Article

Structure and mechanism of the Zorya anti-phage defence system

https://doi.org/10.1038/s41586-024-08493-8

Received: 17 December 2023

Accepted: 4 December 2024

Published online: 11 December 2024

Open access

Check for updates

Haidai Hu[1,10]✉, Philipp F. Popp[2,10], Thomas C. D. Hughes[3], Aritz Roa-Eguiara[1], Nicole R. Rutbeek[1], Freddie J. O. Martin[1], Ivo Alexander Hendriks[4], Leighton J. Payne[3], Yumeng Yan[1], Dorentina Humolli[5], Victor Klein-Sousa[1], Inga Songailiene[1,6], Yong Wang[7], Michael Lund Nielsen[4], Richard M. Berry[8], Alexander Harms[5], Marc Erhardt[2,9]✉, Simon A. Jackson[3]✉ & Nicholas M. I. Taylor[1]✉

Zorya is a recently identified and widely distributed bacterial immune system that protects bacteria from viral (phage) infections. Three Zorya subtypes have been identified, each containing predicted membrane-embedded ZorA–ZorB (ZorAB) complexes paired with soluble subunits that differ among Zorya subtypes, notably ZorC and ZorD in type I Zorya systems[1,2]. Here we investigate the molecular basis of Zorya defence using cryo-electron microscopy, mutagenesis, fluorescence microscopy, proteomics and functional studies. We present cryo-electron microscopy structures of ZorAB and show that it shares stoichiometry and features of other 5:2 inner membrane ion-driven rotary motors. The $ZorA_5B_2$ complex contains a dimeric ZorB peptidoglycan-binding domain and a pentameric α-helical coiled-coil tail made of ZorA that projects approximately 70 nm into the cytoplasm. We also characterize the structure and function of the soluble Zorya components ZorC and ZorD, finding that they have DNA-binding and nuclease activity, respectively. Comprehensive functional and mutational analyses demonstrate that all Zorya components work in concert to protect bacterial cells against invading phages. We provide evidence that ZorAB operates as a proton-driven motor that becomes activated after sensing of phage invasion. Subsequently, ZorAB transfers the phage invasion signal through the ZorA cytoplasmic tail to recruit and activate the soluble ZorC and ZorD effectors, which facilitate the degradation of the phage DNA. In summary, our study elucidates the foundational mechanisms of Zorya function as an anti-phage defence system.

Bacteria face frequent bacteriophage (phage) attacks and have evolved diverse defence strategies, including the restriction-modification and CRISPR–Cas systems[3–5]. Anti-phage defence systems have key roles in modulating phage–bacteria population dynamics and have biotechnological potential[6]. Recent studies have identified many new anti-phage defence systems, with further work revealing how some systems are activated by directly sensing phage-encoded proteins, or through indirect mechanisms[1,2,7–10]. As phage invasion initiates with cell envelope interactions, some defence systems might detect changes in the envelope as early infection signals. However, such defence mechanisms have not yet been identified.

Among newly discovered anti-phage defence systems, Zorya systems are widespread[1,2]. There are three known Zorya types, each encoding ZorA and ZorB, which are thought to form membrane-embedded complexes related to the proton-driven flagellar stator unit MotAB[11–13]. Moreover, each type encodes one or more cytosolic proteins of unknown function, namely ZorC and ZorD for type I Zorya systems. Most anti-phage defence systems encoding membrane-associated proteins are thought to function by disrupting or depolarizing the host membrane, leading to cell death or dormancy before phage replication completes, a mechanism called abortive infection[14]. A similar membrane-depolarization mechanism for ZorAB has been proposed[1], but it has not been ruled out that ZorAB could instead act as the sensor of infection.

Using single-particle cryo-electron microscopy (cryo-EM), mutagenesis, functional assays, proteomics and total internal reflection fluorescence (TIRF) microscopy, we decipher several key aspects of the Zorya defence mechanism. We identified that ZorA and ZorB form a unique 5:2 proton motive force (PMF)-driven motor complex with a long intracellular tail and propose that it acts as a phage infection sensor and signal transduction complex. After phage perturbation of the cell envelope, the peptidoglycan-binding domain (PGBD) of ZorB anchors the complex to the cell wall, and proton flow drives ZorA and its tail to rotate around ZorB. This rotation induces recruitment of the

[1]Structural Biology of Molecular Machines Group, Protein Structure & Function Program, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [2]Institute of Biology/Molecular Microbiology, Humboldt-Universität zu Berlin, Berlin, Germany. [3]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand. [4]Proteomics program, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [5]Institute of Food, Nutrition and Health, ETH Zurich, Zurich, Switzerland. [6]Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [7]College of Life Sciences, Zhejiang University, Hangzhou, China. [8]Department of Physics and Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford, UK. [9]Max Planck Unit for the Science of Pathogens, Berlin, Germany. [10]These authors contributed equally: Haidai Hu, Philipp F. Popp. ✉e-mail: haidai.hu@cpr.ku.dk; marc.erhardt@hu-berlin.de; simon.jackson@otago.ac.nz; nicholas.taylor@cpr.ku.dk

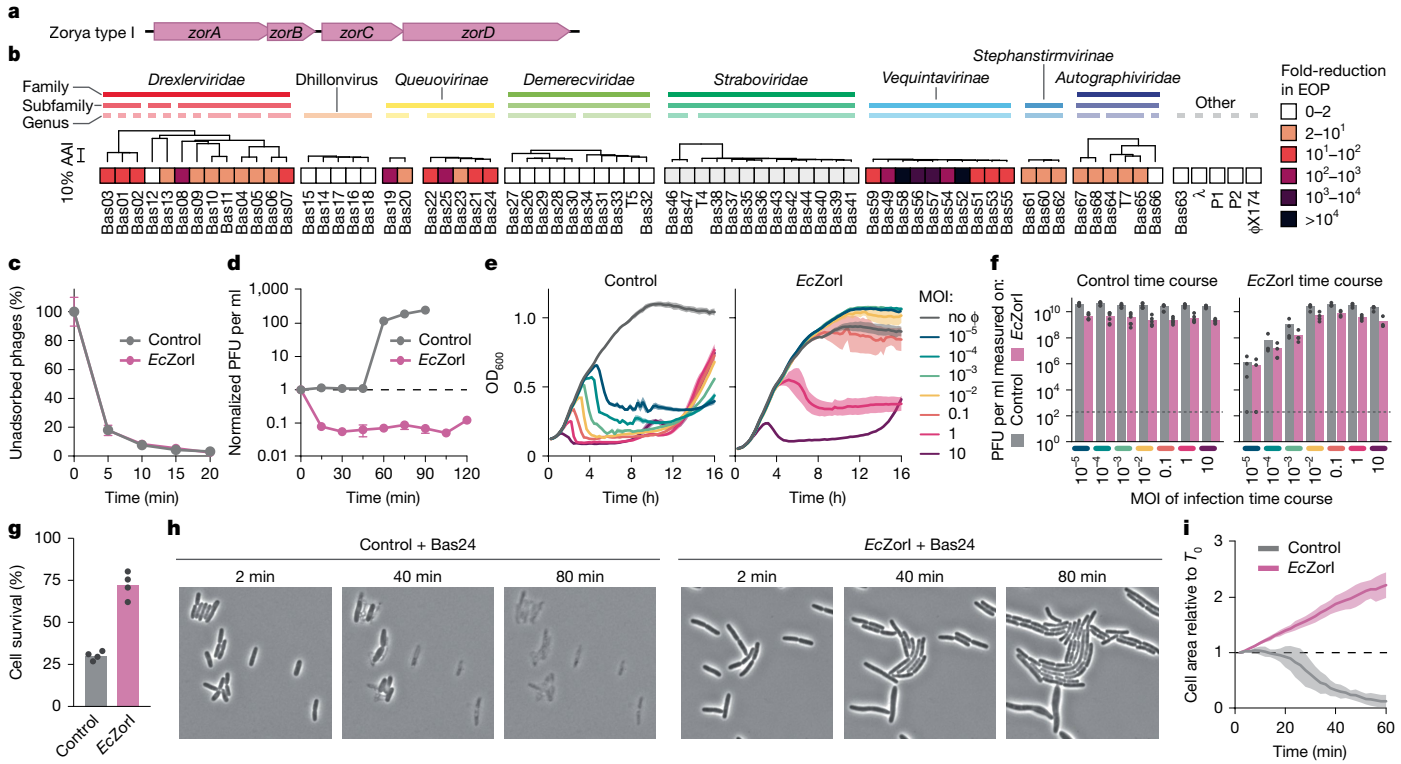


Fig. 1 | Zorya has broad activity against phages through a direct immunity mechanism. a, Schematic of the *Ec*ZorI operon. **b**, *Ec*ZorI defence against diverse *E. coli* phages, determined using efficiency of plaquing (EOP) assays. AAI, the average amino acid identity between proteins encoded by each phage (providing an estimate of the relatedness between phages). **c**, Adsorption of phage Bas24 onto *E. coli* cells possessing or lacking *Ec*ZorI. **d**, One-step phage growth curve for phage Bas24 infection of *E. coli*, with or without *Ec*ZorI, normalized to the plaque-forming units (PFU) per ml at the initial timepoint. **e**, Infection time courses for liquid cultures of *E. coli*, with and without *Ec*ZorI, infected at different MOIs of phage Bas24. ϕ, phage. **f**, Phage titres at the end timepoint for each sample from the infection time courses in **e**, measured as PFU per ml on indicator lawns of *E. coli* either without (control) or with *Ec*ZorI. The limit of detection (LOD) is shown by dotted lines. **g**, The survival of *E. coli* cells, lacking or possessing *Ec*ZorI, that were infected at an MOI of 5 with Bas24. **h**, Time-lapse, phase-contrast microscopy of *E. coli* cells with and without *Ec*ZorI infected with Bas24 at an MOI of 5. **i**, Quantification of the time-lapse microscopy in **h**, displaying the measured cell area relative to the initial timepoint. For **b**–**g**, data are the mean of at least three biological replicates (datapoints indicate replicates) and error bars (**c** and **d**) or shaded regions (**e**) represent the standard s.e.m. For **i**, data are mean ± s.d., derived from independent biological triplicates.

soluble effectors ZorC and ZorD, which have DNA-binding and nuclease activities, leading to the local degradation of invading phage DNA to facilitate direct (non-abortive) defence.

Zorya protects through a direct mechanism

Type I Zorya systems are widely distributed across Gram-negative phyla (Extended Data Fig. 1a–c). We therefore cloned the complete *Escherichia coli* NCTC 9026 type I Zorya system (*Ec*ZorI), including its native promoter, into a low-copy plasmid and used a heterologous *E. coli* strain[15] to examine anti-phage defence. The *Ec*ZorI system provided anti-phage activity against a diverse range of phages, but some phage families tested were unaffected (Fig. 1a,b). *Ec*ZorI did not impair the adsorption of phages to host cells (Fig. 1c) but acted subsequently to prevent phage replication and burst (Fig. 1d). However, *Ec*ZorI did not defend against plasmids introduced by conjugation or transformation (Extended Data Fig. 1d,e), suggesting that some aspect of phage infection, other than the mere introduction of foreign DNA into the cell, triggers Zorya activity.

For population-level defence by *Ec*ZorI in liquid cultures infected at different multiplicities of infection (MOIs), each phage tested affected the control populations to differing extents; however, population growth in the *Ec*ZorI samples was generally unaffected at low (<0.1) MOIs and, in some cases, also at high (>1) MOIs (Fig. 1e and Extended Data Fig. 1f,g). Importantly, the growth kinetics at early timepoints did not reveal any premature host population collapse or delayed growth for cells expressing *Ec*ZorI compared with the negative controls (Extended Data Fig. 1f). *Ec*ZorI also reduced the levels of phages detectable at the end timepoints in most infected cultures (Fig. 1f and Extended Data Fig. 1h). For a single synchronized round of Bas24 infection (at MOI 5), *Ec*ZorI increased the survival rate of infected cells compared with the control populations (Fig. 1g). Together, these results indicate that *Ec*ZorI acts through a direct rather than abortive infection mechanism. This finding was confirmed using single-cell time-lapse microscopy, where cells expressing *Ec*ZorI continued to replicate after exposure to Bas24, whereas cells lacking *Ec*ZorI lysed within 60 min (Fig. 1h,i and Supplementary Videos 1 and 2).

Zorya contains a ZorA$_5$B$_2$ complex

To investigate how ZorA and ZorB support direct defence by Zorya, we purified recombinantly expressed *Ec*ZorAB complex from cell membranes using lauryl maltose neopentyl glycol (LMNG) detergent (Fig. 2a and Extended Data Fig. 2a). Visualized by negative-stain EM, *Ec*ZorAB contains a head domain attached to a long tail, measuring approximately 700 Å (Fig. 2b). We then resolved the *Ec*ZorAB cryo-EM structure to an overall resolution of 2.7 Å, revealing an oligomeric assembly of five ZorA and two ZorB subunits (Fig. 2c–f, Extended Data Fig. 2b–f and Extended Data Table 1). The 5:2 stoichiometry is supported by quantitative mass spectrometry (MS) analyses of *Ec*ZorI-expressing cells and conserved with the flagellar stator unit MotAB complex (MotA$_5$B$_2$) and similar rotary motors[16,17] (Extended Data Fig. 2g,h and Supplementary Table 1). Overall, *Ec*ZorAB comprises four structural layers: a

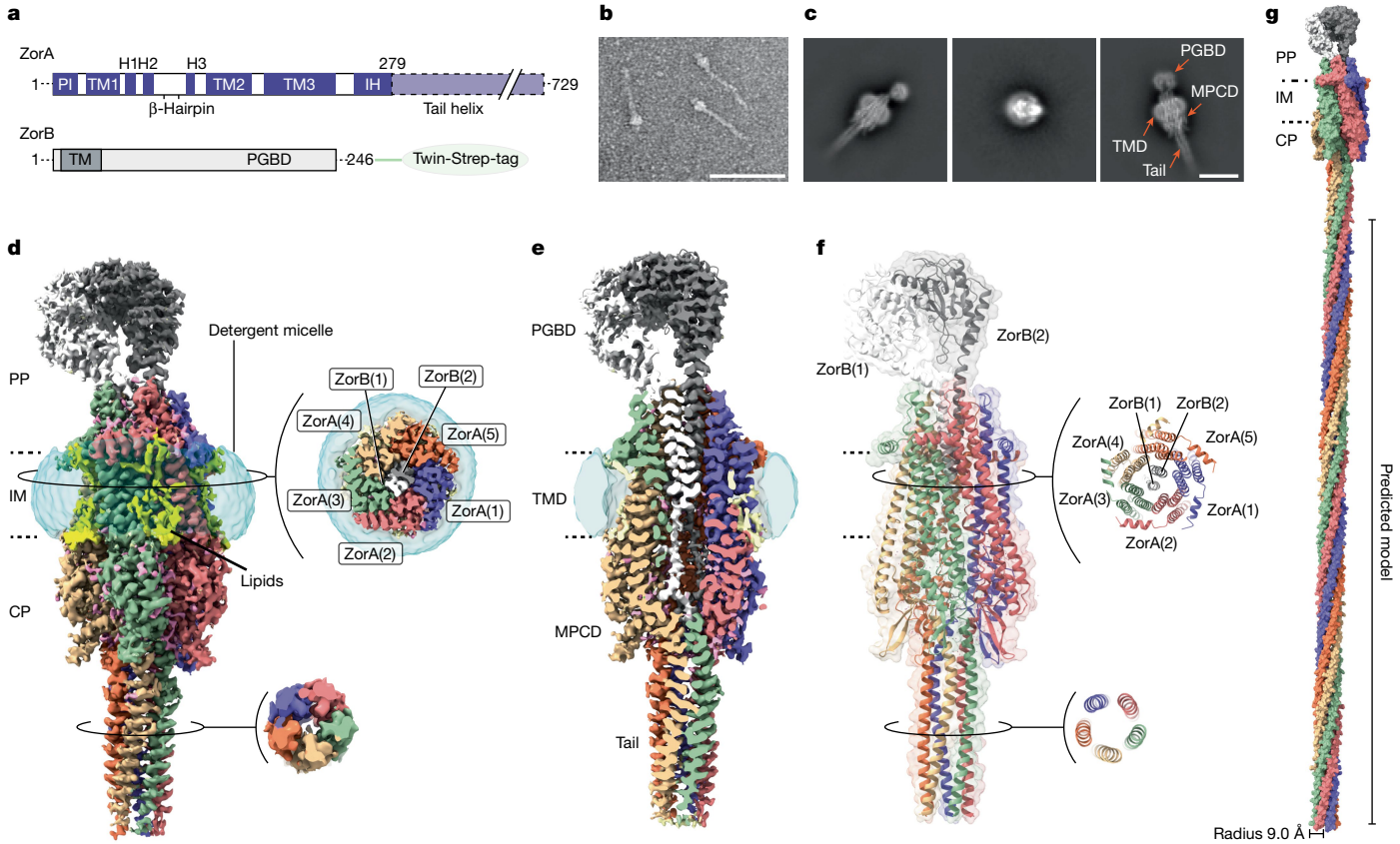


Fig. 2 | Cryo-EM analysis of *Ec*ZorAB and its architecture. a, Schematic of *Ec*ZorA and *Ec*ZorB. **b**, Negative-stain EM image of purified *Ec*ZorAB particles. Scale bar, 1,000 Å. **c**, Representative high-resolution two-dimensional classes of *Ec*ZorAB images from cryo-EM. Domain architectures of the *Ec*ZorAB complex are shown. Scale bar, 100 Å. **d**, Cryo-EM map of *Ec*ZorAB. Five ZorA subunits (purple, salmon, light green, tan and coral) surround two ZorB subunits (white and dark grey) viewed from the plane of the membrane. Membrane-bound lipids are shown in yellow. The detergent micelle is shown as a translucent surface representation in cyan. The dashed lines show inner membrane boundaries. Two cross-section views of the *Ec*ZorAB TMD and tail are shown. **e**, Cross-section view of the EM density map from the plane of the membrane. **f**, Ribbon model representation of *Ec*ZorAB, with two cross-section views of the model shown. **g**, Composite model of the *Ec*ZorAB whole complex, represented as a surface model. The radius of the ZorA tail is indicated. CP, cytoplasm; H, helix; IM, inner membrane; PP, periplasm. Images in **b** are representative of at least three replicates.

predicted PGBD (ZorB Thr47–Leu246), transmembrane domain (TMD), membrane-proximal cytoplasmic domain (MPCD; spanning ZorA residues Gly48–Leu127 and Lys207–Ser222) and a tail-like structure formed by the ZorA C-terminal region (ZorA Gly223–Thr729) extending into the cytoplasm (Fig. 2a,c,e).

