the insertion region forms extensive interactions with the ATP and Mg^{2+}. Overall, the binding of ATP stabilizes the insertion region, which is evident from the ordered insertion region being absent from the apo GajA and GajA–GajB structures owing to flexibility (Fig. 4a,b,e). To validate that the ABC ATPase active site is responsible for ATP-induced inhibition, we designed mutations of the key residues that coordinate ATP binding (Extended Data Fig. 9f–h and Supplementary Table 2). Notably, H320A lifts inhibition of DNA cleavage by ATP (Fig. 4f). Our results suggest that the ATP binding site at ATPase domain contributes to the inactivity of GajA in the presence of ATP.

Together, these results offer an elegant model of how GajA–GajB is maintained in an inactive pre-catalytic state. In this model, ATP binding at the ATPase domain of GajA stabilizes the insertion region, keeping the Toprim domain of GajA in a closed, inactive state. It is conceivable that the inhibitory effect extends allosterically from the insertion region to Toprim.

Finally, we considered an alternative, potential possibility of ATP binding at the Toprim domain. We pursued this possibility because, although ATP was not directly visible in the Toprim domain of ATP/

Mg^{2+}-bound GajA–GajB, we observed ATP binding at the Toprim domain in the structure of ATP-bound GajA without Mg^{2+} (Extended Data Fig. 9i–l). We mutated residues in the ATP-binding pocket in the Toprim domain (Extended Data Fig. 9m–s and Supplementary Table 2), and found that several mutants, including H535R, are catalytically active in the presence of ATP (Extended Data Fig. 9o,p). These results indicate that Toprim may potentially contribute to GajA inhibition, although further investigation will be required to test this.

GajB activation

To investigate whether the functions of GajA and GajB are coupled, we first compared ATPase activity by incubating GajB and DNA with or without GajA (Extended Data Fig. 9t). Although the active site of GajB and UvrD are highly conserved (Extended Data Fig. 5b,c), no ATPase activity was observed in the presence of T7 DNA, which was expected because of the inability of GajB to bind DNA, probably owing to the shorter 2B subdomain. We then tested whether GajA facilitates activation of GajB. As expected, when GajA, GajB and DNA were incubated together, ATP was rapidly hydrolysed. By contrast, no ATPase activity was observed when the GajB active site was mutated (GajB-D167A/E168A) or when GajB was absent (Extended Data Fig. 9t). Finally, we explored whether GajB activation is a result of GajA nicking; we replaced GajA with an active site mutant E379A that has no nuclease activity. This GajA mutant completely abolished the ATPase activity of GajB (Extended Data Fig. 9t). Thus, our data indicate that GajA-nicked DNA
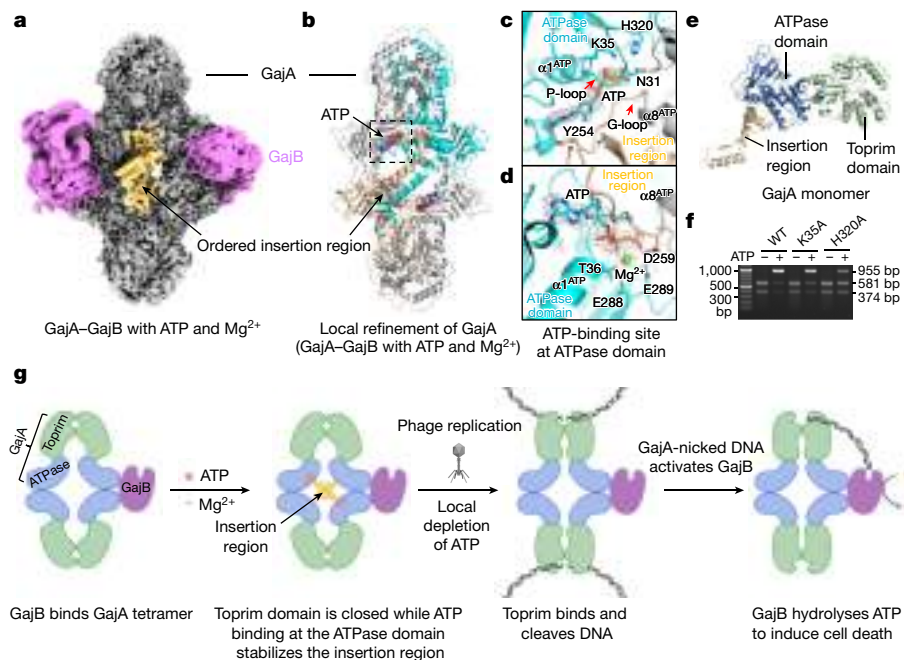
Fig. 4 | Cryo-EM structure of GajA–GajB in complex with ATP and Mg²⁺.
a, Cryo-EM density map of GajA–GajB in complex with ATP and Mg²⁺. GajA is coloured in grey, GajB is in magenta, and the extra density that corresponds to the insertion region is coloured in yellow. **b**, Overview of the structure of GajA–GajB in complex with ATP and Mg²⁺. Four ATP molecule (spheres) binding sites are located at the ABC ATPase domain active site. **c**,**d**, Enlarged view of the ATP binding pocket at the ABC ATPase domain in ATP-bound GajA–GajB. Residues that interact with ATP (**c**) and residues that coordinate Mg²⁺ (**d**) are shown. The insertion region is coloured in wheat and the rest of the ABC ATPase domain are coloured in cyan and grey (neighbouring protomer); ATP is coloured in cyan and Mg²⁺ is coloured in green. The mesh indicates the electron microscopy density of ATP and Mg²⁺. **e**, Structure of a GajA monomer including the insertion region. **f**, Cleavage of pUC19-955 DNA by GajA or GajA mutants with or without 0.5 mM ATP. The assays were repeated three times with similar results. WT, wild type. **g**, Model of anti-phage defence by the Gabija system. The orange circle indicates ATP, the red star indicates Mg²⁺ and the grey circle indicates Pi.

activates GajB. Together, these results suggest that nicked DNA generated by GajA activates the ATPase activity of GajB.

Discussion

Immune signalling often involves higher-order assemblies[33]. Although Gabija is a relatively simple defence system that comprises only two protein components, GajA and GajB, the GajA–GajB supramolecular complex can carry out pattern recognition, signal amplification and execution of cell death. These processes require signalling cascades in mammalian cells. Notably, the Gabija system exerts strong anti-phage activity against a broad range of phages, despite functioning as only one protein complex. Since GajB does not bind DNA on its own[21,22], the GajA–GajB supramolecular complex and the bipartite binding of the GajA tetramer to DNA is likely to facilitate the handover of GajA-bound DNA to GajB.

In summary, we reveal that in normal cellular conditions, the ATPase domain of GajA binds ATP and Mg²⁺, stabilizing the insertion region. ATP binding at ATPase domain and ordered GajA insertion may potentially have an allosteric role in ATP-mediated nuclease inhibition. The two hinged clamps of the Toprim dimer of GajA remain in a closed conformation in the presence of ATP and need to open up to bind and cleave DNA. Upon phage infection, local depletion of ATP enables GajA to cleave DNA, and nicked DNA then activates GajB and ultimately leads to cell death (Fig. 4g). Thus, Gabija may represent a unique prokaryote immune system that senses metabolites as danger signals. Recent studies highlight the similarities between the prokaryotic defence system and the eukaryotic innate immune system[5,7,8,34–39], suggesting that similar principles are applied for immune defence across all organisms. The mechanistic insights presented here expand our knowledge of prokaryotic defence systems and underline the role of metabolites as a potential danger signal for immune defence.

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/s41586-024-07270-x.

1. Stern, A. & Sorek, R. The phage-host arms race: shaping the evolution of microbes. *BioEssays* **33**, 43–51 (2011).
2. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* **577**, 327–336 (2020).
3. Dy, R. L., Richter, C., Salmond, G. P. C. & Fineran, P. C. Remarkable mechanisms in microbes to resist phage infections. *Annu. Rev. Virol.* **1**, 307–331 (2014).
4. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).
5. Millman, A. et al. An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556–1569.e1555 (2022).
6. Rousset, F. et al. Phages and their satellites encode hotspots of antiviral systems. *Cell Host Microbe* **30**, 740–753.e745 (2022).
7. Vassallo, C. N., Doering, C. R., Littlehale, M. L., Teodoro, G. I. C. & Laub, M. T. A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome. *Nat. Microbiol.* **7**, 1568–1579 (2022).
8. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
9. Cheng, R. et al. A nucleotide-sensing endonuclease from the Gabija bacterial defense system. *Nucleic Acids Res.* **49**, 5216–5229 (2021).
10. Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
11. Payne, L. J. et al. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Res.* **50**, W541–W550 (2022).
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. Koonin, E. V. & Gorbalenya, A. E. The superfamily of UvrA-related ATPases includes three more subunits of putative ATP-dependent nucleases. *Protein Seq. Data Anal.* **5**, 43–45 (1992).
14. Berger, J. M., Gamblin, S. J., Harrison, S. C. & Wang, J. C. Structure and mechanism of DNA topoisomerase II. *Nature* **379**, 225–232 (1996).