The periplasmic domain exhibited flexibility relative to the TMD; accordingly local refinement was used to improve the resolution to 3.5 Å, resolving a dimerized ZorB PGBD (Extended Data Fig. 2i–k). The flexibility of the ZorA tail prevented its complete three-dimensional reconstruction, with our cryo-EM map providing density for only the first 56 tail residues, despite the negative-stain EM and MS analyses confirming that the purified complex contained intact full-length ZorA (Fig. 2b–f). Secondary-structure prediction revealed a preference for the tail to adopt α-helical structures (Extended Data Fig. 3a), suggesting that the remaining ZorA tail probably continues the experimentally observed coiled-coil with a right-handed super-helical twist. On the basis of these observations, we constructed an idealized full-length ZorAB model in which the ZorA tail forms a helical bundle projecting 70 nm into the cytoplasm, with a helical pitch of 328 Å and a radius of 9.0 Å (Fig. 2g and Extended Data Fig. 3c–e).

ZorA$_5$B$_2$ is a PG-binding H$^+$-driven motor

On the periplasmic side, the C-terminal PGBDs of the ZorB subunits form a homodimer, with the dimerization interface composed of α3 and β5 from each monomer, driven mainly by van der Waals forces and electrostatic interactions. Furthermore, a C-terminal loop of ZorB caps the side of the dimerization interface (Fig. 3a,b and Extended Data Fig. 4a,b). Each monomer contains two disulfide bridges, potentially contributing to the stability and rigidity of this domain: bridges connect α1 to the β1–β2 loop and α3 to the C-terminal loop (Extended Data Fig. 4b,c). The overall ZorB dimer structure resembles that of the periplasmic domain of MotB of the flagellar stator unit and of other peptidoglycan (PG)-binding proteins[18–20] (Extended Data Fig. 4b,d). MotAB is kept in an inactive state by the MotB 'plug' regions, which are connected through a long linker to the PGBD and inhibit ion flux and rotation of MotA around MotB (Extended Data Fig. 4e,i–m). Only after incorporation of MotAB into the flagellar motor, the MotB plug is released and the PGBDs dimerize to enable PG binding[13]. By contrast, the ZorB PGBDs are already dimerized, and the PGBDs are fused without a linker to the ZorB transmembrane (TM) helices through α1 (Fig. 3a). Mutations of key residues predicted to be involved in the ZorB PGBD dimer interface, PG binding (Tyr151, Asn152, Leu155 and Arg159) and the disulfide bridges abolished Zorya-mediated phage defence, as did truncations of the C-terminal loop region (Fig. 3h and Extended Data Figs. 2k and 6a). In vitro pull-down assays demonstrated that purified ZorAB and purified ZorB PGBD bind to PG, whereas corresponding PG-binding-site mutants displayed reduced binding (Extended Data Fig. 4f–h,q). The ZorB TM helices are asymmetrically surrounded by five ZorA subunits, each containing three TM helices (TM1–TM3).

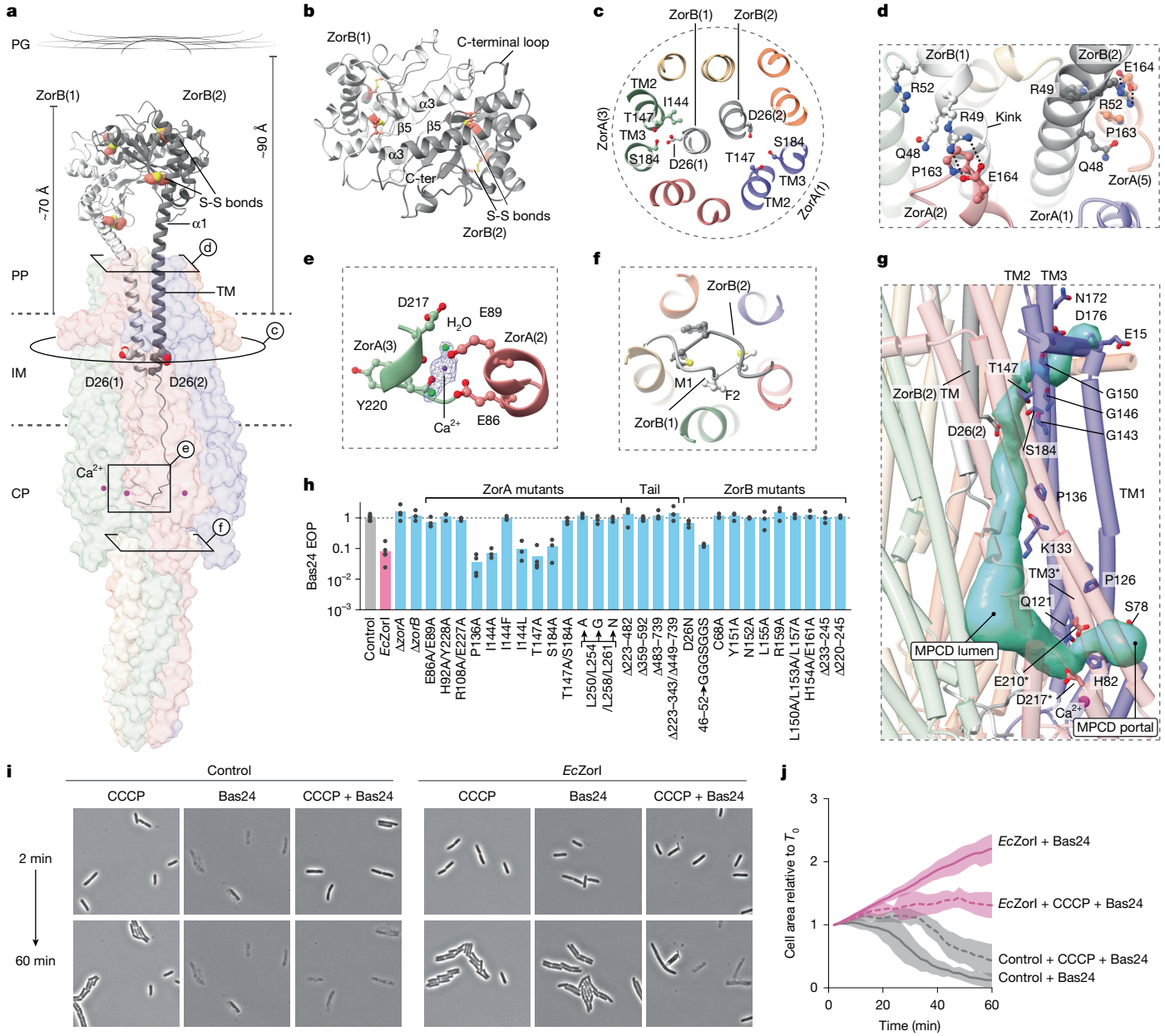


Fig. 3 | ZorAB is a PG-binding, proton-driven motor. a, *Ec*ZorAB viewed from the plane of the membrane, with ZorB shown as ribbons (black and white) and ZorA shown as a translucent surface. The distance between the inner membrane and PG layer in *E. coli* is approximately 90 Å (ref. 32). The cysteine residues from the two disulfide bridges in the ZorB PGBD are indicated and shown as spheres. The Asp26 residues from both ZorB TM helices are indicated and shown. **b**, Top view of the ZorB PGBD. **c**, Cross-section view of the ZorAB TMD, showing ZorB Asp26 and the surrounding residues. **d**, Magnified view of the interactions between ZorB and ZorA at the domain assembly interface in the periplasmic space. **e**, The Ca^{2+}-binding site. EM densities are only overlapped on Ca^{2+} ion, and the two water molecules. **f**, Magnified view of the interactions between the ZorB N terminus and the ZorA tail. **g**, Ion-translocation pathway (semitransparent surface representation in light blue) in ZorAB. Residues along the ion-permeation pathway and from the ion-selectivity filter are shown. Each asterisk indicates residues or structural elements from the neighbouring ZorA subunit. **h**, The effects of ZorA and ZorB mutations on *Ec*ZorI-mediated anti-phage defence against Bas24, as measured using EOP assays. Data are the mean of at least three replicates (datapoints indicate replicates) and are normalized to the control samples lacking *Ec*ZorI. ZorB(46–52>GGGSGGS), replacement of ZorB residues 46–52 with a GGGSGGS linker. Data for additional phages are provided in Extended Data Fig. 6a. **i**, Time-lapse phase-contrast microscopy analysis of *E. coli* cells expressing empty vector control or *Ec*ZorI with or without exposed to Bas24 at an MOI of 5 in the presence or absence of 30 μM CCCP. **j**, Quantification of the time-lapse microscopy images in **i**, displaying the measured cell area relative to the initial timepoint images. For **j**, data are mean ± s.d., derived from independent biological triplicates.

ZorA TM2 and TM3 are lined directly against the ZorB TM helices, while TM1 is peripheral and faces the lipid bilayer. Lipid densities are observed around the ZorAB TMD that probably stabilize the TMD (Fig. 2d and Extended Data Fig. 2l). The ZorAB TMD is structurally related to that of 5:2 ion-driven prokaryotic rotary motors, including possessing the universally conserved and mechanistically essential aspartate residue, Asp26 in ZorB. One ZorB Asp26 is engaged with ZorA chain 3 (ZorA(3)) through contact with TM2 Thr147 and TM3 Ser184 and the other Asp26 is unengaged and points towards a lumen enclosed by the ZorA MPCD (Fig. 3a,c). The interaction modes of these two Asp26 are the same as those in the inactive state of MotAB[11], suggesting a similar conformational state and rotary mechanism (Extended Data Fig. 4i–m). Despite the lack of a MotB-like 'plug' in ZorB, there are several features potentially blocking the rotation of ZorA around ZorB in this state: ZorA(2)

Pro163 induces a kink in ZorB(1) α1 near residue Val46, and two salt bridges, ZorA(2) Glu164–ZorB(1) Arg49 and ZorA(5) Glu164–ZorB(2) Arg52, and several polar interactions are located at the ZorAB periplasmic assembly interface (Fig. 3a,d). Replacing ZorB residues 46–52 with a GGGSGGS linker (to abolish the predicted rotational blockages) and then generating a cryo-EM reconstruction revealed that, for this mutant, the densities for the ZorB TM helices could not be resolved and the ZorB PGBD density is poor, suggesting that ZorA is free to rotate around ZorB (through Brownian motion) in the mutant (Extended Data Fig. 4n–p,r,s).

On the cytoplasmic side, the ZorA TMD and MPCD are connected by TM1 and TM3, the intracellular segments of which are joined by three vertical helices (H1–H3) and a β-hairpin motif. H3 is less ordered owing to the presence of two proline residues, Pro126 and Pro136 (Extended Data Fig. 5a,b). We found five strong spherical densities in the ZorA MPCD, each coordinated by the mainchain carboxylate groups of Asp217 and Tyr220 from the end of TM3, and the side chains of Glu86 and Glu89 from the adjacent subunit, as well as two well-resolved water molecules. On the basis of the strongly negative electrostatic environment and the surrounding coordinating residues, we assigned these densities to Ca^{2+}, bridging the MPCD of two adjacent ZorA subunits and linking ZorA TM3 to its intracellular helix (Fig. 3a,e and Extended Data Fig. 5c). We also observed a water-filled ion-permeation pathway connecting the periplasmic space through the unengaged ZorB Asp26 to the cytosol (Fig. 3g). On the periplasmic side, a cavity lined by several negatively charged residues probably attracts incoming ions. Moving towards the cytoplasmic side, ZorA residues Thr147 and Ser184 resemble an ion-selectivity filter[21] that controls ion access from the periplasm to ZorB Asp26 (Fig. 3c,g). The absence of the additional polar residues in the ion-selectivity filter that is strictly required for sodium coordination[21] indicates that ZorAB is probably a proton-driven motor (Extended Data Fig. 4i,k,m). The pathway extends from ZorB Asp26 in the direction of the cytoplasm to the inner lumen encircled by the ZorA MPCD, where we found a highly hydrated lateral portal that could facilitate ion exit (Fig. 3g).

We next mutated residues along the ion-permeation pathway to probe their role in Zorya defence. ZorB Asp26 is essential for all models of ion translocation and motor rotation and its mutation to asparagine abolishes Zorya defence. In the ion-selectivity filter, mutation of ZorA Thr147 or Ser184 to alanine did not impair Zorya activity against Bas24 but did against other phages, whereas the double-mutant ZorA(T147A/S184A) is non-functional against all of the phages tested. Alanine substitution of ZorA Pro136, which creates a kink in the ZorA MPCD H3 helix, resulted in increased defence activity against some phages. Furthermore, increasing the side-chain size and rigidity of Ile144, near ZorB Asp26, which would sterically hinder ZorA from rotating around ZorB, leads to non-functional Zorya (Fig. 3c,h and Extended Data Fig. 6a,c–e). We further confirmed the necessity of the PMF-driven ZorAB motor function for Zorya-mediated phage protection by performing single-cell time-lapse microscopy in the presence or absence of the PMF-dissipating protonophore carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP). The addition of CCCP did not prevent the growth of cells expressing either *Ec*ZorI or the empty vector control in the absence of phage but did impair *Ec*ZorI-mediated protection against Bas24 (Fig. 3i,j and Supplementary Videos 3–6). These observations support the idea that PMF-driven rotation of ZorA around ZorB is essential for Zorya anti-phage defence.

ZorAB tail controls anti-phage defence

One of the most notable features of the ZorAB complex is its long tail-like structure (Fig. 2b,g). Within the ZorA MPCD, ZorB N-terminal residues Met1 and Phe2 intertwine and hydrophobically block the entrance of the tail (Fig. 3a,f). On the outside of the tail, residue Arg108 from the β-hairpin motif forms a salt bridge with Glu227, and His92 makes electrostatic contact with the hydroxyl group of Tyr228 (Extended Data Fig. 5c). Disrupting these interactions leads to loss of defence against Bas24 (Fig. 3h). Inside the tail, Leu250, Leu254, Leu258 and Leu261 from each ZorA subunit make up continuous hydrophobic pentameric rings. Moreover, we observed an extra density along the tail central axis in this region, which is best modelled as a fatty acid, consistent with a predicted lipid-binding site[22] (Extended Data Fig. 5d,f). Mutants targeting this hydrophobic motif abolished Zorya defence (Fig. 3h) and we observed that, although the tail bundle appeared assembled in purified ZorA(L250G/L254G/L258G/L261G)–ZorB and ZorA(L250N/L254N/L258N/L261N)–ZorB mutants, the TMD domain was perturbed (Extended Data Fig. 5g,h), suggesting that the ZorA tail influences correct ZorAB TMD assembly. To further test this, we deleted the entire tail (ZorA(Δ223–729)) and this mutant also abolished ZorAB TMD motor formation (Extended Data Fig. 5i). Given that the tail structure protrudes into the cytoplasm and is surrounded by aqueous solution, hydrophobic interactions inside the tail seem to be the primary driving force for tail assembly, and it is unlikely that the tail conducts ions or other small soluble molecules (Extended Data Fig. 3d,e). Notably, part of the ZorA tail (residues 540–729) shows homology with the core signalling unit of the bacterial chemosensory array (Extended Data Fig. 3b), which contains a long intracellular helical bundle that is responsible for transferring signal from the extracellular environment into the cell and regulates the activities of the subsequent effectors[23]. Sequence analyses further revealed that long ZorA tails are present in all Zorya subtypes, suggesting that the tail length is functionally important (Extended Data Fig. 6b).

Deletion of any of the Zorya genes results in the loss of anti-phage defence, emphasizing that the complete function of the Zorya system requires communication between the membrane-anchored ZorAB complex and cytosolic soluble proteins (Figs. 3h and 4c,i). Considering the motor-like structural features of the ZorAB TMD and its long cytoplasmic tail, we speculated that the ZorA tail is responsible for transmitting a signal derived from the activated ZorAB motor to the cytosolic proteins ZorC and ZorD. To test whether the length of the ZorA tail is important for Zorya function, we made four ZorA tail truncations: deleting the beginning (ZorA(Δ223–482)), middle (ZorA(Δ359–592)) and tip (ZorA(Δ483–729)) of the tail as well as a combination of deleting the beginning and tip (ZorA(Δ223–343/Δ449–729)) (Extended Data Fig. 5a). Purification and cryo-EM analyses of ZorA(Δ359–592) ZorB and ZorA(Δ435–729) ZorB confirmed that the deletions resulted in shorter tails but did not impair ZorAB TMD motor assembly (Extended Data Fig. 5j–m and Extended Data Table 1). However, all tail truncations abolished Zorya defence (Fig. 3h and Extended Data Fig. 6a,c,d). Mutating the Ca^{2+}-binding site (E86A and E89A) also resulted in the loss of Zorya function (Fig. 3h and Extended Data Fig. 6a) and caused conformational changes in the ZorA MPCD, including to the linker between TM3 and the ZorA tail helix (Extended Data Fig. 5e,n). Thus, inactivation of the Ca^{2+}-binding sites probably disrupts the connection between the ZorAB TMD motor and the tail. These data indicate that ZorAB tail integrity and its motor connection are essential for Zorya anti-phage activity, supporting the role of the ZorA tail in communicating with the cytosolic effector proteins.

ZorC and ZorD DNA interactions enable defence

To better understand the roles of ZorC and ZorD in anti-phage defence, we next obtained their structures and investigated their biological roles. ZorC possesses a domain containing an EH signature motif (Glu400, His443) with unknown function. We determined the cryo-EM structure of *Ec*ZorC to an anisotropic resolution of 3.7 Å, with sufficient density to model residues Arg58–Pro478 (Fig. 4a, Extended Data Fig. 7a–e and Extended Data Table 1). *Ec*ZorC consists of a core domain that connects through a long linker to a C-terminal globular domain, the density of which is blurred, and de novo model building

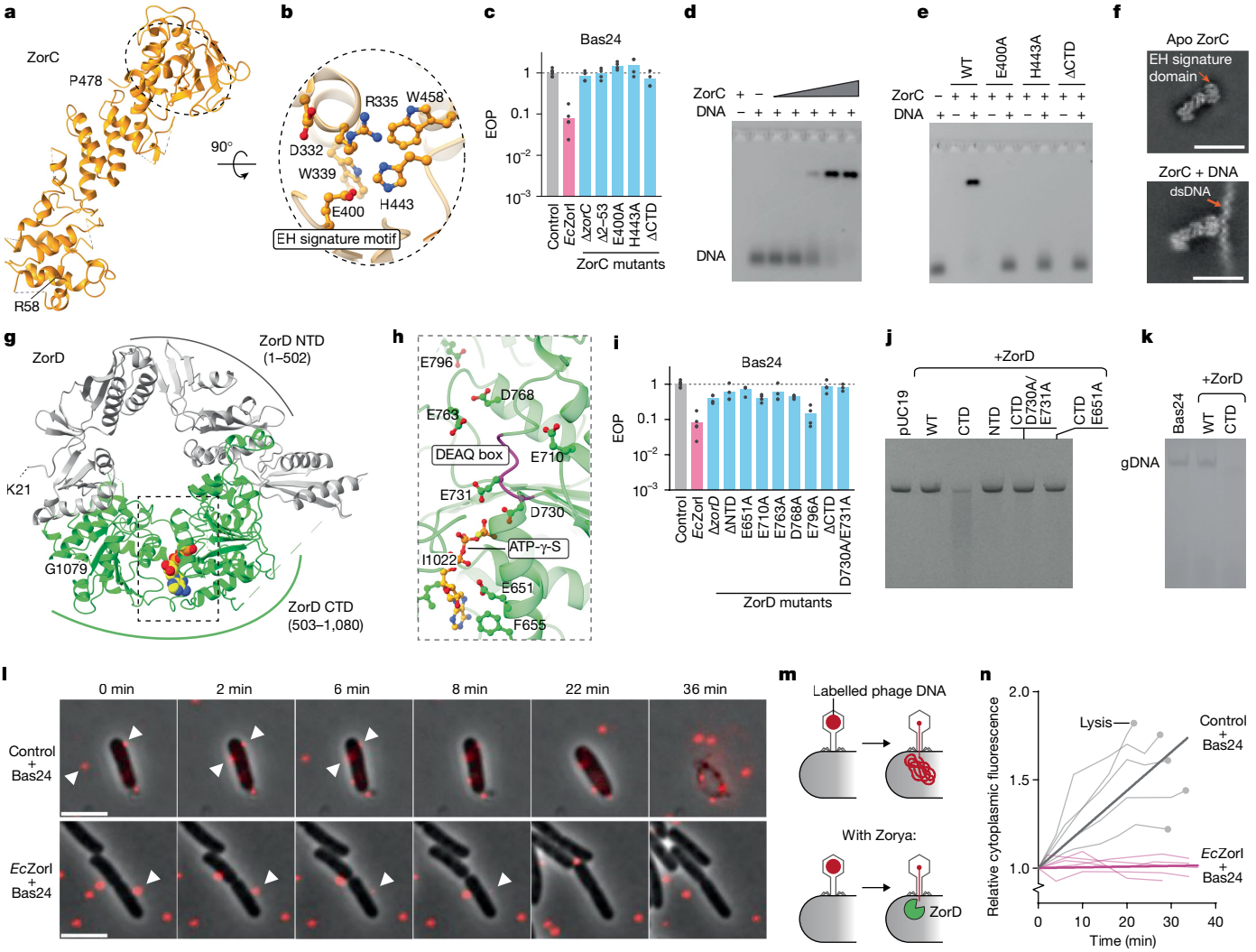


Fig. 4 | Structural and functional characterization of ZorC and ZorD. **a**, Ribbon model representation of ZorC. Residues from Arg58 to Pro478 were modelled based on EM density. **b**, Details of the ZorC EH signature motif. **c**, The effects of ZorC mutations on *Ec*ZorI-mediated anti-phage defence, as measured using EOP assays. **d**, In vitro interaction of *Ec*ZorC with 200 nM dsDNA (18 bp, 50% GC content, 5′ FAM-labelled dsDNA; the sequence is shown in Extended Data Fig. 7h), ZorC concentrations were as follows for lanes 1, 3, 4, 5, 6 and 7, respectively: 2,500, 100, 250, 500, 1,000 and 2,500 nM. **e**, The effects of ZorC mutations on dsDNA-binding activity. All reactions were made to a final concentration of 100 nM of dsDNA and 2,000 nM of protein. **f**, Representatives of high-resolution two-dimensional classes of apo *Ec*ZorC and *Ec*ZorC–dsDNA complex images from cryo-EM. Scale bars, 100 Å. **g**, Ribbon model representation of *Ec*ZorD in a complex with ATP-γ-S, with the bound ATP-γ-S shown as a sphere representation. **h**, Details of ATP-γ-S-binding sites. The backbone of the DEAQ box motif (ZorD residues 730–733) is coloured in magenta. Conserved negatively charged residues surrounding the DEAQ box motif are shown. **i**, The effects of ZorD mutations on *Ec*ZorI-mediated anti-phage defence, as measured using EOP assays. ΔNTD represents ZorD(Δ1–502) and ΔCTD represents ZorD(Δ503–1080). For **c** and **i**, data are the mean of at least three replicates (datapoints indicate replicates) and are normalized to the control samples lacking *Ec*ZorI. Data for additional phages are provided in Extended Data Fig. 6a. **j**, ZorD CTD degrades linear plasmid DNA. **k**, ZorD CTD degrades phage Bas24 genomic DNA (gDNA). Data for additional phages are provided in Extended Data Fig. 8e. **l**, Time-lapse montage of SYTOX-Orange-labelled Bas24 infections. The arrows indicate phage particles that appear to adsorb and inject their DNA. Scale bars, 4 μm. **m**, Schematic of the apparent transfer of labelled phage DNA from the capsid to inside the cell. **n**, Quantification of intracellular fluorescence levels over time in individual *E. coli* cells, comparing the infection dynamics in Zorya-deficient cells and *Ec*ZorI-expressing cells (data from **l**, plus four additional replicates). The dotted points indicate cell lysis of *E. coli* cells lacking *Ec*ZorI. The bold lines represent the mean estimated from a linear regression analysis. Images in **d**, **e**, **j** and **k** are representatives of at least three replicates. CTD, C-terminal domain; NTD, N-terminal domain; WT, wild type.

was not possible (Extended Data Fig. 7b). In the core domain, the EH signature motif Glu400 and His443, together with Asp332, Arg335, Trp339 and Trp458, form an electrostatic network (Fig. 4b). Moreover, in an AlphaFold3-predicted[24] full-length ZorC model, the N-terminal region (residues Met1–Glu48) was modelled with low confidence as two hydrophilic helices that extend from the core domain beyond the density observed in the cryo-EM map (Extended Data Fig. 7b,f). Deletion of the N-terminal helices or the C-terminal globular domain results in loss of Zorya function (Fig. 4c). The ZorC surface contains several patches of net positive charge, including the region containing the EH signature motif (Extended Data Fig. 7g), suggesting that ZorC might interact with nucleic acids. Using electrophoretic mobility shift assays (EMSAs), we observed that ZorC can bind to double-stranded DNA (dsDNA) in a sequence-independent manner (Fig. 4d and Extended Data Fig. 7h). Both the EH signature motif and C-terminal globular domain are required for ZorC DNA-binding activity (Fig. 4e). We further confirmed the ZorC–DNA binding by obtaining a cryo-EM dataset of ZorC in complex with dsDNA. Although we did not obtain a high-resolution reconstruction, the 2D classes clearly indicate that the EH-motif-containing domain interacts with DNA (Fig. 4f and Extended

Data Fig. 7i). Furthermore, ZorC E400A or H443A mutations abolished Zorya function, indicating that ZorC DNA binding is indispensable for Zorya defence (Fig. 4c and Extended Data Fig. 6a).

*Ec*ZorD contains a predicted Snf2-related domain at its C terminus and such domains are known to use ATP hydrolysis to bind to or remodel DNA[25]. We therefore determined the structure of *Ec*ZorD in the absence and presence of a slowly hydrolysable ATP analogue, ATP-γ-S (Fig. 4g, Extended Data Fig. 8a–d and Extended Data Table 1). The *Ec*ZorD N-terminal domain (residues Met1–Asn502) interacts directly with its C-terminal domain (residues Asp503–Ala1080), forming a toroid-shaped molecule. ATP-γ-S is bound within a cleft near the hallmark DEAQ box motif (ZorD residues Asp730–Gln733), surrounded by many conserved negatively charged residues (Fig. 4h). Mutations in both the ATP-binding site (D730A/E731A) and those conserved negatively charged residues resulted in loss of Zorya function (Fig. 4i). We next assessed ZorD DNA-targeting activity by incubating purified ZorD with plasmid DNA in vitro. Full-length ZorD was unable to degrade DNA; however, the ZorD C-terminal domain exhibited nuclease activity and rapidly degraded both plasmid DNA and phage genomic DNA (Fig. 4j,k and Extended Data Fig. 8e). Mutating the DEAQ box motif (D730A, E731A) or a glutamate (Glu651) recognizing the ATP ribose group completely abolished the nuclease activity of the ZorD C-terminal domain (Fig. 4j). Moreover, the presence of ZorC did not activate the autoinhibited ZorD nuclease activity or inhibit the activity of the ZorD C-terminal nuclease domain (Extended Data Fig. 8f,g). This aligns with the finding that ZorC and/or ZorD alone cannot protect against phage infection without ZorAB. Our results suggest that ZorD has nuclease activity and that its full-length form is autoinhibited, probably becoming active once defence is triggered, presumably through a conformational change consistent with AlphaFold3 predictions (Extended Data Fig. 8h–m).

To directly measure ZorD nuclease activity in vivo, we attempted to use the *parS*-ParB system to track phage DNA within infected cells[26]. While the system was not functional with Bas24, we were able to establish activity with Bas54, against which Zorya also confers protection (Fig. 1b). Using time-lapse microscopy to track the presence of phage DNA during Bas54::*parS* infection, we observed Bas54::*parS* DNA within infected cells (detected as fluorescent spots of ParB–mScarlet bound to *parS* loci) in the absence but not presence of Zorya (Extended Data Fig. 7j and Supplementary Video 7). This supports our model for Zorya-mediated phage DNA degradation, but could also be explained if Zorya blocked injection of phage DNA. To exclude this possibility, we labelled Bas24 phage DNA with a fluorescent dye (SYTOX Orange) and used time-lapse imaging to track phage adsorption, DNA injection and subsequent fate of the injected DNA and infected cells. In the absence of Zorya, we observed individual phage particles adsorbing and injecting their DNA, apparent as transfer of fluorescence from the phage capsid to an accumulation of intracellular fluorescence. In the presence of Zorya, DNA injection still occurred (the fluorescence from the adsorbed phage particles decreased over time, consistent with the rate of DNA injection in the absence of Zorya), but we detected no intracellular fluorescence accumulation, suggesting that the injected phage DNA was rapidly degraded and that the cells were protected from lysis (Fig. 4l–n and Supplementary Video 8). These experiments provide further evidence supporting phage DNA degradation by the *Ec*ZorI system.

ZorAB recruits ZorC/D during phage invasion

We next examined how ZorAB, ZorC and ZorD coordinate during phage infection. We examined whether *zorC* and *zorD* from a *Pseudomonas aeruginosa* type I Zorya system could complement corresponding deletions of *Ec*ZorI genes, but this was not the case, suggesting that direct interactions occur between ZorAB and either or both ZorC or ZorD (Extended Data Fig. 9a,b). Moreover, quantitative MS analysis of *Ec*ZorI-expressing cells implied an approximate 1:1:1 $ZorA_5B_2$:ZorC:ZorD stoichiometry (Extended Data Fig. 2g,h). We next used TIRF microscopy to examine the subcellular distributions of functional mNeonGreen (mNG) or HaloTag (HT) fusions to ZorB, ZorC and ZorD (Extended Data Fig. 9c,d). In the presence and absence of phage, ZorB–mNG and ZorB–HT formed distinct membrane-associated foci, while expression of mNG alone from the *Ec*ZorI promoter resulted in uniform, cytoplasmic fluorescence independent of phage (Fig. 5a,c and Extended Data Fig. 9e,f). We observed a slight but significant increase in ZorB–mNG foci in phage-infected bacteria, which was independent of ZorAB motor function, as the non-rotating ZorB(D26N) mutant showed a similar increase (Fig. 5a,b). For both mNG–ZorC and ZorD–mNG, we observed a significant increase in membrane-associated focus formation after phage infection compared with the non-phage control, suggesting that the cytosolic effector proteins ZorC and ZorD are recruited to an activated *Ec*ZorI system (Fig. 5c,d). This result is reinforced by a positive correlation between the number of ZorD foci and the phage MOI (Extended Data Fig. 9h,i). We next used dual-tagged constructs to investigate co-localization of ZorB–HT with either mNG-tagged ZorC or ZorD (Extended Data Fig. 9c). In the absence of phage, ZorB and ZorC or ZorD foci rarely co-localized. However, after Bas24 infection, co-localization of ZorC–mNG or ZorD–mNG with ZorB–HT occurred in approximately 20% or 30% of cells, respectively (Fig. 5e–h). As TIRF microscopy visualizes only about a quarter of the cell depth[27], the observed co-localization is probably an under-representation of the ZorC/D recruitment frequency. No ZorD co-localization was detected in the non-functional motor ZorB(D26N) mutant or in a ZorA tail-tip-deletion mutant (Fig. 5h). Overall, these data suggest that the cytosolic effector proteins ZorC and ZorD are recruited to phage-activated ZorAB complexes, and that both the rotary function of ZorAB and the cytoplasmic tail of ZorA are required to transmit the phage infection signal and recruit the cytosolic nuclease to activated ZorAB.

Discussion

Here we show that an *E. coli* type I Zorya system exhibits defence activity against phylogenetically diverse phages through a direct immunity mechanism but not against bacterial conjugation or plasmid transformation. ZorA and ZorB form an inner-membrane-integrated $ZorA_5B_2$ proton-driven rotary motor complex with a long, intracellular tail structure. We propose that the ZorAB complex acts as a sensor to detect phage infection and transmits the invasion signal through rotation of the ZorA tail to recruit and activate the effectors ZorC and ZorD, which bind to and degrade invading phage DNA (Fig. 5i–k). Our model and data refute a previous hypothesis that ZorC and ZorD are involved in sensing and inactivation of phage DNA, with ZorAB acting as a proton channel to depolarize the membrane potential and induce cell death if the initial ZorC/D protection failed[1]. Instead, we found that the defence is direct, does not induce cell dormancy or death, and that ZorC and ZorD alone (without ZorAB) do not provide protection from phage infection.

We propose that ZorAB senses phage-induced perturbations that reduce the distance between the inner membrane and the PG layer, enabling ZorB to anchor to PG. PG binding and ion flow through the TMD are essential for ZorAB and the flagellar stator unit MotAB to function as rotary motors. However, the short distance of the ZorB PGBD to the membrane motor unit (due to the absence of a flexible linker typical for MotB) means that, normally, the PG layer is too distant for ZorB to reach (Fig. 3a and Extended Data Fig. 4e). We therefore propose that ZorAB complexes are usually inactive and free to diffuse laterally, as observed for MotAB complexes before flagellar incorporation[13,28]. Perturbation of the PG layer or an increase in the local curvature of the inner membrane (which is known to occur during breaching of the cell envelope by some phages)[29], or other phage-induced cell envelope changes that