15. Allemand, F., Mathy, N., Brechemier-Baey, D. & Condon, C. The 5S rRNA maturase, ribonuclease M5, is a Toprim domain family member. *Nucleic Acids Res.* **33**, 4368–4376 (2005).
16. Keck, J. L., Roche, D. D., Lynch, A. S. & Berger, J. M. Structure of the RNA polymerase domain of *E. coli* primase. *Science* **287**, 2482–2486 (2000).
17. Podobnik, M., McInerney, P., O'Donnell, M. & Kuriyan, J. A TOPRIM domain in the crystal structure of the catalytic core of *Escherichia coli* primase confirms a structural link to DNA topoisomerases. *J. Mol. Biol.* **300**, 353–362 (2000).
18. Corbett, K. D. & Berger, J. M. Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 95–118 (2004).
19. 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).
20. 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).
21. Cheng, R. et al. Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host Microbe* **31**, 1331–1344 e1335 (2023).
22. Oh, H. et al. Structural and functional investigation of GajB protein in Gabija anti-phage defense. *Nucleic Acids Res.* **51**, 11941–11951 (2023).
23. Champoux, J. J. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* **70**, 369–413 (2001).
24. Laponogov, I. et al. Trapping of the transport-segment DNA by the ATPase domains of a type II topoisomerase. *Nat. Commun.* **9**, 2579 (2018).
25. Dong, K. C. & Berger, J. M. Structural basis for gate-DNA recognition and bending by type IIA topoisomerases. *Nature* **450**, 1201–1205 (2007).
26. 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).
27. Lopatina, A., Tal, N. & Sorek, R. Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).
28. Sather, L. M. et al. A broadly distributed predicted helicase/nuclease confers phage resistance via abortive infection. *Cell Host Microbe* **31**, 343–355.e345 (2023).
29. Millman, A., Melamed, S., Amitai, G. & Sorek, R. Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. *Nat. Microbiol.* **5**, 1608–1615 (2020).
30. Cohen, D. et al. Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* **574**, 691–695 (2019).
31. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
32. Gao, L. A. et al. Prokaryotic innate immunity through pattern recognition of conserved viral proteins. *Science* **377**, eabm4096 (2022).
33. Wu, H. Higher-order assemblies in a new paradigm of signal transduction. *Cell* **153**, 287–292 (2013).
34. Burroughs, A. M. & Aravind, L. Identification of uncharacterized components of prokaryotic immune systems and their diverse eukaryotic reformulations. *J. Bacteriol.* **202**, e00365–20 (2020).
35. Ye, Q. Z. et al. HORMA domain proteins and a Trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity. *Mol. Cell* **77**, 709–722.e7 (2020).
36. Bernheim, A. et al. Prokaryotic viperins produce diverse antiviral molecules. *Nature* **589**, 120–124 (2021).
37. Hogrel, G. et al. Cyclic nucleotide-induced helical structure activates a TIR immune effector. *Nature* **608**, 808–812 (2022).
38. Johnson, A. G. et al. Bacterial gasdermins reveal an ancient mechanism of cell death. *Science* **375**, 221–225 (2022).
39. Kibby, E. M. et al. Bacterial NLR-related proteins protect against phage. *Cell* **186**, 2410–2424.e18 (2023).

Article

Methods

Protein expression and purification

The following coding sequences were cloned into pET28a with an N-terminal 6×His-tag: GajA (GenBank accession number: MW659467), GajB (GenBank accession number: OM891105, residues 1–499), and GajA–GajB (located at bases 94190–97412 of the *B. cereus* VD045 genome (AHET01000033)). All mutations in this study were introduced using the Mut Express MultiS Fast Mutagenesis Kit V2 (Vazyme), and all plasmids had been verified by sequencing. These plasmids were transformed into *E. coli* BL21(DE3) cells and cultured in 1 l LB medium supplemented with 50 μg ml^{-1} kanamycin at 37 °C, induced at OD_{600} of 0.6 with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cultured for 18 h at 18 °C before collection. Cells were collected and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, and 1 mM DTT), and lysed by ultrasonication. Supernatant was collected after centrifugation (16,000 rpm, 4 °C, 30 min), and incubated for 30 min on Ni-NTA agarose, which had been pre-equilibrated with 10 volumes of lysis buffer. After the mixture had been loaded onto the column, it was washed with 10 volumes of elution buffer containing 20 mM followed by 50 mM imidazole. The majority of GajA was eluted off the column by elution buffer containing 100 mM imidazole. Collected eluates were further purified using a Superdex 200 HiLoad (16/600) size-exclusion column (GE Healthcare Life Sciences) equilibrated with lysis buffer. All purified Gabija proteins were prepared the same as GajA while the concentration of imidazole in elution buffer varied. All purified Gabija protein were concentrated to approximately 1 mg ml^{-1} before use.

Cryo-EM imaging and data processing

For cryo-EM grid preparation, GajA samples (3 μl, 1 mg ml^{-1} protein concentration) were applied on glow-discharged Quantifoil R1.2/1.3 300 mesh copper grids, using a Vitrobot Mark IV (FEI) set at blotting force 3, blotting time 3 s, 100% humidity, and 8 °C. Blotted grids were immediately plunged into liquid ethane and transferred to liquid nitrogen for storage. GajA–DNA complex samples were obtained by mixing GajA and 21-bp dsDNA at molar ratio of 1:1.2 and incubating at room temperature for 5 min with 10 mM Ca^{2+} and 0.1 mM Ni^{2+} before vitrification. GajA–GajB complex samples were obtained by mixing GajA and GajB at molar ratio of 1:2 and incubating at 4 °C for 3 h before vitrification. ATP-bound GajA samples were obtained by mixing GajA with 1 mM ATP at 37 °C for 5 min before vitrification. ATP-bound GajA–GajB samples were obtained by mixing GajA and GajB at molar ratio of 1:1.2 and incubating at 4 °C for 1 h with 10 mM ATP and 10 mM Mg^{2+} before vitrification. These grids were screened on a FEI Glacios microscope (Core Facility of Wuhan University) operated at 200 kV and equipped with a Ceta D CMOS camera. Grids with optimal ice thickness and particle density were chosen for cryo-EM data collection. All data collection was performed using a FEI Titan Krios G4 microscope (Core Facility of Wuhan University) operated at 300 kV and equipped with a Gatan K3 direct electron detector. For GajA, GajA–DNA complex, 4:1 GajA–GajB complex, ATP-bound GajA complex, and ATP-bound GajA–GajB complex, 8,627, 5,441, 5,875, 6,991 and 5,079 movies were collected with the Thermo Scientific EPU software, respectively, in counting mode, with 40 total frames per movie, 50 electrons per Å^2 accumulated dose, and 0.84 Å pixel size. For the 4:4 GajA–GajB complex, 2,872 movies were collected with the grid prepared for ATP-bound GajA–GajB with 1 mM ATP and without Mg^{2+} in the same counting mode as before. All datasets were processed in CryoSPARC[40].

For apo GajA, 7,139 images were yielded after patch-based motion correction and contrast transfer function (CTF) estimation, from which 3,023,069 particles were autopicked by blob picker and template picker. After a round of 2D classification, 1,539,938 particles were selected to do ab initio reconstruction followed by heterogeneous refinement with C2 symmetry. Particles from the best class were chosen for 2 rounds of homogeneous refinement producing 720,926 particles that were used in one additional round of heterogeneous refinement with D_2 symmetry.

The best class, with 275,074 particles, was used to do homogeneous refinement with D_2 symmetry. The resulting map has a resolution of 2.66 Å.

For GajA–DNA complex, after blob picking and 2 rounds of 2D classification, 545,240 particles were selected and used to do heterogeneous refinement with D_2 symmetry. The best class was used to do homogeneous refinement with D_2 symmetry. After homogeneous refinement, 539,168 particles were yielded and then selected to do local refinement with C_1 symmetry, followed by 3D classification. The particles in class 0 were selected to do non-uniform refinement and heterogeneous refinement with D_2 symmetry. Finally, 88,461 particles in the best class were used to do non-uniform refinement with D_2 symmetry. The resulting map has a resolution of 3.12 Å. To improve the density quality of the TOPRIM domain's head region, 570,317 particles in class 1 were selected after 3D classification and used to do local refinement with C_1 symmetry and with a local mask on a dsDNA-bound GajA dimer. The resulting map has a resolution of 2.92 Å. In addition, DeepEMhancer[41] was used to assist model building.

For the 4:1 GajA–GajB complex, after patch-based motion correction, CTF estimation and a round of 2D classification, 360,360 particles were selected to do ab initio reconstruction, followed by heterogeneous refinement with C_1 symmetry. Choosing the particles from the best class for homogeneous refinement produced 246,070 particles that were used to perform 3D classification; the best class with 78,983 particles was used to do homogeneous refinement with C_1 symmetry. The resulting map has a resolution of 3.39 Å.