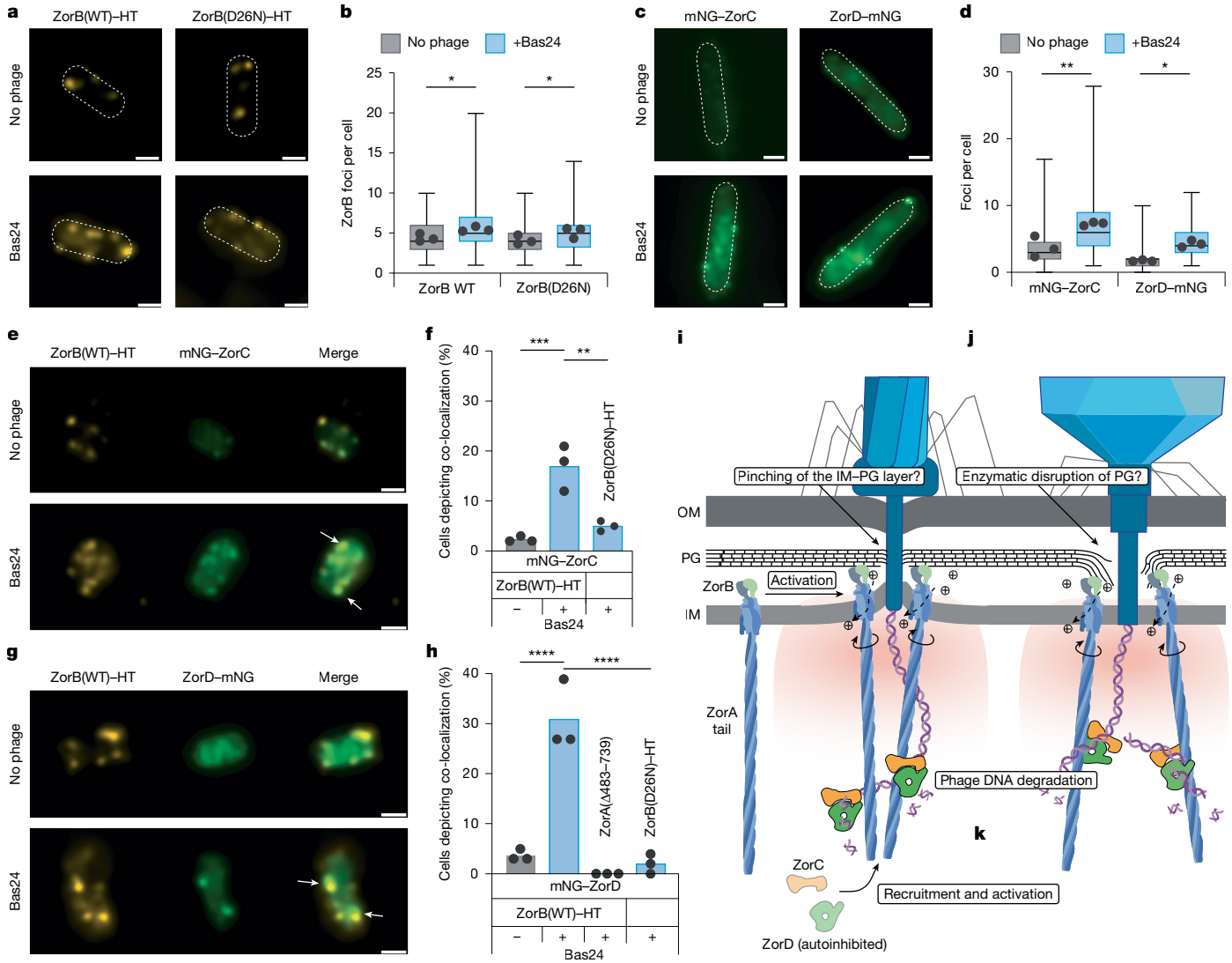


Fig. 5 | Subcellular distributions and co-localization of Zorya components with a proposed model. a, Exemplary denoised TIRF images of the subcellular distributions of WT ZorB and ZorB(D26N) fused to HT with or without Bas24. Scale bars, 1 µm. **b**, Comparison of the detected maxima of the ZorAB complex foci between the untreated or Bas24-exposed (MOI, 5; 30 min) conditions. $n > 250$ cells, 3 replicates. $P = 0.022$ (WT) and $P = 0.027$ (D26N). **c**, Exemplary denoised TIRF images of the subcellular distributions of mNG-tagged ZorC and ZorD, with or without Bas24; mNG was fused to either the ZorC N terminus (mNG–ZorC) or ZorD C terminus (ZorD–mNG). Scale bars, 1 µm. **d**, Comparison of the detected maxima of the ZorC and ZorD foci between the untreated and Bas24-exposed (MOI 5, 30 min) conditions. $n > 250$ cells. $P = 0.004$ and $P = 0.04$. **e**, Exemplary denoised TIRF images of co-localization of mNG–ZorC with ZorB(WT)–HT with or without Bas24. Scale bars, 1 µm. The white arrows highlight co-localization. **f**, Co-localization analysis of ZorB(WT)–HT or ZorB(D26N)–HT with ZorC–mNG with or without Bas24 (MOI, 5; 30 min). $P = 0.0002$ (left), $P = 0.0042$ (right). **g**, Exemplary denoised TIRF images of co-localization of ZorD–mNG with ZorB–HT, with or without Bas24. Scale bars, 1 µm. **h**, Co-localization analysis of ZorB(WT)–HT or mutants (ZorB(D26N)–HT and ZorA(Δ483–739) (tail-tip deletion)) with ZorD–mNG, with or without Bas24 (MOI, 5; 30 min). ****$P < 0.0001$. **i**–**k**, The proposed Zorya defence model. OM, outer membrane. **i**, Inactive ZorAB diffuses laterally within the inner membrane. **j**, Inactive ZorAB detects cell envelope perturbation during phage infection and ZorB PGBDs anchor to PG. Ion translocation through ZorAB triggers ZorA and its tail to rotate around ZorB. **k**, The ZorAB motor signal is transferred through the ZorA tail, which recruits and/or activates ZorC and ZorD. ZorC and ZorD bind to and degrade phage DNA, preventing phage replication. Datapoints represent the mean focus counts for each of three replicates and the shaded bars represent the mean between replicates. For **b** and **d**, data are mean values, with the whiskers representing the minimum to maximum values. Statistical analysis was performed using unpaired *t*-tests (**b**,**d**,**f**) or two-way ANOVA (**h**)[33]. ZorB(WT)–HT and ZorB(D26N)–HT expression data are provided in Extended Data Fig. 6e.

reduce the inner-membrane–PG-layer distance (such as localized PG degradation generating 'frayed' edges)[30], would enable the ZorB PGBD to contact the PG layer, allowing binding and subsequent activation of ZorAB and recruitment and/or activation of effectors (ZorC and ZorD) to clear the phage infection (Fig. 5i–k). Notably, although flagellar stator units are widely distributed among both Gram-positive and Gram-negative bacteria[13], Zorya is under-represented in bacteria with single-membrane cell envelopes, suggesting that differences in cell wall architecture or phage infection mechanisms might prevent effective phage-induced activation of Zorya (Extended Data Fig. 1a–c).

Sensing perturbation of the cell envelope provides an elegant mechanism that exploits the critical early stages of infection to trigger a direct anti-viral defence. While the exact mechanism of phage-induced ZorAB anchoring and activation remains unclear, the need of a rotary motor and the long ZorA tail to recruit and activate effectors suggests an intriguing hypothesis in which the effectors (ZorC/ZorD) are specifically activated in close proximity to the cell membrane. This localized

effector function would protect host DNA from effector activity without relying on epigenetic-based self versus non-self discrimination mechanisms[31] (Supplementary Discussion).

Overall, we provide structural and functional insights into the Zorya defence system and propose that Zorya acts early in infection by sensing perturbation of the cell envelope to initiate a localized anti-phage response near the cell membrane. Our work paves the way for further research to understand the detailed mechanisms of this unique activation signal for anti-phage 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/s41586-024-08493-8.

1. Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
2. Payne, L. J. et al. Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. *Nucleic Acids Res.* **49**, 10868–10878 (2021).
3. Smith, W. P. J., Wucher, B. R., Nadell, C. D. & Foster, K. R. Bacterial defences: mechanisms, evolution and antimicrobial resistance. *Nat. Rev. Microbiol.* **21**, 519–534 (2023).
4. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* **18**, 113–119 (2020).
5. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* **577**, 327–336 (2020).
6. Wang, J. Y. & Doudna, J. A. CRISPR technology: a decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
7. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
8. Millman, A. et al. An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* https://doi.org/10.1016/j.chom.2022.09.017 (2022).
9. Stokar-Avihail, A. et al. Discovery of phage determinants that confer sensitivity to bacterial immune systems. *Cell* **186**, 1863–1876 (2023).
10. 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).
11. Santiveri, M. et al. Structure and function of stator units of the bacterial flagellar motor. *Cell* https://doi.org/10.1016/j.cell.2020.08.016 (2020).
12. Deme, J. C. et al. Structures of the stator complex that drives rotation of the bacterial flagellum. *Nat. Microbiol.* **5**, 1553–1564 (2020).
13. Hu, H. et al. Structural basis of torque generation in the bi-directional bacterial flagellar motor. *Trends Biochem. Sci.* **47**, 160–172 (2021).
14. Rousset, F. & Sorek, R. The evolutionary success of regulated cell death in bacterial immunity. *Curr. Opin. Microbiol.* **74**, 102312 (2023).
15. Maffei, E. et al. Systematic exploration of *Escherichia coli* phage–host interactions with the BASEL phage collection. *PLoS Biol.* **19**, e3001424 (2021).
16. Rieu, M., Krutyholowa, R., Taylor, N. M. I. & Berry, R. M. A new class of biological ion-driven rotary molecular motors with 5:2 symmetry. *Front. Microbiol.* **13**, 948383 (2022).
17. Martin, F. J. O., Santiveri, M., Hu, H. & Taylor, N. M. I. Ion-driven rotary membrane motors: from structure to function. *Curr. Opin. Struct. Biol.* **88**, 102884 (2024).
18. Roujeinikova, A. Crystal structure of the cell wall anchor domain of MotB, a stator component of the bacterial flagellar motor: implications for peptidoglycan recognition. *Proc. Natl Acad. Sci. USA* **105**, 10348–10353 (2008).
19. Zhu, S. et al. Conformational change in the periplamic region of the flagellar stator coupled with the assembly around the rotor. *Proc. Natl Acad. Sci. USA* **111**, 13523–13528 (2014).
20. Trinh, N. T. T. et al. Crystal structure of type IX secretion system PorE C-terminal domain from *Porphyromonas gingivalis* in complex with a peptidoglycan fragment. *Sci. Rep.* **10**, 7384 (2020).
21. Hu, H. et al. Ion selectivity and rotor coupling of the *Vibrio* flagellar sodium-driven stator unit. *Nat. Commun.* **14**, 4411 (2023).
22. Krapp, L. F., Abriata, L. A., Cortés Rodriguez, F. & Dal Peraro, M. PeSTo: parameter-free geometric deep learning for accurate prediction of protein binding interfaces. *Nat. Commun.* **14**, 2175 (2023).
23. Riechmann, C. & Zhang, P. Recent structural advances in bacterial chemotaxis signalling. *Curr. Opin. Struct. Biol.* **79**, 102565 (2023).
24. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
25. Dürr, H., Flaus, A., Owen-Hughes, T. & Hopfner, K.-P. Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* **34**, 4160–4167 (2006).
26. Owen, S. V. et al. Prophages encode phage-defense systems with cognate self-immunity. *Cell Host Microbe* **29**, 1620–1633 (2021).
27. Mattheyses, A. L., Simon, S. M. & Rappoport, J. Z. Imaging with total internal reflection fluorescence microscopy for the cell biologist. *J. Cell Sci.* **123**, 3621 (2010).
28. Wadhwa, N. & Berg, H. C. Bacterial motility: machinery and mechanisms. *Nat. Rev. Microbiol.* https://doi.org/10.1038/s41579-021-00626-4 (2021).
29. Hu, B., Margolin, W., Molineux, I. J. & Liu, J. Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl Acad. Sci. USA* **112**, E4919–E4928 (2015).
30. Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y. & Drulis-Kawa, Z. Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl. Microbiol. Biotechnol.* **101**, 3103–3119 (2017).
31. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
32. Matias, V. R. F., Al-Amoudi, A., Dubochet, J. & Beveridge, T. J. Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. *J. Bacteriol.* **185**, 6112–6118 (2003).
33. GraphPad Prism v.9.0.0 for Mac OS X. www.graphpad.com (GraphPad Software, 2019).

Methods

Phylogenetic analysis of Zorya systems

To create the phylogenetic tree shown in Extended Data Fig. 1b, Zorya systems were identified by running PADLOC[34] v.1.1.0 with PADLOC-DB v.1.4.0 across RefSeq v209 bacterial genomes[35]. For a representative sample of high-quality Zorya systems, the results were filtered for systems with canonical gene order (*zorABCD*, *zorABE* or *zorFABG* for types I, II and III, respectively) that were not at the edge of a contig and where the ZorA protein had minimum PADLOC-DB target and hmm alignment coverages of 80%. The respective ZorA sequences were trimmed to 250 amino acids to roughly isolate the head domain (representing the TMD motor unit). The ZorA head sequences were then clustered using MMseqs2[36] v.15.6f452 with the options --min-seq-id 0.8 --coverage 0.8, and the representative sequences for each resulting cluster were used for subsequent analyses. For the representative ZorA sequences, a diversified ensemble of 100 replicate alignments was built using the Super5 algorithm of muscle (v.5.1)[37] by perturbing each guide tree with 25 different seeds. *E. coli* MotA (WP_000906340.1) was included in the alignments as an outgroup. The alignment with the greatest column confidence was used to build a phylogenetic tree with FastTree v.2.1.11 with the option -lg (ref. 38). Bacteria encoding each protein were assigned taxonomy based on GTDB v214.1[39], grouping phyla with alphabetic suffixes into their base phylum. To create the phylogenetic tree shown in Extended Data Fig. 1c, the GTDB v214.1 bacterial reference tree was filtered for genomes present in RefSeq v209 and collapsed to the phylum level.

Cloning of the Zorya defence system and mutagenesis

The *Ec*ZorI operon with its native promoter region was amplified by PCR from the *E. coli* strain NCTC9026 genome (purchased from the National Collection of Type Cultures (NCTC)) and subcloned into a modified pACYC vector using the In-Fusion cloning strategy (In-Fusion Snap Assembly Master Mix; TaKaRa, 638947). The *Pa*ZorI operon was amplified from the *P. aeruginosa* strain DSM24068 genome (DSMZ-German Collection of Microorganisms and Cell Cultures; Leibniz Institute) and was subcloned into a modified pACYC vector under the *E. coli* ZorI native promotor using the In-Fusion cloning strategy. For generating mutations (point mutations, deletions, mNeonGreen or Halo tag insertions, where *Ec*ZorI *zorC* or *zorCD* genes were replaced by *Pa*ZorI *zorC* or *zorCD*), plasmids were constructed based on standard cloning techniques (In-fusion snap assembly). All plasmids were verified by either Sanger or Nanopore sequencing.

Phage infectivity assays

The host *E. coli* ΔRM (hereafter ΔRM, a derivative of *E. coli* MG1655 engineered to remove multiple restriction modification systems, which was used isolate the BASEL phage collection)[15], possessing either pControl (pACYC) or p*Ec*ZorI (or mutants thereof) were grown overnight in LB + chloramphenicol (25 μg ml^{-1}). EOP assays were performed using bacterial lawns of the host strain in 0.35% LB agar + 10 mM $MgSO_4$ + 2 mM $CaCl_2$ overlaid onto 1.5% LB agar + chloramphenicol. Tenfold dilution series of phages were spotted onto the overlays, air-dried, then the plates were incubated overnight at 30 °C. Liquid culture infection time courses were performed in 96-well plates in an incubated shaking plate reader at 30 °C. The time courses were begun with cells at an optical density at 600 nm (OD_{600}) of 0.05 and phages were added at the indicated MOI, assuming an OD_{600} to cell ratio of 3×10^8 cells per OD_{600} unit. The average amino acid identity between phages (indicative of relatedness) and hierarchical clustering of phage genomes was calculated using EzAAI[40], separating clusters with proteome coverage <40%.

Phage adsorption and one-step growth curve assays

Overnight cultures of ΔRM possessing either pControl or p*Ec*ZorI were used to inoculate fresh LB + chloramphenicol cultures at a 1:100 dilution. The inoculated cultures were grown at 30 °C with shaking until reaching an OD_{600} of 0.4–0.6, then collected by centrifugation, washed with LB + chloramphenicol and resuspended at an OD_{600} of 1.0 in LB + 10 mM $MgSO_4$ + 2 mM $CaCl_2$. For phage adsorption assays, 10 ml samples of resuspended cells were infected with phage Bas24 at an MOI of 10^{-4}, then the samples were mixed and incubated at 30 °C without shaking. For the 0 min timepoint (total input phages), 100 μl samples were removed and added to 0.35% LB agar seeded with ΔRM + pControl (as an indicator lawn), then poured on top of 1.5% LB agar + chloramphenicol. For each subsequent timepoint, 1 ml samples were taken, centrifuged to pellet cells and the supernatant (containing unabsorbed phages) was then filtered through a 0.2 μm PES syringe filter. Samples (100 μl) of the filtered supernatant were added to indicator overlays (as above) poured onto 1.5% LB agar + chloramphenicol. All overlay plates were incubated overnight at 30 °C before counting plaques. For each timepoint, the percentage of unabsorbed phages was calculated as the timepoint plaque count/plaque count for the time 0 min pControl sample. For the one-step phage growth curves (burst time and size), 2 ml samples of the cells resuspended at and OD_{600} of 1.0 in LB + 10 mM $MgSO_4$ + 2 mM $CaCl_2$ (as above) were infected with phage Bas24 at an MOI of 10^{-4}, then the samples were mixed and two tenfold diluted samples were prepared, and the dilution series for each sample was then incubated at 30 °C without shaking. At the indicated timepoints, 100 μl samples of each dilution were removed and added to 0.35% LB agar seeded with ΔRM + pControl (as an indicator lawn), then poured on top of 1.5% LB agar + chloramphenicol. All overlay plates were incubated overnight at 30 °C before counting plaques. For each timepoint, the PFU was normalized to the PFU of the 0 min pControl samples.

Conjugation assays

Plasmids encoding kanamycin resistance and each possessing different origins of replications (ColE1, pMAT16; RSF1010, pPF1825[41]; pBBR1, pSEVA237R[42]; or RK2, pPF1619) were conjugated from the *E. coli* donor ST18[43,44] (an auxotroph requiring supplementation with 5-aminolevulinic acid; ALA) into the *E. coli* recipient ΔRM possessing either pControl or p*Ec*ZorI. Matings were performed at the indicated donor to recipient ratios (D:R) and incubated overnight on LB agar + chloramphenicol + ALA at 30 °C. The conjugation efficiency was determined by plating dilution series of the matings onto LB agar + chloramphenicol + kanamycin (transconjugants) and LB agar + chloramphenicol (total recipients). The transconjugant frequency was defined as the transconjugant CFU/recipient CFU.

Transformation assays

Chemically competent cells of ΔRM possessing either pControl or p*Ec*ZorI were prepared according to the Inoue method[45], with HEPES-KOH pH 6.8 used for the transformation buffer. Cells were stored in 200 μl aliquots at −80 °C before use. For each transformation assay, 5 ng of plasmid (quantified using a Qubit BR kit) was used. Plasmids used were as described above for the conjugation assays (ColE1, pMAT16; or pBBR1, pSEVA237R).

Cell survival assays

Overnight cultures of ΔRM possessing either pControl or p*Ec*ZorI were used to inoculate fresh LB + chloramphenicol cultures at a 1:100 dilution. The inoculated cultures were grown at 30 °C with shaking until reaching an OD_{600} of 0.4–0.6, then collected by centrifugation, washed with LB + chloramphenicol, and resuspended at an OD_{600} of 0.2. Phage Bas24 was then added at an MOI of 5 to each sample; control samples without phage addition were also included. After 20 min adsorption, tenfold serial dilutions of each sample were plated (100 μl each) onto LB + chloramphenicol, then incubated overnight at 30 °C. The cell survival rate was calculated as [CFU obtained + Bas24]/[CFU obtained without phage addition].

Phages and phage genome purification

Phage primary stocks were prepared using the double-agar method, with near-confluent plaque overlays. The phages were collected by adding SM buffer (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-HCl pH 7.5, 5 mM $CaCl_2$) on top of the overlay agar and mixed for 4 h at 4 °C. The suspension was collected and centrifuged for 15 min at 4,000*g*. High-titre phage samples were obtained by inoculating 1–3 l of LB with a 10^3 dilution of an overnight culture of ΔRM and grown at 37 °C to an OD_{600} of 0.3. The bacterial culture was inoculated with the primary stock to an MOI of 0.025 and infection was carried out at 37 °C at 90 rpm until a clear lysate was obtained. The lysate was collected at 4,000*g*, for 15 min at 4 °C. After decanting the supernatant, 1 μg ml^{-1} of DNase I and 1 μg ml^{-1} of boiled RNase A were added to the cleared lysate. The lysate was gently stirred at 90 rpm for 30 min at room temperature.