For the 4:4 GajA–GajB complex, 2,872 movies were recorded. After patch-based motion correction, CTF estimation, blob picking and three rounds of 2D classification, 632,358 particles were selected to do ab initio reconstruction followed by heterogeneous refinement with C_1 symmetry. Particles from the best class were chosen for homogeneous refinement producing 243,515 particles that were used in one additional round of ab initio reconstruction, followed by heterogeneous refinement with D_2 symmetry. The volume of the best class was used to do 2 rounds of heterogeneous refinement with D_2 symmetry. Finally, 21,115 particles were used to do homogeneous refinement with D_2 symmetry, and then subjected to do local resolution estimation and local filtering. The resulting map has a resolution of 3.20 Å.

For the ATP-bound GajA complex, 3,000 images were yielded after patch-based motion correction and CTF estimation. After blob picking and a round of 2D classification, 501,032 particles were selected to do ab initio reconstruction followed by heterogeneous refinement with D_2 symmetry. Particles from the best class were chosen for homogeneous refinement producing 220,381 particles that were used in one additional round of heterogeneous refinement with D_2 symmetry. The best class with 56,959 particles was used to do homogeneous refinement with D_2 symmetry and then subjected to do local resolution estimation and local filtering. The resulting map has a resolution of 2.81 Å.

For the ATP-bound GajA–GajB complex, 504,335 particles were selected after patch-based motion correction, CTF estimation, blob picking and two rounds of 2D classification, and then used to do ab initio reconstruction followed by heterogeneous refinement with C_1 symmetry. The best class with 235,072 particles was used to perform homogeneous refinement with C_1 symmetry. The resulting map has a resolution of 2.81 Å. To improve the extra density quality of the central cavity of GajA tetramer, 146,586 particles were used to do local refinement with C_1 symmetry. Subsequently, they were subjected to a further round of focused 3D classification. The best class, with 41,197 particles, was used to do local refinement of the GajA region with C_1 symmetry, and then subjected to do local resolution estimation and local filtering. The resulting map has a resolution of 3.01 Å.

Model building and refinement

There was no crystallographic structural information available for GajA and GajB. The AlphaFold[42] and ColabFold[43] predicted model of dimer

GajA and monomer GajB were used as a guide together with secondary structure prediction program, Jpred[44]. All models were built in Coot[45]. The final models of GajA, GajA–DNA complex, and GajA–GajB complexes (4:1 and 4:4) and ATP-bound GajA contain residues of GajA from 1 to 156 and 281 to 578 (missing the ATPase insertion region). GajA–GajB complexes (4:1 and 4:4) contain residues of GajB from 1 to 499. The maps of ATP/Mg^{2+}-bound GajA–GajB (full map and focused refinement map) have the electron density of the insertion region which is sufficient for model building. The maps of ATP/Mg^{2+}-bound GajA–GajB lack density for the model building of the entire GajB. As a result, the ATP and Mg^{2+}-bound GajA–GajB model (full map) contains GajB residues 1–200 and the ATP/Mg^{2+}-bound GajA model (focused refinement of GajA region based on the full map) does not include GajB. The model refinement was carried out by real-space refinement in PHENIX[46]. The quality of the model was assessed by MolProbity[47]. Statistical details can be found in Extended Data Table 1. Figures were prepared in PyMOL and UCSF ChimeraX[48].

Plaque assays

Phages were propagated by picking a single phage plaque into a liquid culture of *E. coli* B grown at 37 °C to an OD_{600} of 0.3–0.4 in LB medium until culture collapse. The culture was then centrifuged for 4 min at 12,000 rpm and the supernatant was filtered through a 0.2-μm filter to remove remaining bacteria and bacterial debris. Plaque assays were carried out as previously described[8,49,50]. The sequences of GajA–GajB or its mutants were cloned into the pQE82L vector. The recombinant vectors were transformed into *E. coli* B. A single bacterial colony was picked from a fresh LB agar plate and grown in LB broth containing ampicillin (100 μg ml^{-1}) at 37 °C to an OD_{600} of ~0.4. Protein expression was induced by the addition of 0.2 mM IPTG. After further growth for ~1 h, 500 μl of the bacterial cultures was mixed with 14.5 ml of 0.5% LB top agar, and the samples were poured onto LB plates containing ampicillin (100 μg ml^{-1}) and IPTG (0.1 mM). Plates were spotted with 4 μl of the two phages diluted in LB at 8 tenfold dilutions, namely, 10^{-1} to 10^{-8} for T7 and 10^{0} to 10^{-7} for T4. Plates were incubated at 37 °C overnight and imaged. All original source images of plaque assays are in Supplementary Fig. 1.

Preparation of double-stranded DNA

The synthetic double-stranded DNA (dsDNA) was prepared by mixing equimolar amounts (50 μM) of complementary single-stranded oligonucleotides in a total volume of 100 μl in annealing buffer (10 mM Tris-HCl pH 7.4, 50 mM NaCl). The synthetic 21-bp dsDNA was prepared by mixing equimolar amounts (100 μM) of complementary single-stranded oligonucleotides in a total volume of 20 μl in ddH$_2$O. Complementary oligonucleotides were annealed by heating at 95 °C for 5 min followed by gradient cooling to 25 °C over a period of 110 min.

Electrophoretic mobility shift assay

The DNA substrates pUC19-955 were amplified from plasmid pUC19 by primers pUC19-F/R containing two overlapping GajA recognition sequences AATAA*CCCGG*TTATT (one recognition site in the plus strand shown in bold and the other in the unbolded minus strand). pUC19-955 DNA (125 ng) was incubated with 2.5 μM GajA or its mutants in a final volume of 10 μl binding buffer (20 mM Tris-HCl pH 8.0 and 5 mM CaCl$_2$) at 4 °C for 30 min. Reactions were stopped by adding 2 μl of 6× loading dye containing 20 mM EDTA. Samples were analysed by native agarose gel electrophoresis. Samples with substrates of synthetic double-stranded DNA were analysed by PAGE. All original source images of electrophoretic mobility shift assays are in Supplementary Fig. 1.

Hydrolytic activity assays

Hydrolytic activity was determined by the PiColorLock phosphate detection system kit (Expedeon), which measures the amount of free phosphate released. The reactions were performed in hydrolysis reaction buffer (20 mM Tris-HCl pH 8.0, 2 mM MnCl$_2$, and 100 μg ml^{-1} BSA) with 0.2 mM ATP and 50 ng T7 DNA in the presence of 1 μM GajA, GajB, or their mutants at 37 °C for 15 min, unless stated otherwise. The reaction was stopped by adding Malachite Green solution at a sample:dye ratio of 4:1. Subsequent processing was performed following the kit manual and the samples were quantified by a NanoPhotometer (Implen) at 650 nm.

DNA cleavage assays

DNA cleavage experiments were performed in 10-μl reaction volumes with 125 ng of pUC19-955 DNA (or 0.6 μM synthetic dsDNA) and 0.3 μM GajA or its mutants in the reaction buffer (20 mM Tris-HCl pH 9.0, 1 mM MgCl$_2$, and 1 mM DTT, with or without 0.5 mM ATP). Reactions were carried out at 37 °C for 5 min and then stopped by adding 2 μl of 6× loading dye. Samples were analysed by native agarose gel electrophoresis. Samples with substrates of synthetic double-stranded DNA were analysed by polyacrylamide gel electrophoresis. All original source images of electrophoretic separation are in Supplementary Fig. 1.

Reporting summary

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

Data availability

Cryo-EM maps and the atomic coordinates of GajA have been deposited in the Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB) under the accession codes EMD-36541 and 8JQ9. Cryo-EM maps and the atomic coordinates of ATP-bound GajA have been deposited in the EMDB and PDB under the accession codes EMD-37915 and 8WY4. Cryo-EM maps and the atomic coordinates of GajA–dsDNA have been deposited in the EMDB and PDB under the accession codes EMD-37916 and 8WY5. Cryo-EM maps and the atomic coordinates of dsDNA-bound GajA dimer (focused refinement) have been deposited in the EMDB and PDB under the accession codes EMD-38058 and 8X51. Cryo-EM maps and the atomic coordinates of the 4:1 GajA–GajB complex have been deposited in the EMDB and PDB under the accession codes EMD-36569 and 8JQC. Cryo-EM maps and the atomic coordinates of the 4:4 GajA–GajB complex have been deposited in the EMDB and PDB under the accession codes EMD-36563 and 8JQB. Cryo-EM maps and the atomic coordinates of ATP and Mg^{2+}-bound GajA–GajB have been deposited in the EMDB and PDB under the accession codes EMD-38071 and 8X5N. Cryo-EM maps and the atomic coordinates of ATP and Mg^{2+}-bound GajA (focused refinement) have been deposited in the EMDB and PDB under the accession codes EMD-38070 and 8X5I. Source data are provided with this paper.

40. 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).
41. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
42. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
43. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
44. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).
45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
46. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
47. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
48. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
49. 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).
50. 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).

Article

Author contributions Longfei Wang and B.Z conceived the project. J.L., R.C., Z.W., W.Y., J.X. and X.Z. purified proteins. J.L. prepared the cryo-EM samples and performed data collection. Longfei Wang and J.L. solved the cryo-EM structures. X.D. and Longfei Wang carried out structure predictions. Z.W. and Longfei Wang performed model building. J.L. and W.Y. performed molecular cloning for mutants. Longfei Wang and Z.W. analysed the structures and designed experiments. R.C. performed phage, ATPase, gel shift and DNA cleavage assays. Longfei Wang, J.L. and Z.W. wrote the manuscript with input from R.C., W.Y., S.X., Lianrong Wang and B.Z.

**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/s41586-024-07270-x.
Correspondence and requests for materials should be addressed to Bin Zhu or Longfei Wang.
Peer review information *Nature* thanks Jack Bravo, Ryan Jackson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.
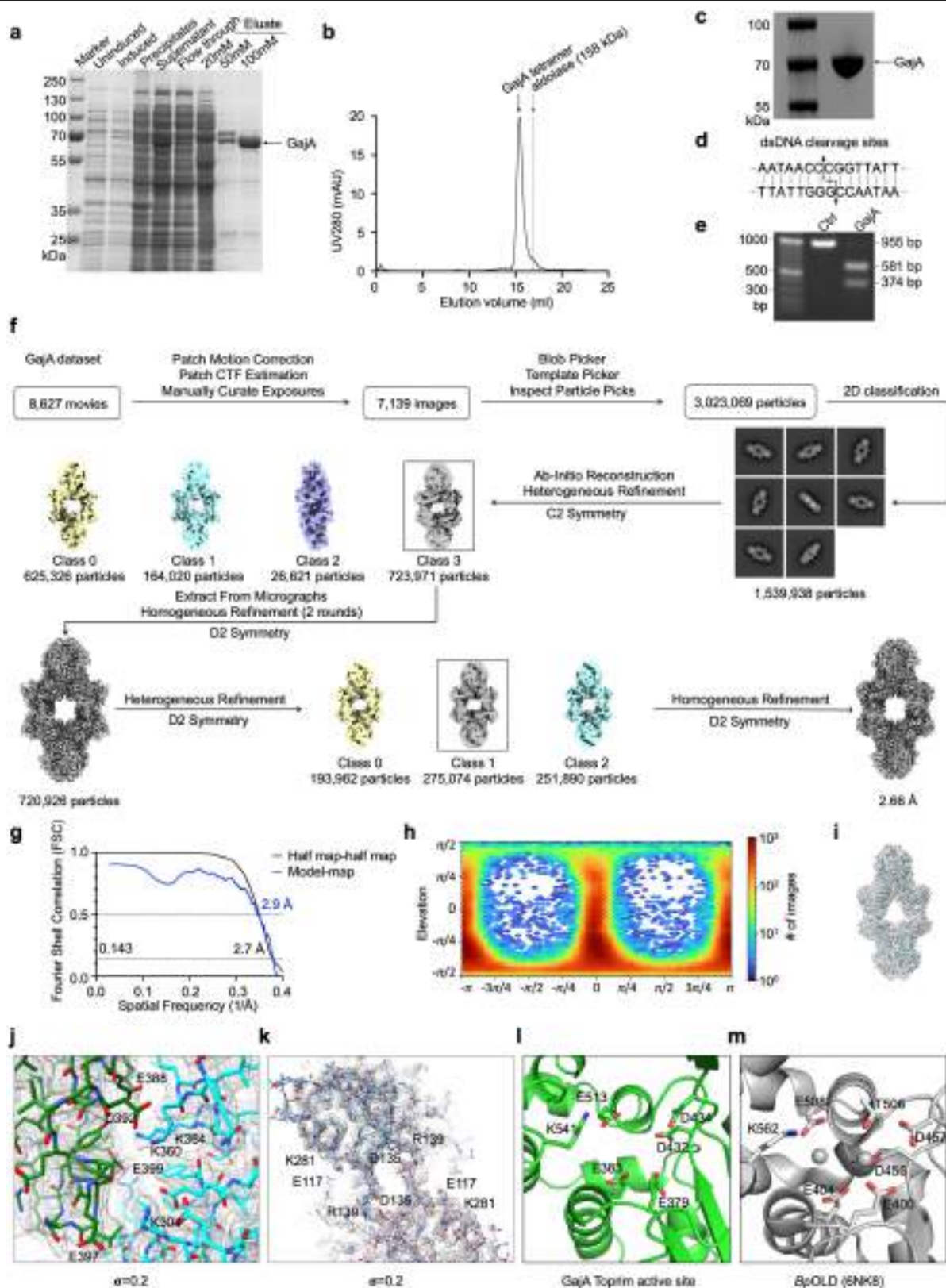
Extended Data Fig. 1 | See next page for caption.

Article

**Extended Data Fig. 1 | DNA cleavage activity and 3D reconstruction of GajA.**
a, Full-length GajA was expressed in *E.coli* BL21 (DE3) cells and was purified
using Ni-NTA agarose. SDS-PAGE showing the fractions collected at each step
of the purification with concentrations of imidazole labeled in eluate. Images
are representative of three biological replicates. **b**, The eluate was subjected to
size-exclusion chromatography (10/300 Superose 6). The molecular weight of
GajA tetramer was about 276 kDa. The retention volume of rabbit aldolase
(158 kDa) is indicated by the dotted line. **c**, The protein purity was visualized by
SDS-PAGE followed by Coomassie-blue staining. Images are representative of
three biological replicates. **d**, The recognize sequence and cleavage site of
GajA. **e**, Cleavage of pUC19-955 DNA by GajA in the presence of 1 mM Mg^{2+}. The
assays were repeated three times with similar results. **f**, Flow chart of cryo-EM
data processing and 3D reconstruction of GajA. **g**, The gold standard threshold
FSC curve for the map of GajA and the map to model FSC curve for model
refined against the overall map of GajA. **h**, The orientation distribution plot of
the 3D reconstruction of GajA. **i**, The final model of GajA tetramer fitted in the
density. **j**, Cryo-EM density validation and close-up view of Toprim/ATPase
dimer interface. The Cryo-EM densities are displayed as gray mesh. **k**, Cryo-EM
density validation and close-up view of the ATPase dimer interface. **l**, Residues
in the GajA Toprim active site. **m**, Metal coordinating residues in the *Bp*OLD
active site.

a

Bacillus cereus
Innestialibacterbartlettii
Clostridium perfringens
Streptococcus oralis
Mollicella sp.

- GajA active site
- GajA-GajA interface
- DNA binding site
- GajA-GajB interface
- ATP binding site
- Mg²⁺ binding site