Phages were concentrated by polyethylene glycol (PEG) precipitation. NaCl was gradually added to a final concentration of 1 M, followed by gradual addition of 10% PEG 8,000 with continuous stirring at room temperature until dissolved. After obtaining a clear solution, the lysate was stirred (100 rpm, 30 min; 4 °C) and left overnight at 4 °C. The lysate was centrifuged (15,000*g*, 1 h, 4 °C) and the clear supernatant was removed. The precipitate was resuspended in the minimal amount (up to 2 ml) of SM buffer that allowed solubilization. Insoluble materials were removed by adding 20% (v/v) of chloroform and centrifuged (8,000*g*, 10 min). The supernatant was stored at 4 °C to be used as phage sample for the next step. The phage was then purified by rate zonal separation using OptiPrep Density Gradient Medium (Sigma-Aldrich) in a density gradient ranging from 50 to 10%, diluted in SM medium. The phage sample was applied on the top of the gradient and centrifuged (150,000*g*, 18 h, 4 °C). The phage was extracted, dialysed against SM buffer and the samples were stored at 4 °C. The phage genomes were extracted using the Phage DNA isolation kit from Norgen Biotek, aliquoted and stored at −20 °C.

Phage genome labelling

Stocks of Bas24 were treated with Pierce Universal Nuclease according to the manufacturer's protocol for 1 h at 37 °C. To stain the phage genomic DNA, SYTOX Orange (Invitrogen) stock solution was added to 10 ml of the phage lysate at a concentration of 1:2,000 and incubated overnight at 4 °C in the dark. Stained phage particles were subsequently purified by PEG precipitation. PEG 6,000 was added to the lysate to a final concentration of 10% (w/v) and incubated overnight at 4 °C to allow for phage aggregation and precipitation. The lysate was centrifuged at 4,000*g* for 30 min at 4 °C to pellet the phages and the supernatant was carefully discarded without disturbing the phage pellet. The phage pellet was then washed by gently adding a 1 ml SM buffer, centrifuged at 6,000*g* for 2 min and used for subsequent time-lapse microscopy experiments.

Protein expression and purification

ZorAB. The full-length genes of *E. coli* ZorA and ZorB code for 729 and 246 residues, respectively. The tandem gene was PCR amplified from the *E. coli* strain NCTC9026 genome and subcloned into a modified pET vector containing a C-terminal human rhinovirus (HRV) 3C protease cleavage site and a twin-Strep-tag II (resulting in pET11a-ZorA-ZorB-3C-TSII). The plasmids containing the recombinant genes were transfected into *E. coli* C43(DE3) competent cells and the proteins were expressed in LB medium. When the culture OD_{600} reached 0.6–0.8, the temperature was decreased from 37 °C to 24 °C, then grown until the OD_{600} reached approximately 0.8–1.0, before 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added for overnight protein induction. The culture was collected, and the cell pellet was resuspended in buffer A containing 20 mM HEPES-NaOH pH 7.5, 300 mM NaCl supplemented with EDTA-free protease inhibitor (Thermo Fisher Scientific) and lysozyme from chicken white egg (Sigma-Aldrich) to a final concentration of 50 μg ml^{-1} and deoxyribonuclease I from bovine (Sigma-Aldrich) to a final concentration of 30 μg ml^{-1}. The mixture was disrupted by high-pressure homogenizer and centrifuged at 185,000*g* for 1 h. The pellet containing the membrane was collected and solubilized using buffer B containing 30 mM HEPES-NaOH pH 7.5, 300 mM NaCl, 10% glycerol, 2% LMNG (Anatrace), supplemented with EDTA-free protease inhibitor at 4 °C for 2 h. The solubilized membrane was then centrifuged at 90,000*g* for 40 min and the supernatant was loaded onto a gravity flow column containing 2 ml (resin volume) of Strep-Tactin Superflow high-capacity resin (IBA), pre-equilibrated with wash buffer containing 20 mM HEPES-NaOH pH 7.5, 300 mM NaCl, 10% glycerol and 0.005% LMNG. The resins were washed five times with 2–3 resin volumes of the wash buffer and elution was carried out five times with 0.5 resin volume (1 ml) of elution buffer containing 20 mM HEPES-NaOH pH 7.5, 300 mM NaCl, 10% glycerol, 0.005% LMNG and 10 mM desthiobiotin). The recombinant protein was then concentrated and loaded onto a pre-equilibrated (20 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 0.002% LMNG) Superose 6 Increase 10/300 GL size-exclusion chromatography column. The fractions from the elution peak corresponding to the molecular mass of the ZorAB complex were pooled, and the protein was concentrated for cryo-EM grid preparation and functional experiments. The procedures for expression and purification of ZorAB mutants were similar to those for the ZorAB wild type.

ZorB and MotB soluble PGBDs. The PGBDs of each ZorB and ZorB (Y151A/N152A/L155A/R159A) (from Arg52 to C terminus) and *E. coli* MotB (from Glu96 to C terminus) were cloned from the existing full-length expression constructs[11] into pET11a by PCR amplification-based deletion mutagenesis (TakaraBio). The ZorB and MotB PGBDs were purified similarly to ZorC and ZorD with a few modifications. The ZorB and the ZorB(Y151A/N152A/L155A/R159A) PGBD vectors were transformed into Rosetta-gami-2(DE3) competent *E. coli* (Novagen). Cells were grown in LB medium supplemented with 100 μg ml^{-1} ampicillin, 34 μg ml^{-1} chloramphenicol and 10 μg ml^{-1} tetracycline at 37 °C to an OD_{600} of 0.7. The cells were then induced with 0.5 mM IPTG and allowed to grow for 16 h at 18 °C. The cultures were collected and the cell pellets resuspended in lysis/wash buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl supplemented with EDTA-free protease inhibitor cocktail (Roche)). Moreover, 1 mg of DNase I and 0.5 mM $MgCl_2$ was added to the resuspended cells. The cells were lysed using the Avestin Emulsiflex C3 homogenizer, cooled to 4 °C and soluble lysates were cleared by centrifugation at 30,000*g* at 4 °C for 30 min. The supernatant was then run over a gravity-flow column containing 2 ml (resin volume) of Strep-Tactin 4Flow high-capacity resin (IBA), pre-equilibrated with wash buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl). The resin was washed with 20 ml lysis/wash buffer and protein was eluted in 12 ml elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM desthiobiotin). The elution was then concentrated and run over a Superose 6 Increase 10/300 GL size-exclusion chromatography column into gel-filtration buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl). The MotB PGBD was expressed and purified identically to the ZorB PGBDs, with the exception that it was expressed in BL21(DE3) gold *E. coli* in LB medium supplemented with 100 μg ml^{-1} ampicillin.

ZorC. The predicted *zorC* gene encodes 560 residues. The *zorC* gene together with a short region upstream of the *zorC* N terminus that encodes 7 residues (LPVGYAT) was PCR amplified from the DNA genome of *E. coli* strain NCTC9026 and subcloned into the modified pET vector (resulting in pET11a-ZorC-3C-TSII). *E. coli* BL21 (DE3) gold chemically competent cells were transformed with the plasmids and the protein was expressed in LB medium with the presence of 100 μg ml^{-1} of ampicillin. In brief, when the OD_{600} reached 1.0–1.2, the temperature was decreased to 16 °C and 0.5 mM IPTG was added for overnight protein induction. The culture was collected, and the cell pellet was resuspended using buffer containing 20 mM Tris-HCl pH 7.5, 10% glycerol and 500 mM NaCl

supplemented with EDTA-free protease inhibitor (Thermo Fisher Scientific). The cells were lysed using an Avestin Emulsiflex C3 homogenizer, cooled to 4 °C and centrifuged at 18,000*g* for 40 min. The supernatant was then added to a gravity-flow column containing 3 ml (resin volume) of Strep-Tactin Superflow high-capacity resins (IBA), pre-equilibrated with wash buffer (20 mM Tris-HCl pH 7.5, 10% glycerol and 500 mM NaCl). Resins were washed five times with 2–3 resin volumes of wash buffer and elution was performed with 4 CV of elution buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl and 10 mM desthiobiotin). The recombinant protein was pooled and concentrated and was loaded onto a pre-equilibrated (20 mM Tris-HCl pH 7.5, 500 mM NaCl) Superose 6 Increase 10/300 GL size-exclusion chromatography column. Peak fractions were pooled, and another round of size-exclusion chromatography was carried out with buffer 20 mM HEPES-NaOH pH 7.5, and 150 mM NaCl to decrease the NaCl concentration. The fractions from the elution peak corresponding to the molecular mass of ZorC were pooled and the protein was concentrated to approximately 1 mg ml^{-1} for cryo-EM grid preparation and functional experiments. ZorC proteins used for EMSAs were exchanged into buffer containing 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol and 1 mM TCEP after elution. Pure fractions were concentrated and flash-frozen in small aliquots and stored at −80 °C until use. The sample purity was assessed using SDS–PAGE. The procedures of expression and purification of ZorC mutants were similar to those for the ZorC wild type.

ZorD. The predicted *zorD* gene encoding 1,086 residues was PCR amplified from the DNA genome of *E. coli* strain NCTC9026 and was subcloned into the modified pET vector, resulting in pET11a-ZorC-3C-TSII. The expression and purification of ZorD protein were similar to those for the ZorC protein, except for the purification buffer. In brief, the suspension buffer contained 150 mM NaCl, 20 mM HEPES-NaOH pH 7.5 and 10% glycerol; the wash buffer was the same as the suspension buffer and the elution buffer contained 150 mM NaCl, 20 mM HEPES-NaOH pH 7.5, 10% glycerol and 10 mM desthiobiotin; and the size-exclusion chromatography buffer contained 150 mM NaCl and 20 mM HEPES-NaOH pH 7.5. Purified ZorD was concentrated to 0.4–0.6 mg ml^{-1} for functional experiments and cryo-EM grid preparation. For the nuclease experiment, ZorD protein was kept in the elution buffer and flash-frozen in small aliquots and stored at −80 °C until use.

Cryo-EM grid preparation, data collection, model building and refinement

ZorAB. Freshly purified ZorAB sample was concentrated to 2–3 mg ml^{-1} and 2.7 µl protein was applied onto glow-discharged (30 s, 5 mA) grids (Quantifoil R0.6/1300 mesh Au) and plunge-frozen into liquid ethane using the Vitrobot Mark IV (FEI, Thermo Fisher Scientific) with the following settings: 100% humidity, 4°C, blotting force 25, 4–6 s blot time and 7 s wait time. Videos were collected using the semi-automated acquisition program EPU (FEI, Thermo Fisher Scientific) on the Titan Krios G2 microscope operated at 300 keV paired with a Falcon 3EC direct electron detector (FEI, Thermo Fisher Scientific). Images were recorded in electron-counting mode, at ×96,000 magnification with a calibrated pixel size of 0.832 Å and an underfocus range of 0.7 to 2.5 µm. The number of micrographs and the total exposure values for the different datasets are summarized in Supplementary Table 1. Grid preparation and data collection strategies for the ZorAB mutants were similar to those for the ZorAB wild type.

ZorC. Purified ZorC (3 µl at around 1 mg ml^{-1}) was applied onto glow-discharged (30 s, 5 mA) grids (UltrAuFoil R 0.6/1, 300 mesh, Gold) and plunge-frozen into liquid ethane using the Vitrobot Mark IV (FEI, Thermo Fisher Scientific), with the following settings: 100% humidity, 4 °C, blotting force 20, 4 s blot time and 10 s wait time. Videos were collected using the semi-automated acquisition program EPU (FEI, Thermo Fisher Scientific) on the Titan Krios G2 microscope operated at 300 keV paired with a Falcon 3EC direct electron detector (FEI, Thermo Fisher Scientific). Images were recorded in electron-counting mode, at ×96,000 magnification with a calibrated pixel size of 0.832 Å and an underfocus range of 1 to 2.5 µm. The number of micrographs and total exposure values for the datasets are summarized in Extended Data Table 1. For the ZorC + DNA sample, ZorC (final concentration 0.6 mg ml^{-1}) was mixed with commercial pUC19 plasmid (NEB) (final concentration of 0.5 µg µl^{-1}). The samples were incubated at room temperature for 30 min, followed by 30 min at 4 °C before grid preparation. The samples (3 µl) were applied to UltraAuFoil R 2/2, 200 mesh Gold grids (glow discharged 60 s at 10 mA) and plunge-frozen as described above, but with the following settings: blotting force 15, 3 s blot time.

ZorD. ZorD showed preferred orientation of particles on ice. Zwitterionic detergent (0.5% CHAPSO; Anatrace) was added to the purified sample to a final concentration of 0.0125% before cryo-EM grid preparation. For the apo form, the preparation of grids was similar to ZorC. For ZorD in complex with ATP-γ-S, 4 µl of 0.1 mM ATP-γ-S was added to 400 µl of purified ZorD at 0.0375 mg ml^{-1}. The mixture was concentrated to 15 µl to reach a ZorD concentration of around 0.6 mg ml^{-1}. The grid preparation was similar to for ZorC. The number of micrographs and total exposure values for the different datasets are summarized in Extended Data Table 1.

Cryo-EM data processing

All datasets were processed using cryoSPARC[46] v.4.2.1, unless otherwise stated. We started by using Patch motion correction to estimate and correct for full-frame motion and sample deformation (local motion). Patch contrast function (CTF) estimation was used to fit local CTF to micrographs. Micrographs were manually curated to remove low-quality data (relative ice thickness value greater than 1.1 and CTF value worse than 3.5 Å). We performed particle picking by template picking or using topaz particle picking[47]. Particles were extracted with a box size of 500 pixels for ZorAB datasets, 256 pixels for the ZorC dataset and 400 pixels for the ZorD dataset. One round of 2D classification was performed followed by ab initio reconstruction. Heterogeneous refinement was used to exclude broken particles. Non-uniform refinement was applied with a dynamic mask to obtain a high-resolution map. Local refinement was additionally performed with a soft mask to achieve a higher-resolution map of some flexible regions. For all datasets, the number of videos, the number of particles used for the final refinement, map resolution and other values during data processing are summarized in the Extended Data Table 1.

Model building and validation

We used AlphaFold2 or AlphaFold3[24] to predict all of the initial models. The predicted models were manually fit into the cryo-EM density by using UCSF ChimeraX[48]. The model was refined in Coot[49], or using StarMap[50] in the case of ZorC, for which the map is anisotropic and the resolution is modest. The model was then refined against the map using PHENIX real space refinement[51]. The ZorAB composite model was constructed by extending the pentameric tail as an idealized right-handed super-helical coiled-coil. Local conformations were manually adjusted in PyMol[52] (v.2.5) and optimized through energy minimization using GROMACS (v.2022.5)[53] with CHARMM27 force field. However, it is worth noting an irregularity in the AF2 model, specifically in residues 312 to 322, which introduces a substantial twist in the ZorA tail, raising possibilities of other pentameric forms of the ZorA tail and further reflecting its conformational dynamics.

PG purification and pull-down experiments

PG was purified from *E. coli* ΔRM with the protocol adapted from a previous study[54]. In brief, ΔRM cells were incubated in 1 l LB medium until the OD_{600} reached 0.8. Cells were collected and resuspended in 12 ml PBS buffer and split into two 50 ml Falcon tubes, then 10% (w/v)

SDS solution (in PBS) was added to a final concentration of 6% (w/v). The Falcon tubes were boiled for 1 h with stirring at 500 rpm. The heat was turned off and the tube was allowed to cool to ambient temperature overnight. The next day, the solutions from both Falcon tubes were pooled into one 50 ml Falcon tube, and centrifuged at room temperature for 45 min at 108,000*g*. The pellet was washed five times with 5 ml Milli-Q water. The PG was resuspended in 20 ml of buffer containing 50 mM Tris-HCl pH 7.0, and α-amylase was added (Sigma-Aldrich) to a final concentration of 100 μg ml^{-1} and incubated for 2 h at 37 °C. Next, 50 μg ml^{-1} RNase A (Roche) and 10 μg ml^{-1} DNase (Sigma-Aldrich) were added and incubated for an additional 2 h at 37 °C. The mixture was then supplemented with 20 mM $MgSO_4$, 10 mM $CaCl_2$ and 100 μg ml^{-1} trypsin (Sigma-Aldrich), and incubated at 37 °C overnight. The next day, EDTA at pH 8 was added to a final concentration of 10 mM and 10% (w/v) SDS solution to a final concentration of 1% (w/v). The mixture was boiled for 20 min in a water bath and allowed to cool to ambient temperature. The tube was centrifuged at 108,000*g* for 45 min. The resulting pellet was washed five times with Milli-Q water to remove residual SDS. Finally, the pellet was resuspended in 300 μl of Milli-Q water, aliquoted into 35 μl portions and stored at −20 °C. For PG pull-down assays, the purified PG was washed with 1 ml PBS + 0.002% LMNG buffer and centrifuged at 20,000*g* for 30 min. Purified ZorAB and ZorAB ZorB(Y151A/N152A/L155A/R159A) mutant (10 μl, at a concentration of 2 mg ml^{-1}; ZorAB ZorB(Y151A/N152A/L155A/R159A) mutant is less stable, requiring the use of freshly purified protein) was incubated with the PG at room temperature for 1 h, then centrifuged at 20,000*g* for 30 min. The pellet was washed three times with 700 μl of the pull-down buffer by mixing and centrifugation (10 min, 12 °C, 20,000*g*). The supernatant was retained for SDS gel analysis. The pellet was resuspended with 20 μl of buffer and 5 μl of loading dye was added for SDS gel analysis.

For the pull-downs of ZorB PGBD, mutant ZorB PGBD (ZorB(Y151A/N152A/L155A/R159A) PGBD), MotB PGBD (positive control) and ZorE (negative control), the protocol was similar to as above. In brief, PG (PGN-ECndi ultrapure peptidoglycan (InvivoGen); due to the low yield of the laboratory-purified PG) was washed and resuspended in resuspension buffer (20 mM potassium phosphate, pH 7, 150 mM NaCl). Each pull-down reaction contained 10 μl of washed 25 mg ml^{-1} PG, 4 μl of the indicated protein (each added from a 5 mg ml^{-1} stock) and the pull-down buffer (20 mM potassium phosphate pH 6, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$) to final volume of 100 μl. The samples were incubated for 30 min at 20 °C with gentle mixing. The insoluble PG was pelleted by centrifugation at 20,000*g* at 12 °C for 10 min, and the soluble supernatant was retained for SDS–PAGE analysis. The pellet was washed three times with 700 μl of the pull-down buffer by mixing and centrifugation (10 min, 12 °C, 20,000*g*). Finally, the pellet was resuspended in 100 μl of pull-down buffer and 15 μl of each sample (soluble supernatant and resuspended pellet) was mixed with 3 μl loading dye for SDS–PAGE analysis.