b

	Phage T4	Phage T7
Vector		
GajAB		
E397K /E396K		
E388K /D392R/E529K		

c

d

e

f

g

Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | The effect on phage resistance and DNA cleavage activity of GajA mutations. a, Structure-guided alignment of GajA proteins from indicated bacteria colored according to amino acid conservation. The determined *Bacillus cereus* GajA secondary structure is displayed, and active-site and interface residues are annotated according to the key below. **b**, Plaque assay with 10-fold serial dilutions using T4 and T7 phages to infect *E. coli* B expressing the WT Gabija system and indicated mutants of GajA. Images are representative of three biological replicates. **c**, The purified GajA and GajA E117K/D135R, E397K/E399K and E388K/D392R/E529K mutants were visualized by SDS-PAGE followed by Coomassie-blue staining. **d**, The molecular weight of GajA and indicated mutants were shown on native PAGE. **e**, GajA-E399K was subjected to size-exclusion chromatography (10/300 Superose 6) (Top). The line-dashed line indicated GajA tetramer was about 276 kDa. The retention volume of rabbit aldolase (158 kDa) is indicated by the dotted line. SDS-PAGE gel at the bottom shows the fractions of GajA-E399K collected in size-exclusion chromatography. For **c**–**e**, images are representative of three biological replicates. **f**, **g**, Cleavage of pUC19-955 DNA by GajA or indicated mutants in the presence of 1 mM Mg^{2+}. The assays were repeated three times with similar results.
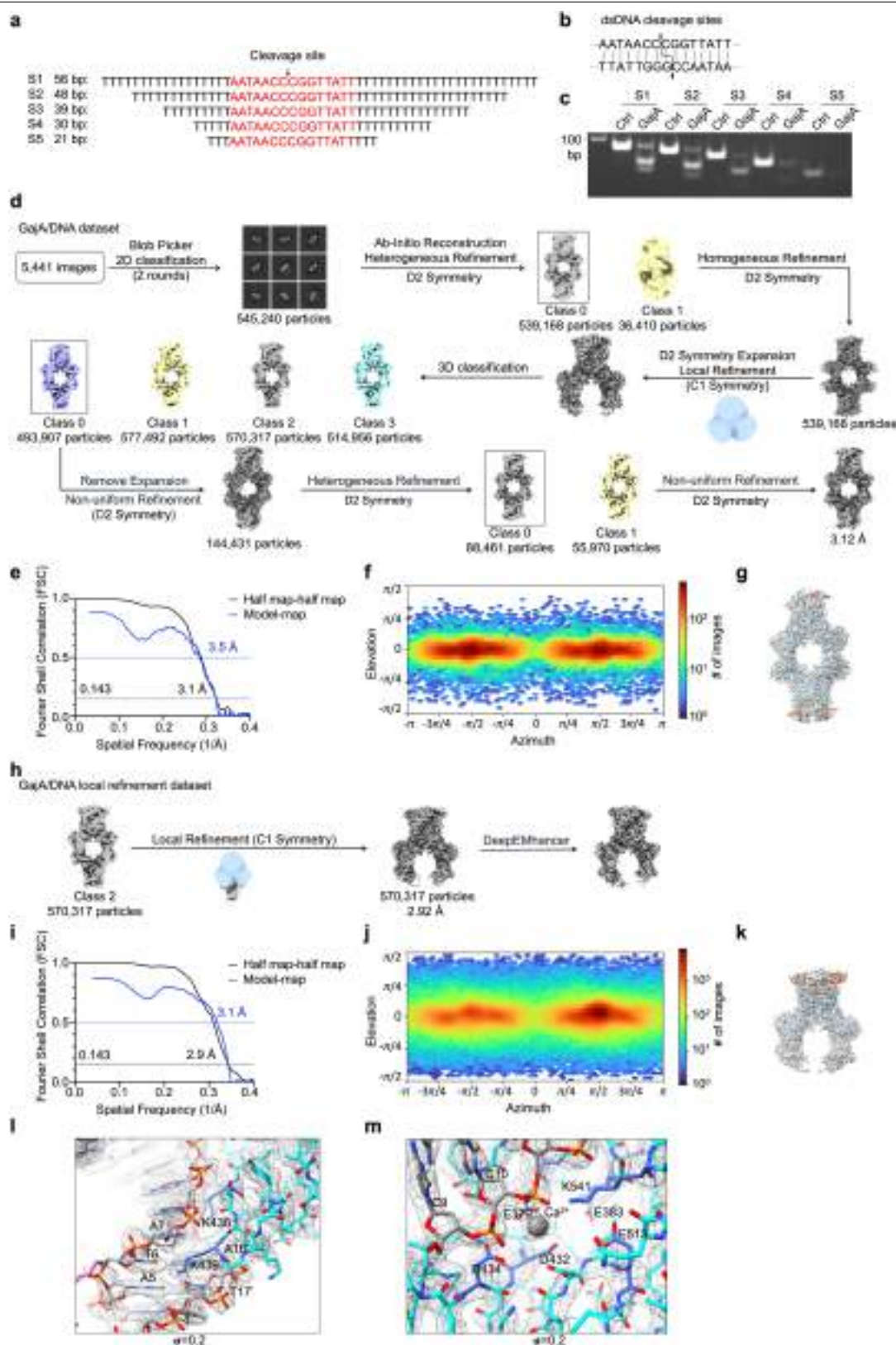
Extended Data Fig. 3 | See next page for caption.

Article

**Extended Data Fig. 3 | 3D reconstruction of the GajA/DNA complex.**
a, The sequences of designed dsDNA substrates in different lengths. **b**, The recognition sequence and cleavage sites of dsDNA substrate for GajA. **c**, Cleavage of the designed dsDNA substrates by GajA in the presence of 1 mM Mg^{2+}. The assays were repeated three times with similar results. **d**, Flow chart of cryo-EM data processing and 3D reconstruction of the GajA/DNA complex. **e**, The gold standard threshold FSC curve for the map of the GajA/DNA complex and the map to model FSC curve for model refined against the overall map of the GajA/DNA complex. **f**, The orientation distribution plot of the 3D reconstruction of the GajA/DNA complex. **g**, The final model of GajA/DNA complex fitted in the density. **h**, Local refinement of the GajA/DNA complex. **i**, The gold standard threshold FSC curve for the map of the GajA/DNA complex after local refinement and the map to model FSC curve for model refined against the overall map of the GajA/DNA complex after local refinement. **j**, The orientation distribution plot of the 3D reconstruction of the GajA/DNA complex after local refinement. **k**, The final model of GajA/DNA complex after local refinement fitted in the density. **l**, Cryo-EM density validation and close-up view of interface between KxxK motif (cyan) and bound-DNA. **m**, Cryo-EM density validation and close-up view of catalytic residues at the Toprim active site (cyan) that can interact with bound-DNA via Ca^{2+} (gray).
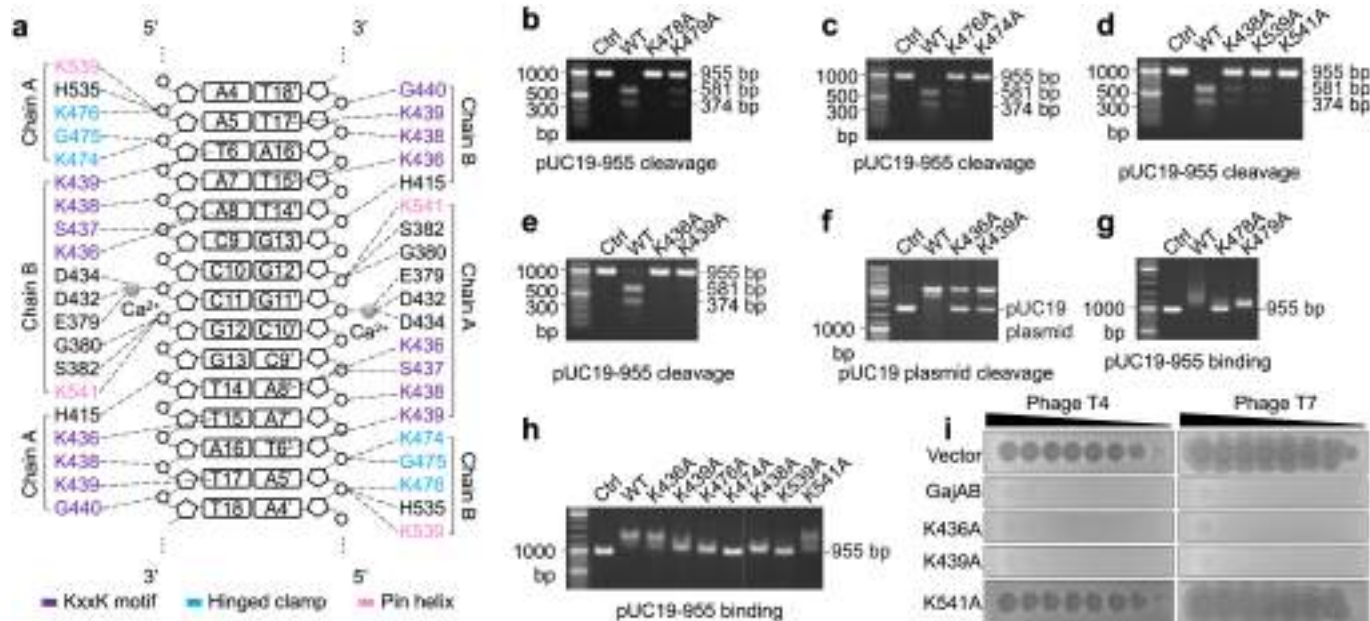
Extended Data Fig. 4 | Interaction between the GajA tetramer and dsDNA.
a, A schematic of the intermolecular interactions between GajA and the 21-bp dsDNA in the GajA/DNA complex. **b**–**e**, Cleavage of pUC19-955 DNA by GajA or indicated mutants in the presence of 1 mM Mg²⁺. **f**, Cleavage of pUC19 plasmid by GajA or indicated mutants in the presence of 1 mM Mg²⁺. **g**, **h**, Binding of pUC19-955 DNA by GajA or indicated mutants in the presence of 5 mM Ca²⁺. For **b**–**h**, the assays were repeated three times with similar results. **i**, Plaque assay with 10-fold serial dilutions using T4 and T7 phages to infect *E. coli* B expressing the WT Gabija system and indicated mutants of GajA. Images are representative of three biological replicates.