DNA-binding experiments

Frozen aliquots of ZorC and ZorC mutants in ZorC buffer (20 mM Tris-HCL pH 7.5, 300 mM NaCl, 10% glycerol, 1 mM TCEP) were thawed and centrifuged to remove possible aggregates. ZorC was diluted to 10× stocks for each concentration used, in the same ZorC buffer. The final EMSA reaction buffer contained phosphate-buffered saline (15.2 mM $NaHPO_4$, 0.90 mM $CaCl_2$, 2.7 mM KCl, 1.47 mM KH_2PO_4, 8.1 mM Na_2HPO_4, 0.49 mM $MgCl_2$, 137.9 mM NaCl) at pH 7.4 and 10% glycerol. ZorC or ZorC buffer was then added, followed by DNA substrate (10× stock in H_2O) or H_2O. DNA oligos with 5′-FAM modifications were synthesized by TAG Copenhagen and double-stranded constructs were obtained by annealing the unlabelled complement sequence (the dsDNA sequences used in this study are shown in Extended Data Fig. 7h). All components were incubated at 4 °C for 30 min and loaded onto a 1.5% (w/v) agarose gel made with 20 mM sodium phosphate buffer (pH 7.2). The samples were run for 30 min at 100 V and 4 °C using 20 mM sodium phosphate (pH 7.2) as the running buffer. The gels were visualized using the Odyssey XF Imaging System at 600 nm.

Nuclease assays

ZorD was incubated with 200 ng pUC19 (linearized by KpnI) in the reaction buffer containing 1× Cutsmart buffer (NEB) and 2 mM ATP (NEB) in a total volume of 25 μl. The reactions were incubated at 37 °C for 1 h with shaking at 600 rpm using the Eppendorf ThermoMixer. DNA product was purified using the NucleoSpin Gel and PCR Clean-up kit (Machery Nagel) using the standard protocol and was analysed with 1% E-Gel EX. For the reaction with the phage genomes, 200 nM proteins were incubated with around 100 ng purified phage genomic DNA in the same reaction buffer indicated above. The reactions were terminated by adding 1× E-gel loading buffer and product was analysed with 1% E-Gel EX.

MS sample preparation

Overnight cultures of *E. coli* ΔRM transformed with p*Ec*ZorI (or pControl), were used to inoculate (at a 1:1,000 dilution) 3 ml LB medium with antibiotics, then grown to an OD_{600} of approximately 0.4. The cell pellet was collected, resuspended in 500 μl of 0.2 M Tris-HCl pH 8.0, and incubated for 20 min. Next, 250 μl of buffer (0.2 M Tris-HCl pH 8.0, 1 M sucrose and 1 mM EDTA) was added, along with 3 μl of 10 mg ml^{-1} lysozyme. The mixture was incubated for 30 min, and 250 μl of 6% (w/v) SDS was added to a final concentration of 1%, after which the sample was heated at 99 °C for 10 min. The mixture was sonicated (Misonix Ultrasonic Liquid Processor with microtip Probe) to fragment DNA and RNA with the following settings: amplitude 10, time 5 s sonication and 5 s pause, 5 cycles.

For MS analysis, we performed protein aggregate capture digestion of proteins[55]. To this end, 250 μl of bacterial lysate was taken from the total sample, and 750 μl of acetonitrile was added into the mixture, along with 50 μl magnetic microspheres that had been prewashed with PBS buffer. The mixture was allowed to settle for 10 min, before retention of the magnetic microspheres by a magnetic plate. Beads were washed once with 1 ml acetonitrile and once with 1 ml of 70% ethanol, after which all ethanol was removed and the beads were stored at −20 °C until further processing. The frozen beads were thawed on ice, supplemented with 100 μl ice-cold 50 mM Tris-HCl pH 8.5 buffer supplemented with 2.5 ng $μl^{-1}$ Lys-C and gently mixed (on ice) every 5 min for 30 min. Digestion was performed for 3 h using the Eppendorf ThermoMixer with shaking at 1,250 rpm and 37 °C. Next, the beads were chilled on ice and 250 ng of sequencing-grade trypsin was added, after which samples were gently mixed (on ice) every 5 min for 30 min. Final digestion was performed overnight using the Eppendorf ThermoMixer with shaking at 1,250 rpm and 37 °C. Peptides were separated from magnetic microspheres using 0.45 μm filter spin columns, and peptides were reduced and alkylated by adding TCEP and chloroacetamide to 5 mM for 30 min before peptide clean-up using the low-pH C18 StageTip procedure. C18 StageTips were prepared in-house by layering four plugs of C18 material (Sigma-Aldrich, Empore SPE Disks, C18, 47 mm) per StageTip. Activation of StageTips was performed with 100 μl 100% methanol, followed by equilibration using 100 μl 80% acetonitrile in 0.1% formic acid, and two washes with 100 μl 0.1% formic acid. Samples were acidified to pH < 3 by addition of trifluoroacetic acid to a concentration of 1%, after which they were loaded on StageTips. Subsequently, StageTips were washed twice using 100 μl 0.1% formic acid, after which peptides were eluted using 80 μl 30% acetonitrile in 0.1% formic acid. All fractions were dried to completion using a SpeedVac at 60 °C. Dried peptides were dissolved in 25 μl 0.1% formic acid and stored at −20 °C until MS analysis.

Approximately 1 μg of peptide was analysed per injection. All samples were analysed on the EASY-nLC 1200 system (Thermo Fisher Scientific) coupled to the Orbitrap Astral mass spectrometer (Thermo Fisher Scientific). The samples were analysed on 20-cm-long analytical

columns, with an internal diameter of 75 µm, and packed in-house using ReproSil-Pur 120 C18-AQ 1.9 µm beads (Dr Maisch). The analytical column was heated to 40 °C, and elution of peptides from the column was achieved by application of gradients with stationary phase buffer A (0.1% formic acid) and increasing amounts of mobile phase buffer B (80% acetonitrile in 0.1% formic acid). The primary analytical gradient ranged from 10% B to 38% B over 57.5 min, followed by a further increase to 48% B over 5 min to elute any remaining peptides, and by a washing block of 15 min. Ionization was achieved using a NanoSpray Flex NG ion source (Thermo Fisher Scientific), with a spray voltage of 2 kV, ion transfer tube temperature of 275 °C and RF funnel level of 50%. All full precursor (MS1) scans were acquired using the Orbitrap mass analyser, while all tandem fragment (MS2) scans were acquired in parallel using the Astral mass analyser. Full scan range was set to 300–1,300 *m/z*, MS1 resolution to 120,000, MS1 AGC target to 250 (2,500,000 charges) and MS1 maximum injection time to 150. Precursors were analysed in data-dependent acquisition mode, with charges 2–6 selected for fragmentation using an isolation width of 1.3 *m/z* and fragmented using higher-energy collision disassociation with normalized collision energy of 25. Monoisotopic precursor selection was enabled in peptide mode. Repeated sequencing of precursors was minimized by setting expected peak width to 20 s, and dynamic exclusion duration to 20 s, with an exclusion mass tolerance of 10 ppm and exclusion of isotopes. MS2 scans were acquired using the Astral mass analyser. MS2 fragment scan range was set to 100–1,500 *m/z*, MS2 AGC target to 50 (5,000 charges), MS2 intensity threshold to 50,000 charges per second and MS2 maximum injection time to 5 ms, therefore requiring a minimum of 250 charges for attempted isolation and identification of each precursor. Duty cycle was fixed at 0.3 s, acquiring full MS scans at ~3.3 Hz and with auto-fitting of Astral scans resulting in MS2 acquisition at a rate of ~100–200 Hz.

MS data analysis

All RAW files were analysed using MaxQuant software (v.2.4.3.0)[56], the earliest release version to support Astral RAW files. The default MaxQuant settings were used, with exceptions outlined below. For generation of the in silico spectral library, the four full-length Zorya protein sequences were entered into a FASTA database, along with all (23,259) Swiss-Prot-reviewed *E. coli* sequences (taxonomy identifier 562) downloaded from UniProt[57] on 7 September 2023. The data were first searched using pFind (v.3.2.0)[58], using the Open Search feature to determine the overall peptide properties and commonly occurring (affecting >1% of PSMs) peptide modification in an unbiased manner. For searching Astral.RAW files using pFind, .RAW files were first converted to .mzML using OpenMS (v.3.0.0)[59]. For the main data search using MaxQuant, digestion was performed using 'Trypsin/P' with up to 2 missed cleavages (default), with a minimum peptide length of 6 and a maximum peptide mass of 5,000 Da. No variable modifications were considered for the first MS/MS search, which is only used for precursor mass recalibration. For the MS/MS main search a maximum allowance of three variable modifications per peptide was set, including protein N-terminal acetylation (default), oxidation of methionine (default), deamidation of asparagine, peptide N-terminal glutamine to pyroglutamate and replacement of three protons by iron (cation Fe[III]) on aspartate and glutamate. Unmodified and modified peptides were stringently filtered by setting a minimum score of 10 and 20, and a minimum delta score of 20 and 40, respectively. First-search mass tolerance was set to 10 ppm, and the maximum charge state of considered precursors was set to 6. Label-free quantification (LFQ) was enabled, Fast LFQ was disabled. iBAQ was enabled. Matching between runs was enabled with a match time window of 1 min and an alignment time window of 20 min. Matching was allowed only between same-condition replicates. Data were filtered by posterior error probability to achieve a false-discovery rate of <1% (default), at the peptide-spectrum match, protein assignment and site-decoy levels.

MS data statistics

All statistical data handling was performed using the Perseus software[60], including data filtering, $\log_2$-transformation, imputation of missing values (down shift 1.8 and width 0.15) and two-tailed two-sample Student's *t*-testing with permutation-based false-discovery rate control. To determine the relative concentration of all proteins in the samples, LFQ-normalized intensity values for each protein were adjusted by molecular mass. To approximate the absolute copy numbers, we extracted known protein copy numbers based on the 'LB' condition as reported previously[61], $\log_2$-transformed them and aligned them to the molecular-mass-adjusted LFQ intensity values from our own data, resulting in 1,901 out of 2,418 quantified protein groups receiving a known copy-number value ($R^2 = 0.6129$). We next subtracted the overall median from all $\log_2$ values and determined the absolute delta between the values of each pair. Out of all pairs, 459 had a $\log_2$ delta of <0.5, which we considered to be a 'proteomic ruler'. Linear regression was performed on the remaining pairs ($R^2 = 0.9868$) to determine a conversion factor between molecular-mass-adjusted LFQ intensity and absolute copy number.

TIRF microscopy cultivation conditions

Overnight cultures of *E. coli* strains expressing ZorB–HaloTag and ZorC or ZorD translational fusions to mNG were incubated shaking at 180 rpm in LB Lennox containing 20 mM $MgSO_4$, 5 mM $CaCl_2$ and supplemented with 12.5 µg ml^{-1} chloramphenicol at 30 °C. The next day, a subculture was inoculated 1:100 and grown at 30 °C until an OD_{600} between 0.3–0.5 was reached. Subsequently, cells were diluted to an OD_{600} of 0.2. For HaloTag fusions, cells were washed once in PBS supplemented with a final concentration of 0.2% glucose and stained with 1 µM TMR ligand for 30 min. To remove excess of TMR after staining, cells were washed twice with PBS supplemented with 0.2% glucose. Cells were then exposed to phages at indicated MOIs or incubated untreated for 30 min in a 2 ml Eppendorf tube under shaking conditions (<650 rpm in an Eppendorf ThermoMixer) at 30 °C. For TIRF microscopy, 1 µl of cells and phage mix was spotted onto an agarose pad (1.2% in Milli-Q water of UltraPure agarose, Invitrogen) and directly imaged.

TIRF microscopy acquisition and data evaluation

TIRF microscopy was performed using the Nikon Eclipse Ti2 inverted microscope equipped with the ILAS 2 TIRF module (Gataca Systems) and a TIRF ×100/1.49 NA oil objective. The samples were excited at 50 ms exposure for 14 frames using a 488 nm laser (power of 550 µW with a sensor area of 283.5 mm^2) at 80% and emission was recovered through a quad TIRF filter cube (emission: 502–549 nm). For the HaloTag constructs, five frames were acquired using a 561 nm laser (power of 930 µW with a sensor area of 283.5 mm^2) at 80% with an exposure time of 50 ms. Emission was recovered through a quad TIRF filter cube (emission: 581–625 nm, followed by an emission wheel filter: 580–611 nm). The second frame (for HaloTag constructs, the third frame) in the fluorescent channel of the acquired TIRF microscopy images was denoised using the Nikon software package Denoise.ai. Subsequently, fluorescent maxima of ZorB, ZorC or ZorD translational fusions to mNG were detected using MicrobeJ[62] run in Fiji[63]. For the co-localization analysis of ZorB–HT with either ZorC–mNG or ZorD–mNG, ZorB focus detection was performed using ilastik[64]. Subsequently, the obtained ZorB binary masks were overlayed in MicrobeJ onto the original image and cells were manually detected in MicrobeJ. From this, parameters such as ZorB focus count, intensity in both HaloTag and mNG channel per cell and mNG signal of the entire cell body were extracted. We defined co-localization of ZorB with a cytoplasmic Zorya component (ZorC or ZorD) if mNG signal within the detected ZorB complex area was 1.5-fold greater than the average cytoplasmic mNG fluorescence of the complete cell. Finally, cells containing at least one ZorB with either ZorD or ZorC were considered to be co-localized. Statistics were calculated in Prism GraphPad 9 using the in-built unpaired *t*-tests or one-way ANOVA.

Time-lapse microscopy and data evaluation

Cells were prepared as stated above for TIRF microscopy. In the case of the Bas54-*parS* phage infection experiments, the *E. coli* cells expressed an IPTG-inducible $ParB^{P1}$-mScarlet fusion protein from pALA2703-mScarlet-$ParB^{P1}$ in addition to EcZorI or the empty vector control[65]. After diluting the subculture to an OD_{600} of 0.2, cells were exposed to phages at an MOI of 5 (or unknown if labelled) for 5 min (or none) under shaking conditions and, subsequently, 1 µl was spotted onto a 1.2% (w/v) agarose pad (dissolved in LB:MQ with a ratio of 1:5). Expression of $ParB^{P1}$–mScarlet was induced by addition of 500 µM IPTG in the subculture. For the time-lapse experiments involving CCCP, cells were exposed to 30 µM CCCP and Bas24 (MOI of 5) for 30 min (or CCCP only) under shaking conditions and followed by spotting 1 µl onto an agarose pad (composition as above, supplemented with 30 µM CCCP). Microscopy slides were then mounted into an incubation chamber (preheated to 37 °C) surrounding the microscope. Acquisition was performed using the Nikon Eclipse Ti2 inverted microscope equipped with a CFI Plan Apochromat DM ×60 Lambda oil Ph3/1.40 objective and phase-contrast images were obtained every 2 min for the indicated time span. Obtained time-lapse images were segmented using ilastik to detect the cell areas[63]. Time-lapse microscopy videos of DNA-labelled Bas24 and Bas54-*parS* phage infection experiments were *x*,*y* drift-corrected using the Fiji Fast4DReg plugin[66,67]. The intensity decay due to photobleaching of the SYTOX Orange fluorescent channel was subsequently corrected using the Bleach Corrector Fiji plugin[68]. The final graphs and videos were prepared using a custom Python script.

Construction of the phage Bas54-*parS*

A *parS* site was introduced between gp69 and gp70 of phage MaxBurger (Bas54)[15] by homologous recombination with a synthetic template followed by CRISPR–Cas13a selection against the parental wild type using a setup similar to the procedure described elsewhere[69]. In a first step, the *parS* sequence (TCGCCATTCAAATTTCACTATTAACTGACTGTTTTTAA AGTAAATTACTCTAAAATTTCAAGGTGAAATCGCCACGATTTCAC) was inserted between plasmid-encoded homology arms of 100–200 bp length by PCR[70], creating pAH210_Bas54_parS_H1. Subsequently, an *E. coli* K-12 host was transformed with this plasmid and then infected with phage Bas54 to enable homologous recombination. The lysate from this infection—containing wild-type and recombinant phages—was then subjected to CRISPR–Cas13a selection against the parental wild type using a crRNA targeting CTCTGAAGACCTCCAGTAGTAAGAT GTAAGT (5′–3′), which includes the 3′ end of gp69 and the downstream region, which is disrupted by insertion of the *parS* site. CRISPR–Cas13a selection was performed using a two-plasmid setup of pAH221 (expressing LbuCas13a) and pAH218_LbuCas13a_parS_H1 (expressing the crRNA). Plaques growing after CRISPR–Cas13a selection were screened for successful insertion of the *parS* site by PCR and the sequence of the recombinant was confirmed by Sanger sequencing.

Bioinformatic analyses of ZorA motor and tail lengths

PADLOC (v.1.1.0)[34] with PADLOC-DB (v.1.4.0) was used to identify Zorya systems in RefSeq v209[35] bacterial and archaeal genomes. Of the systems identified, we excluded those containing pseudogenes or more than one copy of each Zorya gene, or systems with non-canonical gene arrangements (Zorya genes are typically on the same strand, in type-specific conserved gene orders, for example, *zorABCD*, *zorABE* or *zorGABF* for types I–III, respectively). To reduce redundancy due to highly related genome sequences in the RefSeq database, we next selected representative Zorya systems by first clustering the sequences (using MMseqs2 (v.14.7e284)[36] with the options --min-seq-id 0.3 --coverage 0.8) of the proteins encoded by the three adjacent open reading frames on either side of each Zorya system, then randomly selecting one representative system for each unique genetic context observed. The ZorA and ZorB sequences from the representative Zorya systems were then clustered using MMseqs2 with the options --min-seq-id 0.3 --coverage 0.95. The structures were predicted for each MMseqs2 cluster representative of each ZorA and ZorB family using ColabFold (v.1.5.2)[71] with the options --num-recycle 3 --num-models 1 --model-type auto --amber --use-gpu-relax. Structure predictions were run as $ZorA_5ZorB_2$ multimers. The resulting structures were inspected manually (using PyMOL, v.2.5.4)[52] to identify the start of the ZorA tail. The rest of the sequences in each cluster were aligned to the representative sequence using MUSCLE (v.5.1)[37] using the Parallel Perturbed ProbCons algorithm (default) or the Super5 algorithm if the cluster contained more than 100 sequences. The start of the tail for the representative was used to infer the start of the tail for each other protein in the respective alignment.

Figure preparation

Structural figures were prepared using ChimeraX v.1.8[48], PyMOL v.2.5.4 and v.3.0.2[52], GraphPad Prism 9 or GraphPad Prism 10 (one-way ANOVA followed by Dunnett's multiple-comparisons test)[33] and Adobe Illustrator[72]. The ion-permeation pathway shown in ZorAB was analysed using MOLEonline[73]. The hydrophobicity and polarity of the ZorAB tail was calculated using MOLEonline[73]. The electrostatic potential maps were calculated using the APBS[74] electrostatic Plugin integrated inside PyMOL. Full gel/blot images for all relevant figures are provided 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

Atomic coordinates for ZorAB WT, ZorA(E86A/E89A)–ZorB, ZorA(Δ359–592)–ZorB and ZorA(Δ435–729)–ZorB were deposited at the Protein Data Bank (PDB) under accession codes 8QYD, 8QYH, 8QYK and 8QYY, respectively. The corresponding electrostatic potential maps were deposited at the Electron Microscopy Data Bank (EMDB) under accession codes EMD-18751, EMD-18754, EMD-18756 and EMD-18766, respectively. The local refinement map of ZorB PGBD in ZorAB WT was deposited at the EMDB under accession code EMD-18752. Atomic coordinates for ZorC were deposited at the PDB under accession code 8R68. The corresponding electrostatic potential maps were deposited at the EMDB under accession code EMD-18848. Atomic coordinates for ZorD apo form and its complex with ATP-γ-S were deposited at the PDB under accession codes 8QY7 and 8QYC, respectively. The corresponding electrostatic potential maps were deposited at the EMDB under accession codes EMD-18747 and EMD-18750. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[75] partner repository with the dataset identifier PXD047450. Source data are provided with this paper.

34. 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).
35. Haft, D. H. et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.* **46**, D851–D860 (2018).
36. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
37. Edgar, R. C. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 6968 (2022).
38. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
39. Parks, D. H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).
40. Kim, D., Park, S. & Chun, J. Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. *J. Microbiol.* **59**, 476–480 (2021).
41. McBride, T. M. et al. Diverse CRISPR-Cas complexes require independent translation of small and large subunits from a single gene. *Mol. Cell* **80**, 971–979 (2020).

42. Martínez-García, E. et al. SEVA 4.0: an update of the Standard European Vector Architecture database for advanced analysis and programming of bacterial phenotypes. *Nucleic Acids Res.* **51**, D1558–D1567 (2023).
43. Thoma, S. & Schobert, M. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol. Lett.* **294**, 127–132 (2009).
44. Jackson, S. A., Fellows, B. J. & Fineran, P. C. Complete genome sequences of the *Escherichia coli* donor strains ST18 and MFDpir. *Microbiol. Resour. Announc.* **9**, e01014-20 (2020).
45. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28 (1990).
46. 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).
47. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
48. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
50. Lugmayr, W. et al. StarMap: a user-friendly workflow for Rosetta-driven molecular structure refinement. *Nat. Protoc.* **18**, 239–264 (2023).
51. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
52. The PyMOL Molecular Graphics System, v.2.5. pymol.org/2 (Schrödinger, 2021).
53. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
54. Kirchweger, P. et al. Structural and functional characterization of SiiA, an auxiliary protein from the SPI4-encoded type 1 secretion system from *Salmonella enterica*. *Mol. Microbiol.* **112**, 1403–1422 (2019).
55. Batth, T. S. et al. Protein aggregation capture on microparticles enables multipurpose proteomics sample preparation. *Mol. Cell. Proteomics* **18**, 1027–1035 (2019).
56. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
57. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).
58. Chi, H. et al. Comprehensive identification of peptides in tandem mass spectra using an efficient open search engine. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4236 (2018).
59. Röst, H. L. et al. OpenMS: a flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **13**, 741–748 (2016).
60. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
61. Schmidt, A. et al. The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
62. Ducret, A., Quardokus, E. M. & Brun, Y. V. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat. Microbiol.* **1**, 16077 (2016).
63. Schindelin, J. et al. Fiji—an open source platform for biological image analysis. *Nat. Methods* **9**, 676–682 (2012).
64. Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).
65. Li, Y. & Austin, S. The P1 plasmid is segregated to daughter cells by a 'capture and ejection' mechanism coordinated with *Escherichia coli* cell division. *Mol. Microbiol.* **46**, 63–74 (2002).
66. Laine, R. F. et al. NanoJ: a high-performance open-source super-resolution microscopy toolbox. *J. Phys. Appl. Phys.* **52**, 163001 (2019).
67. Jw, P. et al. Fast4DReg—fast registration of 4D microscopy datasets. *J. Cell Sci.* **136**, jcs260728 (2023).
68. K, M. Bleach correction ImageJ plugin for compensating the photobleaching of time-lapse sequences. *F1000Res.* **9**, 1494 (2020).
69. Adler, B. A. et al. Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing. *Nat. Microbiol.* **7**, 1967–1979 (2022).
70. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8**, 91 (2008).
71. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
72. Adobe Illustrator Version 25.3 (Adobe, 2021).
73. Pravda, L. et al. MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res.* **46**, W368–W373 (2018).
74. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).
75. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
76. Li, L. et al. Structure of the ISW1a complex bound to the dinucleosome. *Nat. Struct. Mol. Biol.* **31**, 266–274 (2024).

Acknowledgements The Novo Nordisk Foundation Center for Protein Research is supported financially by the Novo Nordisk Foundation (NNF14CC0001). N.M.I.T. acknowledges support from an NNF Hallas-Møller Emerging Investigator grant (NNF17OC0031006), an NNF Hallas-Møller Ascending Investigator grant (NNF23OC0081528) and an NNF Project grant (NNF21OC0071948). N.M.I.T. is also a member of the Integrative Structural Biology Cluster (ISBUC) at the University of Copenhagen. H.H. acknowledges support from Lundbeck Foundation postdoc fellowship R347-2020-2429. S.A.J. acknowledges support from the Health Research Council of New Zealand (Sir Charles Hercus Fellowship) and from Bioprotection Aotearoa (Tertiary Education Commission, New Zealand). We acknowledge the use of the New Zealand eScience Infrastructure (NeSI) high-performance computing facilities in this research. T.C.D.H. was supported by a University of Otago Doctoral Scholarship. P.F.P. and M.E. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 548567920 in the framework of the priority program SPP2330. M.E. acknowledges funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 864971) and by the Max Planck Society as Max Planck Fellow. V.K.-S. and N.R.R. acknowledge support from Novo Nordisk Bioscience PhD program, NNF0069780 and NNF0078229, respectively. Y.W. acknowledges support from the National Key Research and Development Program of China (2021YFF1200404), the National Science Foundation of China (32371300) and computational resources from the Information Technology Center and State Key Lab of Computer-Aided Design (CAD) & Computer Graphics (CG) of Zhejiang University. The M.L.N. laboratory was supported by the Novo Nordisk Foundation (NNF14CC0001), The Danish Council of Independent Research (grant agreement numbers 4002-00051, 4183-00322A and 8020-00220B) and The Danish Cancer Society (grant agreement R146-A9159-16-S2). A.H. and D.H. acknowledge funding by a Starting Grant (TMSGI3_211369) of the Swiss National Science Foundation (SNSF). We thank the staff at the Danish Cryo-EM Facility at the Core Facility for Integrated Microscopy (CFIM) at the University of Copenhagen and Tillmann Pape and Nicholas Heelund Sofos for support during data collection; and B. Lopez Mendez and M. I. Rasmussen for their support in MS and mass photometry.

Author contributions N.M.I.T. and H.H. conceived the project. H.H., A.R.-E., N.R.R. and F.J.O.M. performed molecular biology and mutagenesis. H.H. expressed, purified, optimized and prepared cryo-EM grids, collected cryo-EM data and determined all of the structures presented in this study. P.F.P. performed and analysed time-lapse, TIRF and labelled phage microscopy experiments. P.F.P. together with M.E. interpreted the microscopy studies. M.E. labelled phages and performed the western blotting of the ZorB–HaloTag fusions. S.A.J. and T.C.D.H. performed phage infectivity, adsorption and phage-burst assays. S.A.J. performed cell survival, conjugation and transformation assays. S.A.J. and L.J.P. performed bioinformatic analyses. N.R.R. and F.J.O.M. assisted with protein purification. N.R.R., F.J.O.M. and H.H. optimized the nuclease and EMSA experiments. N.R.R. obtained the ZorC–DNA complex dataset. A.R.-E. and V.K.-S. purified phage genomes. D.H. and A.H generated the Bas54-*parS* phage. Y.W. and Y.Y. helped with ZorA tail structure modelling. H.H. prepared samples for MS. I.A.H. and M.L.N. performed MS and analysed the data. N.R.R., H.H. and I.S. performed PG-binding assays. H.H. and S.A.J. prepared figures and H.H. wrote the first draft of the manuscript together with N.M.I.T. and S.A.J. with input from all of the authors. This draft was then edited by M.E., P.F.P. and R.M.B., and all of the other authors. H.H., S.A.J., M.E., P.F.P. and N.M.I.T. supervised the generation of the manuscript.

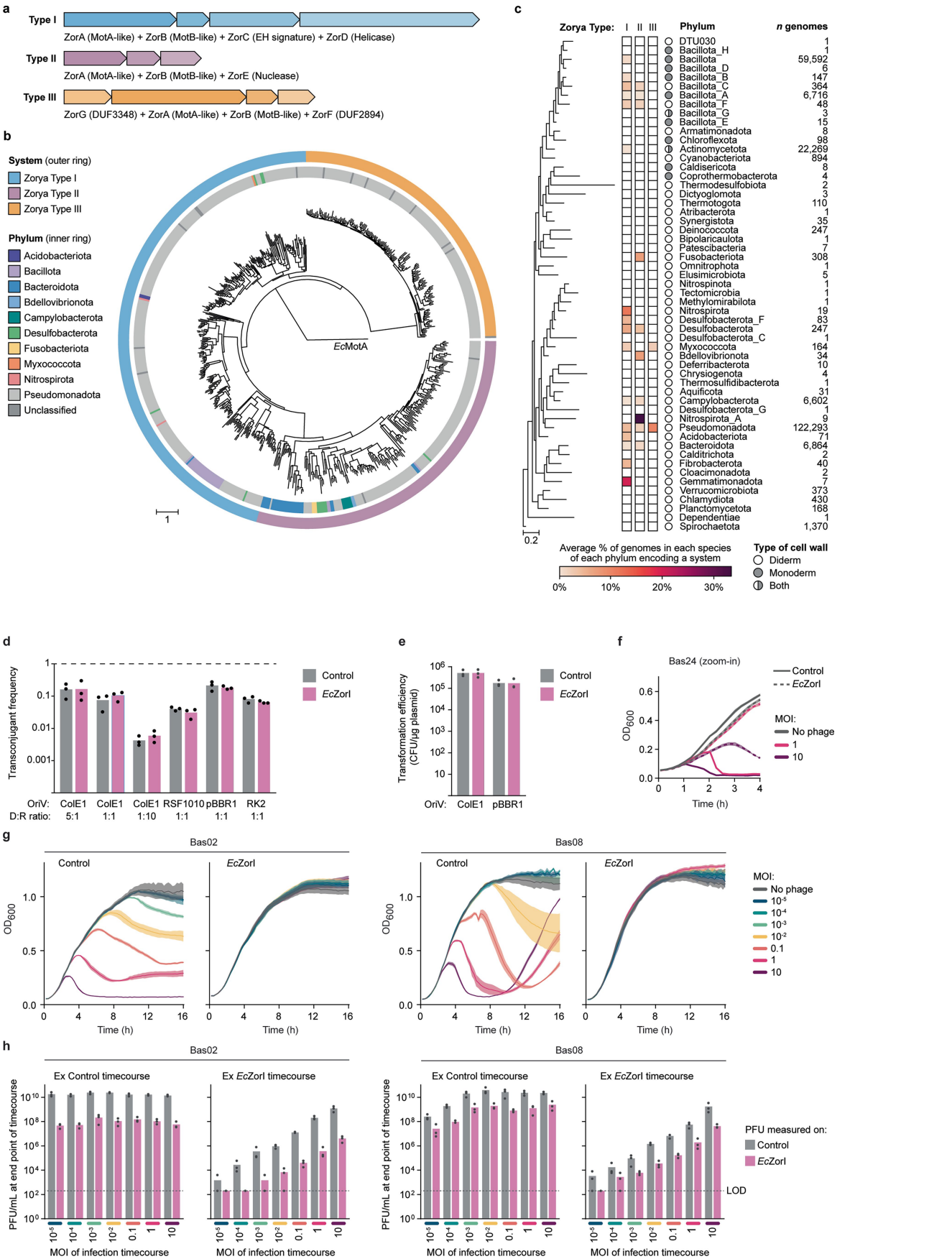
Competing interests The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08493-8.

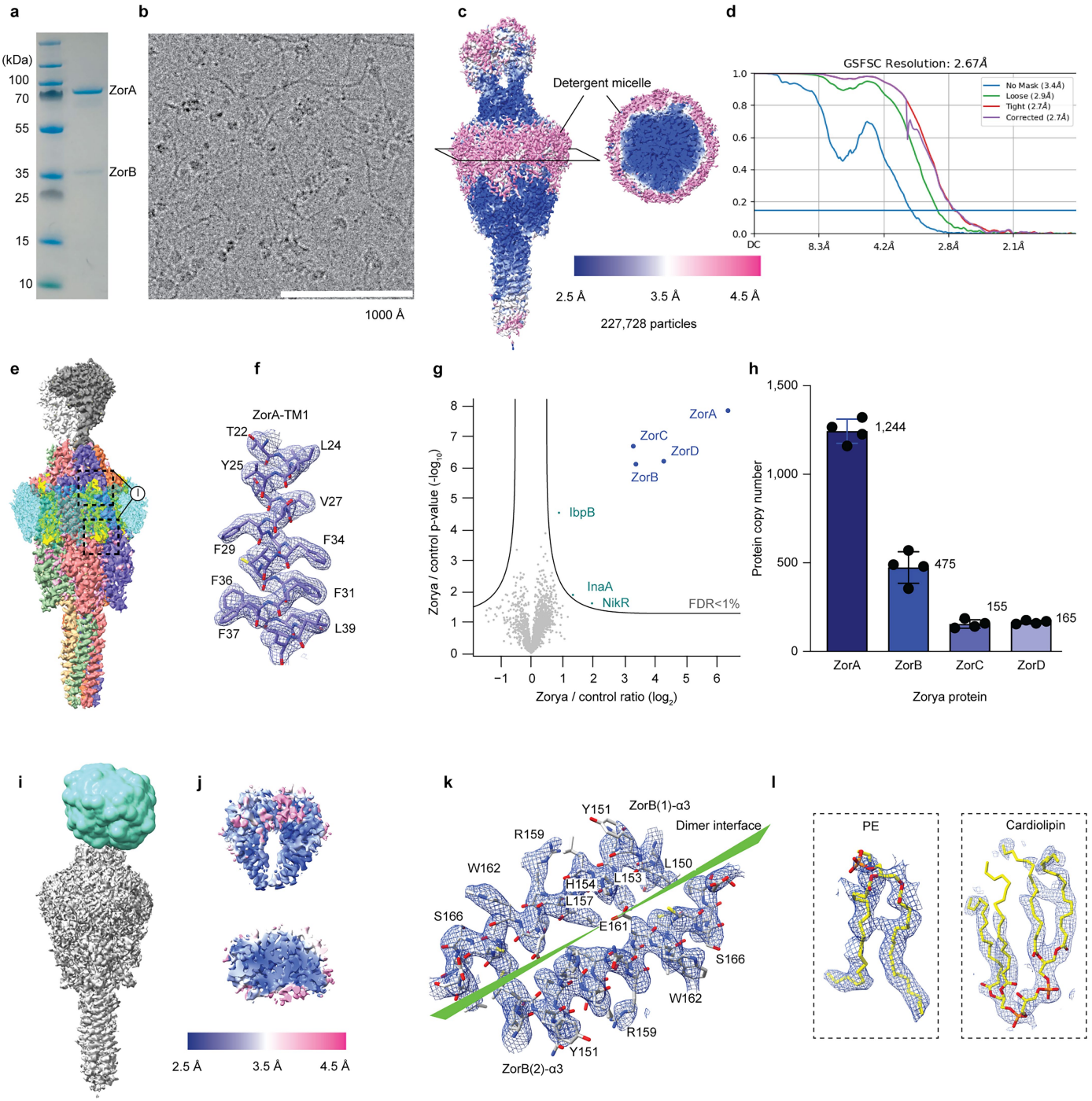
Correspondence and requests for materials should be addressed to Haidai Hu, Marc Erhardt, Simon A. Jackson or Nicholas M. I. Taylor.