a

- GajB active site
- GajB-GajA interface

b

GajB active site

c

EcUvrD (2IS6)
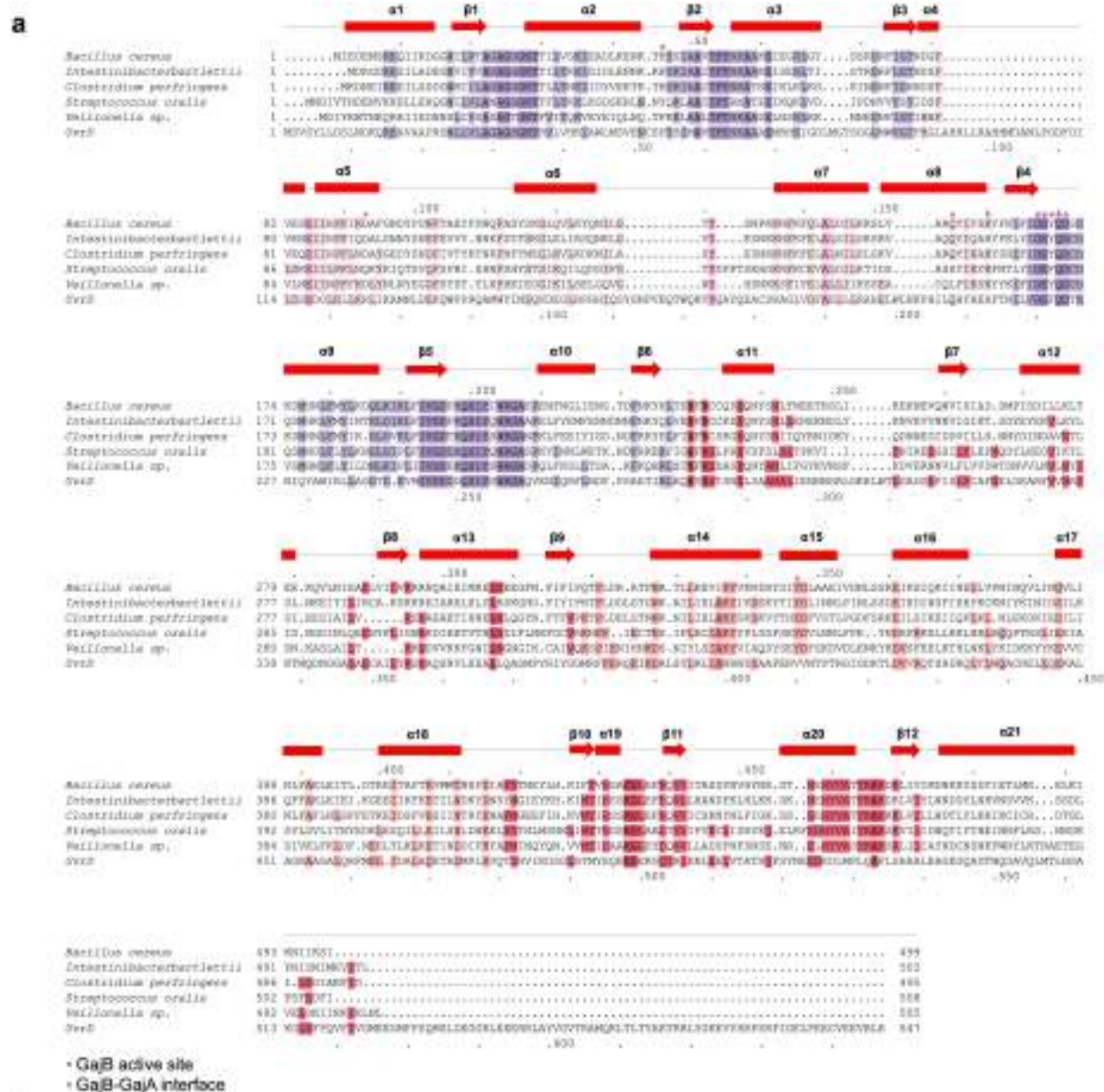
Extended Data Fig. 5 | Structure-guided alignment and active site of GajB.
a, Structure-guided alignment of GajB proteins from indicated bacteria colored according to amino acid conservation. The determined *Bacillus cereus* GajB secondary structure is displayed, and active-site and GajB-GajA interface residues are annotated according to the key below. **b**, Residues in the GajB active site. **c**, Residues in the *Ec*UvrD active site.
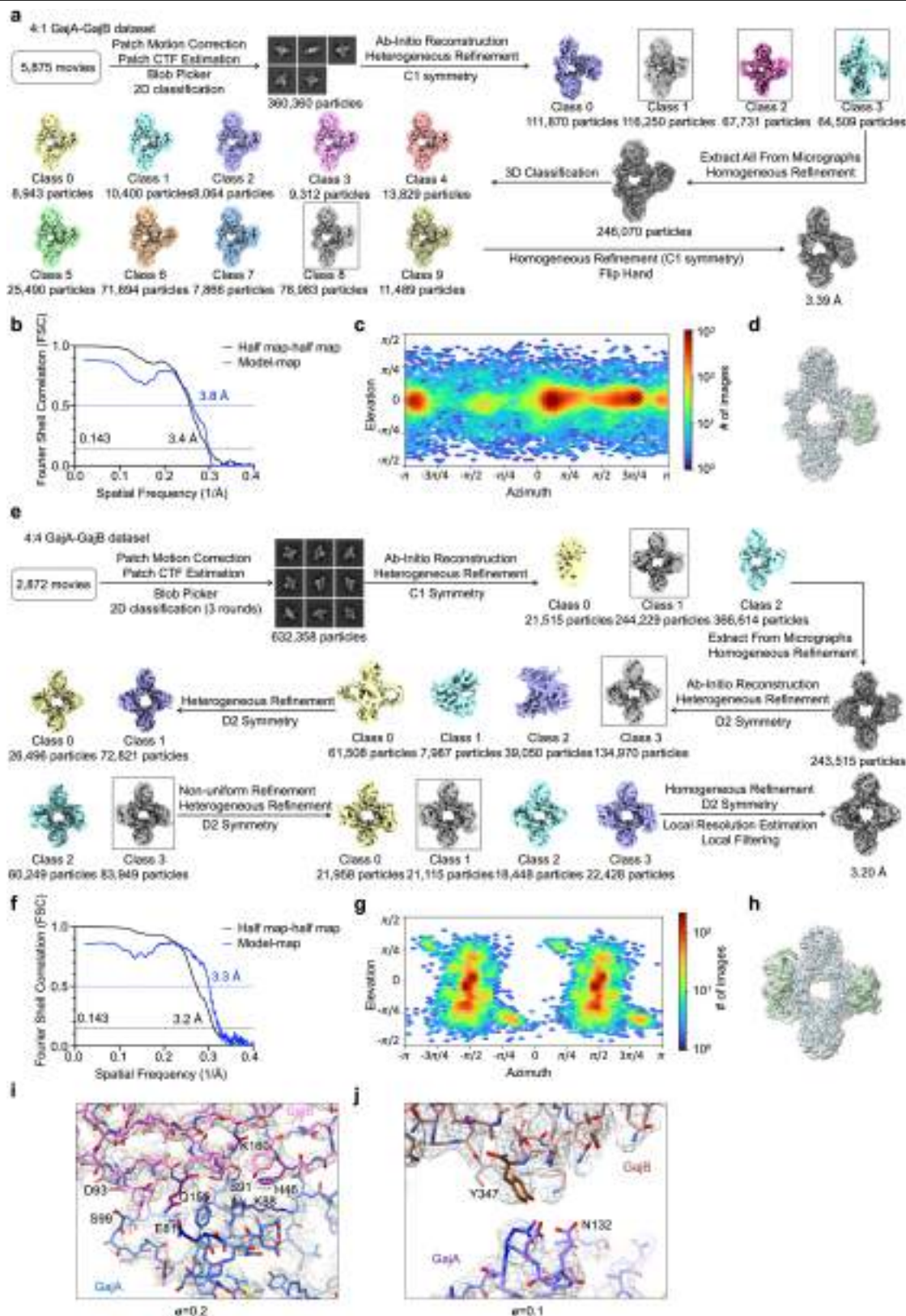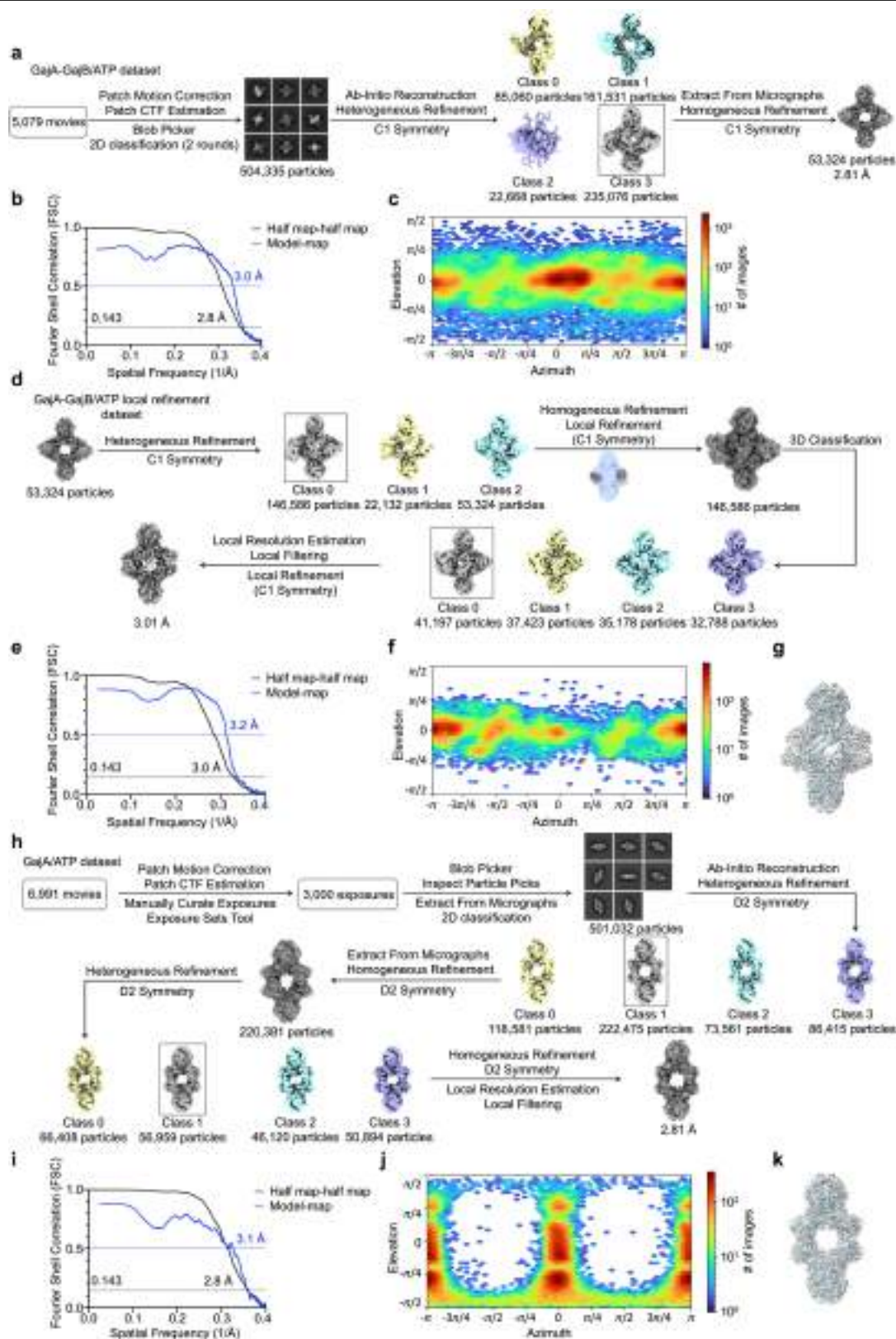
Extended Data Fig. 6 | 3D reconstruction of the 4:1 GajA-GajB complex and 4:4 GajA-GajB complex. a, Flow chart of cryo-EM data processing and 3D reconstruction of the 4:1 GajA-GajB complex. **b**, The gold standard threshold FSC curve for the map of the 4:1 GajA-GajB complex and the map to model FSC curve for model refined against the overall map of the 4:1 GajA-GajB complex. **c**, The orientation distribution plot of the 3D reconstruction of the 4:1 GajA-GajB complex. **d**, The final model of the 4:1 GajA-GajB complex fitted in the density. **e**, Flow chart of cryo-EM data processing and 3D reconstruction of the 4:4 GajA-GajB complex. **f**, The gold standard threshold FSC curve for the map of the 4:4 GajA-GajB complex and the map to model FSC curve for model refined against the overall map of the 4:4 GajA-GajB complex. **g**, The orientation distribution plot of the 3D reconstruction of the 4:4 GajA-GajB complex. **h**, The final model of the 4:4 GajA-GajB complex fitted in the density. **i, j**, Cryo-EM density validation and close-up view of GajA-GajB interfaces at 1B (**i**) and 2B (**j**) domains of GajB.

Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | The function of GajB and GajB-Q155A/Y347A. a, SDS-PAGE showing the fractions collected at each step of the purification of GajB. **b**, The eluate of GajB was subjected to size-exclusion chromatography (16/600 Superdex 200). The molecular weight of GajB was about 59.9 kDa. **c**, The protein purity of GajB was visualized by SDS-PAGE. **d**, Overlaid structures of 4:4 GajA-GajB complex and *Ec*UvrD with DNA. **e**, **f**, Superimposed structures of GajB (colored) and *Ec*UvrD in apo (gray) (**e**) or *Ec*UvrD with DNA (gray) (**f**). **g**–**i**, Structures of GajB (**g**), *Ec*UvrD in apo (**h**), and *Ec*UvrD with DNA (**i**). **j**, The diagram of various Gabija constructs. **k**, The protein expression levels of GajA and GajB in different constructs are indicated in the SDS-PAGE gel. **l**, SDS-PAGE gel showing purified GajA-GajB (His-tagged GajA and co-purified non-tagged GajB) and GajA+B (His-tagged GajA and co-purified non-tagged GajB). **m**, Plaque assay with 10-fold serial dilutions using T4 and T7 phages to infect *E. coli* B expressing the WT Gabija system and indicated mutants. Images are representative of three biological replicates. **n**, Superimposed cryo-EM structures of the 4:1 GajA-GajB complex (brown) and 4:4 GajA-GajB complex (gray). **o**, SDS-PAGE showing the fractions collected at each step of the purification of GajB-Q155A/Y347A. **p**, The eluate of GajB-Q155A/Y347A was subjected to size-exclusion chromatography (16/600 Superdex 200). The molecular weight of GajB-Q155A/Y347A was about 59.7 kDa. **q**, The protein purity of GajB-Q155A/Y347A was visualized by SDS-PAGE. **r**, GajA with GajB or GajB-Q155A/Y347A at different molar ratio was incubated and shown on native PAGE. **s**, **t**, Co-purification of His-tagged GajA and non-tagged GajB (**s**) or GajB-Q155A/Y347A (**t**) from bacteria expressing the WT GajA-GajB gene cassette under T7 bacterial promoter (schematic is shown above the gel). SDS-PAGE showing the fractions collected at each step during the purification, different concentrations of imidazole in eluate has been labeled. For **a**, **c**, **k**–**l**, **o**, **q**–**t**, images are representative of three biological replicates.

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | 3D reconstruction of the ATP-bound GajA complex and ATP-bound GajA-GajB complex. a, Flow chart of cryo-EM data processing and 3D reconstruction of the ATP-bound GajA-GajB complex. **b**, The gold standard threshold FSC curve for the map of the ATP-bound GajA-GajB complex and the map to model FSC curve for model refined against the overall map of the ATP-bound GajA-GajB complex. **c**, The orientation distribution plot of the 3D reconstruction of the ATP-bound GajA-GajB complex. **d**, Local refinement of the ATP-bound GajA-GajB complex. **e**, The gold standard threshold FSC curve for the map of the ATP-bound GajA-GajB complex after local refinemen and the map to model FSC curve for model refined against the overall map of the ATP-bound GajA-GajB complex after local refinement. **f**, The orientation distribution plot of the 3D reconstruction of the ATP-bound GajA-GajB complex after local refinement. **g**, The final model of the ATP-bound GajA-GajB complex after local refinement fitted in the density. **h**, Flow chart of cryo-EM data processing and 3D reconstruction of the ATP-bound GajA complex. **i**, The gold standard threshold FSC curve for the map of the ATP-bound GajA complex and the map to model FSC curve for model refined against the overall map of the ATP-bound GajA complex. **j**, The orientation distribution plot of the 3D reconstruction of the ATP-bound GajA complex. **k**, The final model of the ATP-bound GajA complex fitted in the density.
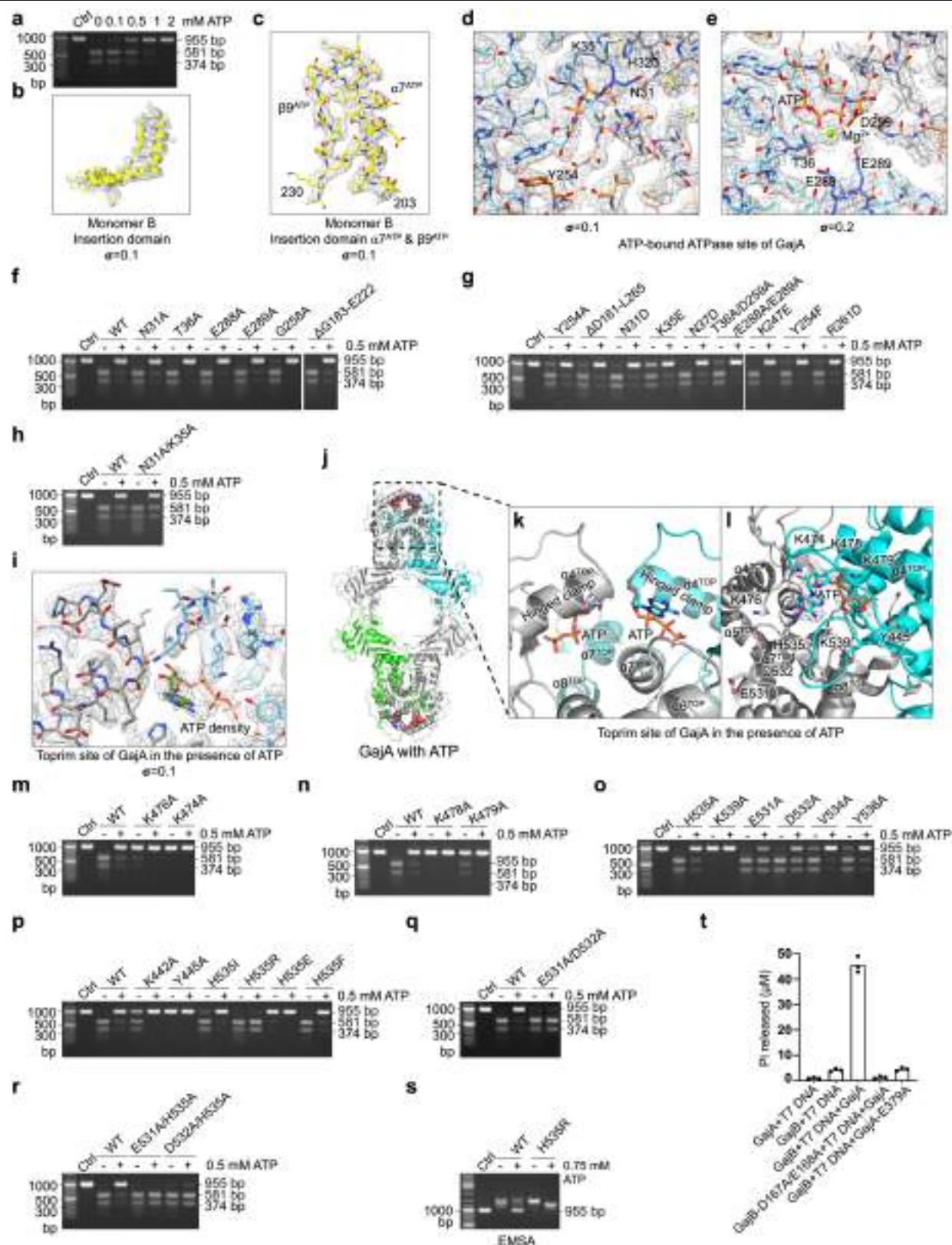
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Interaction between the GajA tetramer and ATP.
a, Titration of ATP in various concentrations on GajA endonuclease activity in the presence of 5 mM Mg^{2+}. **b**, **c**, Cryo-EM density validation of GajA tetramer and close-up view of insertion region. **d**, **e**, Cryo-EM density validation and close-up view of the ATP binding pocket at the ABC ATPase domain. **f**–**h**, Cleavage of pUC19-955 DNA by GajA or indicated mutants in the presence of 1 mM Mg^{2+} and with or without 0.5 mM ATP. **i**, Cryo-EM density validation and close-up view of Toprim domain of ATP-bound GajA. **j**, Overview of the structure of GajA in complex with ATP. **k**, **l**, Enlarged view of the ATP-binding site at the Toprim domain. Two GajA protomers are colored in cyan and gray. The mesh indicates the EM density of ATP. **m**–**r**, Cleavage of pUC19-955 DNA by GajA or indicated mutants in the presence of 1 mM Mg^{2+} and with or without 0.5 mM ATP. **s**, Binding of pUC19-955 DNA by GajA or GajA-H535R with or without 0.75 mM ATP. **t**, ATP hydrolysis assays to compare the activation of GajB or GajB-D167A/E168A by intact T7 genomic DNA or T7 genomic DNA treated by GajA or GajA-E379A. GajB-D167A/E168A is an active site mutant of GajB and GajA-E379A is an Toprim active site mutant of GajA. Data are the average of three biological replicates, with individual data points shown. For **a**, **f**–**h**, and **m**–**s**, the assays were repeated three times with similar results.

Article

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

	GajA	ATP-bound GajA	GajA/DNA	DNA-bound GajA dimer (focused refinement)	4:1 GajA-GajB	4:4 GajA-GajB	ATP/Mg^{2+}-bound GajA-GajB	ATP/Mg^{2+}-bound GajA (focused refinement)
EMDB ID	36541	37915	37916	38058	36569	36563	38071	38070
PDB ID	8JQ9	8WY4	8WY5	8X51	8JQC	8JQB	8X5N	8X5I
Data collection and processing								
Magnification	105,000	105,000	105,000		105,000	105,000	105,000	
Voltage (kV)	300	300	300		300	300	300	
Electron exposure (e–/Å^2)	50	50	50		50	50	50	
Defocus range (μm)	-1.5 to -1.8	-1.5 to -1.8	-1.5 to -1.8		-1.5 to -1.8	-1.5 to -1.8	-1.5 to -1.8	
Pixel size (Å)	0.84	0.84	0.84		0.84	0.84	0.84	
Symmetry imposed	D2	D2	D2	C1	C1	D2	C1	C1
Initial particle images (no.)	1,539,938	501,032	1,151,527	1,151,527	360,360	632,358	504,335	504,335
Final particle images (no.)	275,074	56,959	88,461	570,317	78,983	21,115	190,783	41,197
Map resolution (Å)	2.7	2.8	3.1	2.9	3.4	3.2	2.8	3.0
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.4 to 3.8	2.6 to 4.5	2.9 to 5.0	2.7 to 5.5	3.3 to 6.8	3.2 to 6.5	2.8 to 6.2	3.0 to 7.2
Refinement								
Initial model used (PDB)	AlphaFold	8JQ9	8JQ9	8JQ9	8JQ9	8JQ9	8JQ9	8JQ9
Model resolution (Å)	2.9	3.1	3.5	3.1	3.8	3.3	3.0	3.2
FSC threshold	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Model resolution range (Å)	50-2.9	50-3.1	50-3.5	50-3.1	50-3.8	50-3.3	50-3.0	50-3.2
Map sharpening B factor (Å^2)	-104.7	-80.9	-106.9	-87.3	-60.1	-59.9	-65.9	-56.0
Model composition								
Non-hydrogen atoms	14832	14864	16038	8101	18899	31664	22192	17006
Protein residues	1816	1804	1770	885	2315	3876	2708	2063
Nucleotide residues	0	0	76	42	0	0	0	0
Ligands	0	4	4	2	0	0	8	8
B factors (Å^2)								
Protein	48.19	33.52	89.44	44.24	116.4	137.01	58.47	86.34
Nucleotide			99.51	43.09				
Ligand		70.66	111.83	23.64			42.23	78.36
R.m.s. deviations								
Bond lengths (Å)	0.004	0.006	0.006	0.007	0.004	0.006	0.006	0.006
Bond angles (°)	0.760	1.068	1.175	1.243	0.950	1.178	1.071	1.059
Validation								
MolProbity score	1.44	1.31	1.46	1.42	1.59	1.81	1.31	1.35
Clashscore	3.18	2.35	2.88	1.70	4.24	5.71	2.32	2.20
Poor rotamers (%)	0	0	0	0	0.24	0.31	0	0
Ramachandran plot								
Favored (%)	95.28	95.75	94.39	92.31	94.25	91.73	95.72	94.92
Allowed (%)	4.72	4.25	5.15	7.35	5.62	7.47	4.14	4.79
Disallowed (%)	0	0	0.46	0.34	0.13	0.81	0.15	0.29

This table shows the statistics for the cryo-EM structures of the Gaibja system from *Bacillus cereus* VD045.

nature portfolio

Corresponding author(s):	Longfei Wang, Bin Zhu
Last updated by author(s):	Feb 20, 2024

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	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	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)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	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-3.2.0.4776REL.
Data analysis	AlphaFold-2.3.2, ColabFold-1.5.5, JPred-4, cryoSPARC-4.4.1, DeepEMhancer-0.14, Coot-0.9.8.92 EL, PHENIX-1.20.1, PyMOL-2.5.0, UCSF ChimeraX-1.7, MolProbity-4.2.

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

The cryo-EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession codes: EMDB-36541, EMDB-37915, EMDB-37916, EMDB-38058, EMDB-36569,EMDB-36563,EMDB-38071,EMDB-38070.
The coordinates for the GajA have been deposited in the Protein Data Bank (PDB) under accession codes 8JQ9, 8WY4, 8WY5, 8X51, 8JQC,8JQB,8X5N,8X5I.

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

Life sciences study design

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

Sample size	No sample size calculations were performed as this is not relevant to biochemical, structural or phage plaque assays.
Data exclusions	In cryo-EM data processing, we discarded "junk" particles that could not be classified. This is a widely used and accepted practice in the cryo-EM field. No other data were excluded from the analyses.
Replication	All data were performed with replicates as described. Phage plaque assays or biochemical assays were performed with three independent biological replicates. Figures show one example of a biologically independent replicate and other replicates not shown showed similar results within expected variation. The numbers of replication were described in figure legends.
Randomization	Division of particles into random halves is automatically performed during 3D reconstruction by cryoSPARC.
Blinding	Blinding is not relevant for protein structure determination and functional assays because these results are not subjective.

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	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Clinical data
☒ ☐	Dual use research of concern
☒ ☐	Plants

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A