Peer review information *Nature* thanks David Taylor and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

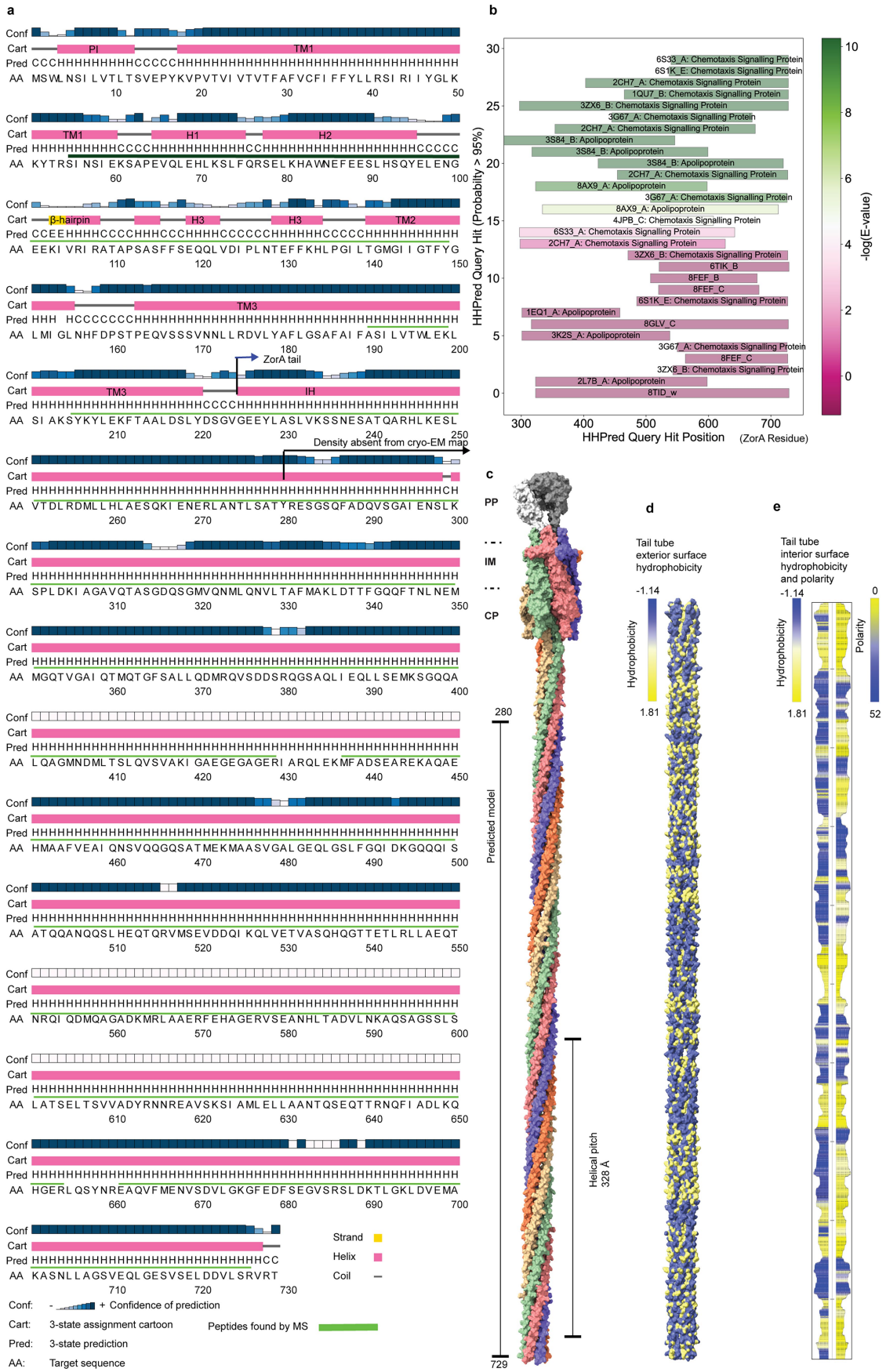


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | *E. coli* Zorya type I protects against phage invasion but not bacterial conjugation or plasmid transformation. a, Zorya system gene arrangements for each Zorya type (I–III), with typical gene annotations shown. **b**, Phylogeny of the ZorA motor sequence (excluding the ZorA tail) rooted with the *E. coli* MotA (*Ec*MotA) sequence. **c**, Taxonomic analyses of Zorya hosts, using the GTDB phyla-level taxa. Each phylum was assigned as either possessing a single membrane cell envelope (Monoderm) or double membrane envelope (Diderm). In some cases, there are both monoderm and diderm species within phyla, which are labelled as 'Both'. The number of genomes analysed for each phyla (*n* genomes) is based on the genomes present in the GTDB v214.1 and RefSeq v209. **d**, The impact of *Ec*ZorI on the uptake of plasmid DNA via conjugation from an *E. coli* donor strain, measured as the transconjugant frequency (number of transconjugants/total recipients). Four plasmids with different origins of replication (OriV) were tested (ColE1, RSF1010, pBBR1 and RK2), at the indicated donor to recipient cell ratios (D:R) for the matings. Data represent the mean of three replicates. **e**, The impact of *Ec*ZorI on the uptake of plasmid DNA via transformation. Chemically competent *E. coli* without (control; empty vector) or with *Ec*ZorI were transformed with plasmids possessing either ColE1 or pBBR1 origins of replication. Data represent the mean of three replicates, with each replicate being a different batch of competent cells. **f**, Infection time courses for liquid cultures of *E. coli*, with and without *Ec*ZorI, infected at different multiplicities of infection (MOI) of phage Bas24 (early timepoints for MOI 1 and 10, from Fig. 1e). **g**, Infection time courses for liquid cultures of *E. coli*, with and without *Ec*ZorI, infected at different multiplicities of infection (MOI) of phage Bas02 and Bas08. **h**, Phage titres at the end timepoint for each sample from the infection time courses (**g**), measured as EOP on indicator lawns of *E. coli* either without (control) or with *Ec*ZorI. LOD: Limit of detection. Data in **f-h** represent the means of 3 replicates and the shaded regions represent the SEM.

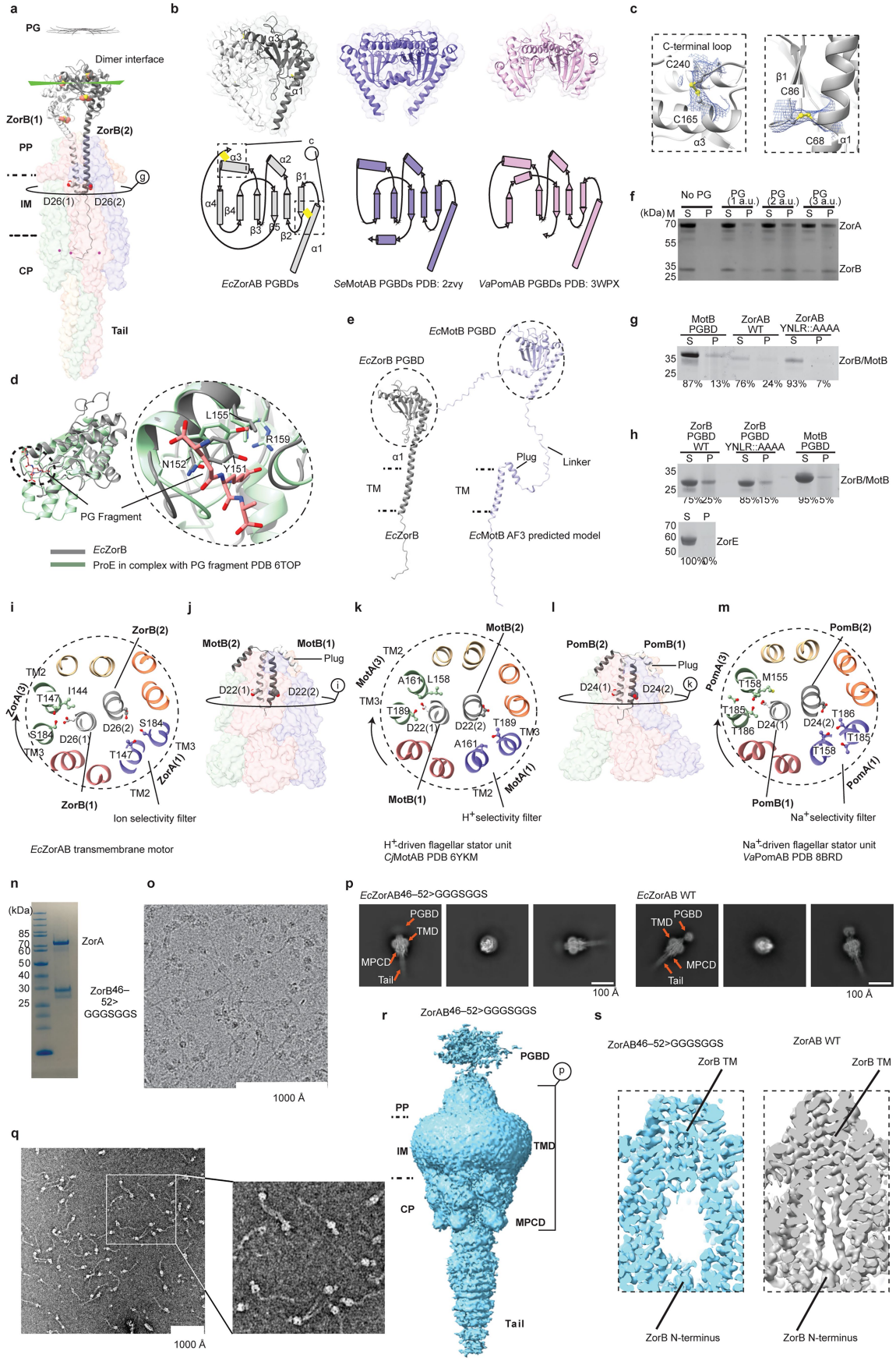


Extended Data Fig. 2 | Cryo-EM dataset processing results and resolution of *Ec*ZorAB. a, A representative SDS gel of the purified *Ec*ZorAB complex. **b**, An EM image of the *Ec*ZorAB sample under cryogenic condition. **c**, Cryo-EM density map of *Ec*ZorAB coloured by local resolution (in Å) estimated in cryoSPARC. **d**, Gold standard (0.143) Fourier Shell Correlation (GSFSC) curves of refined *Ec*ZorAB complex. **e**, Cryo-EM map of *Ec*ZorAB. **f**, Representative model segment of ZorA fit into EM density, focusing on TM1 of one of the ZorA subunits. **g**, Volcano plot analysis, visualizing ratio and significance of change between all proteins quantified by mass spectrometry in *E. coli* total lysates either transformed with p*Ec*ZorI plasmids or not (Supplementary Table 1). Significance was tested via two-tailed two-sample Student's t-testing with permutation-based FDR control, ensuring a corrected p-value of <0.01. n = 4 technical replicates derived from n = 3 culture replicates. **h**, Absolute copy number analysis of Zorya proteins expressed in *E. coli*. Determined via comparison of molecular weight-adjusted label-free quantified protein abundance values from this study, to known copy numbers reported by Schmidt et al. [57] and establishing a "proteomic ruler" for conversion of measured abundance values to approximate copy numbers (Supplementary Table 1). n = 4 technical replicates derived from n = 3 culture replicates. **i**, Soft mask used for local refinement of the ZorB PGBDs. **j**, Local refinement map of the ZorB PGBDs (made with the soft mask shown in **i**), coloured by local resolution. **k**, A representative of a model segment of the ZorB PGBDs fitted into of the local refinement EM density map shown in **j**, focusing on the PGBD dimerized interface. **l**, Fit of lipids found in the TMD of ZorA in the *Ec*ZorAB cryo-EM map. Images in **a** and **b** are representatives of at least 3 replicates.



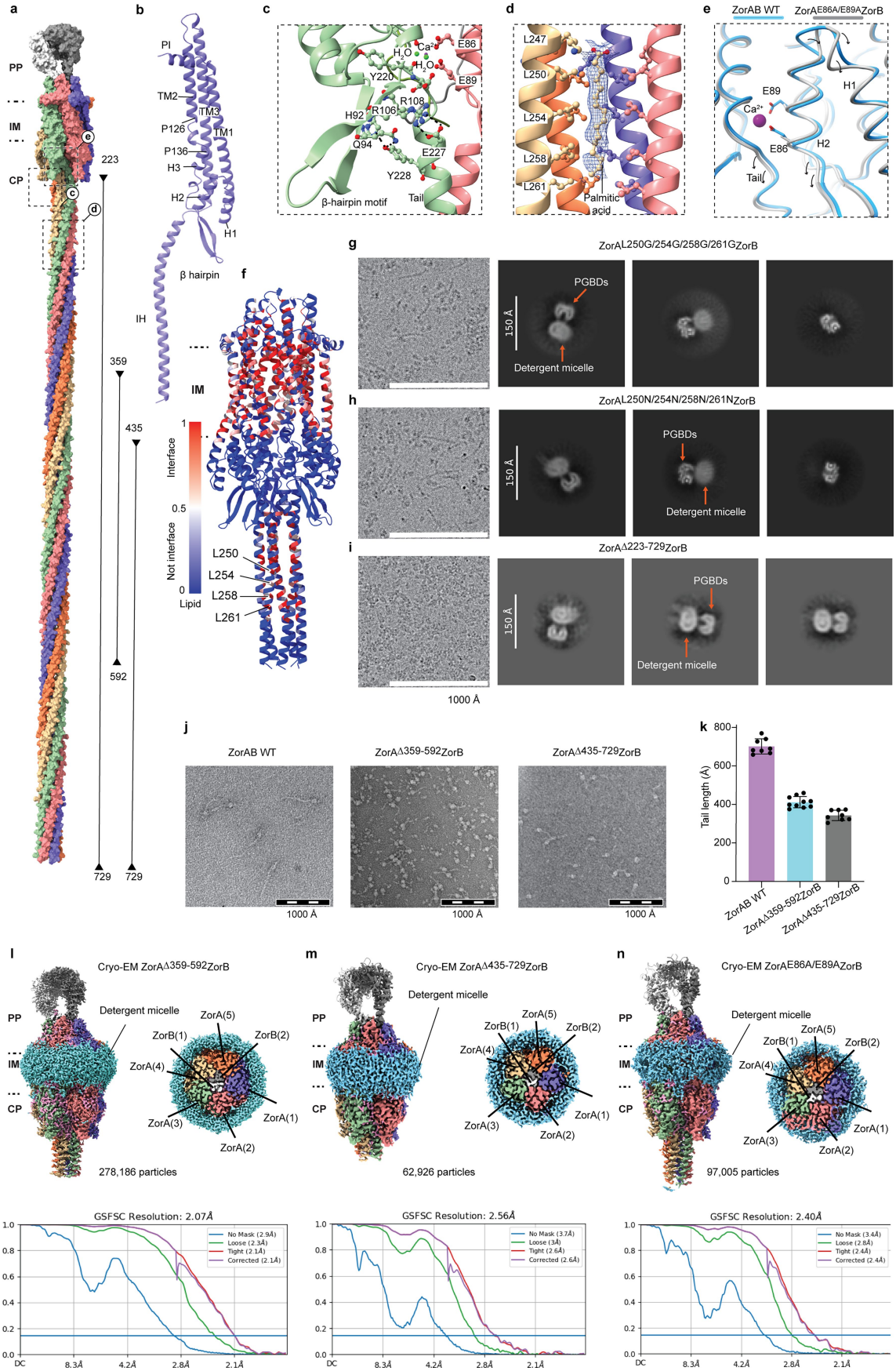
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | *Ec*ZorA tail secondary structural prediction and a complete composite model of *Ec*ZorAB complex. a, Amino acids and secondary structural predictions (Psipred) of the *Ec*ZorA. The peptides found by mass spectrometry that covered ZorA protein are indicated as green lines above the amino acids. **b**, Top hits from an HHpred sequence homology search of the ZorA tail are shown. **c**, A composite model of *Ec*ZorAB with the ZorA tail folding into a pentameric super coiled-coil, with the helical pitch of the tail α-helix shown. **d**, Hydrophobicity of the tail tube exterior surface calculated by ChimeraX. **e**, Hydrophobicity and polarity of the tail tube interior surface calculated by MOLE*online*.



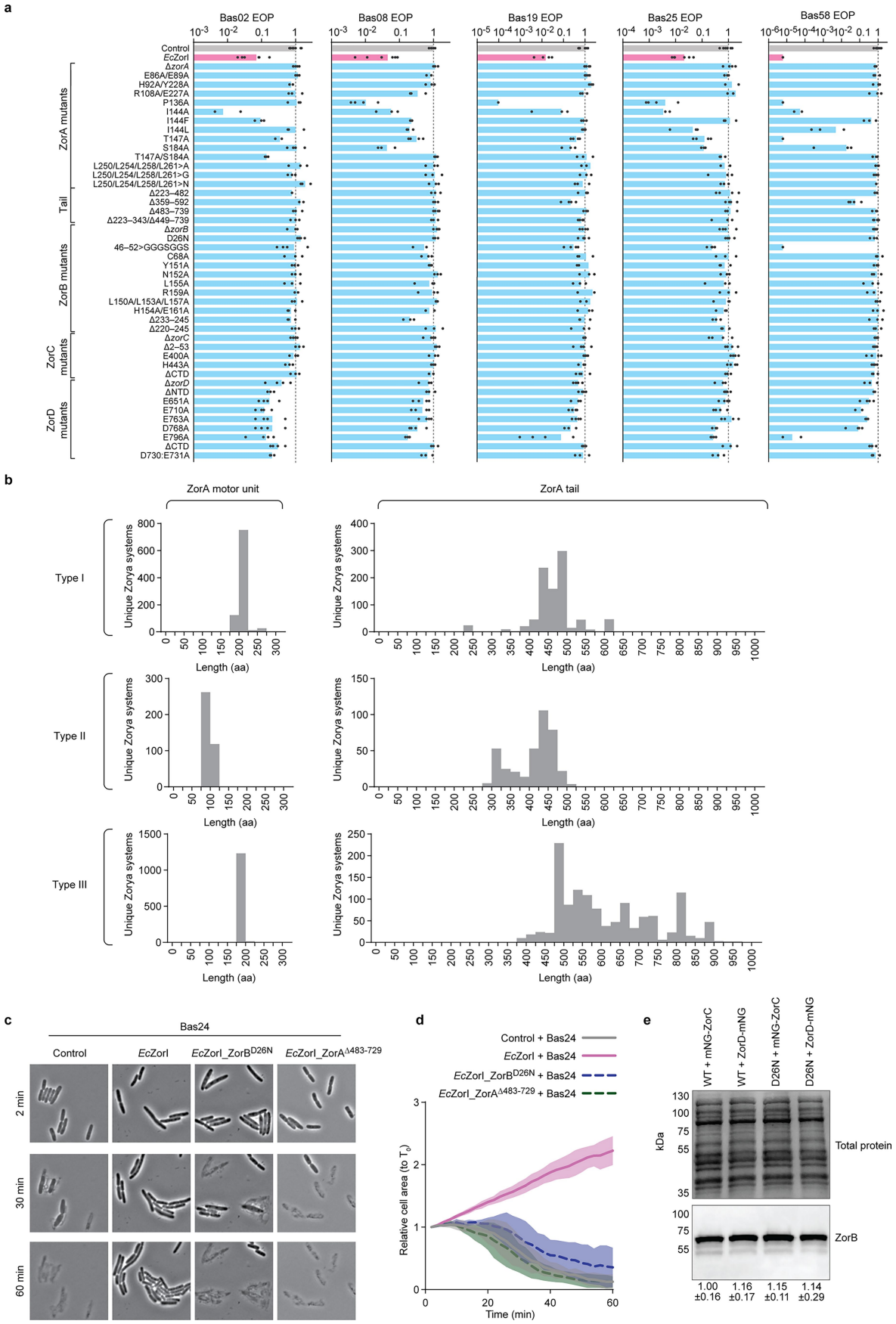
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Structure of *Ec*ZorAB and its function as a peptidoglycan binding rotary motor. a, Cartoon representation of the *Ec*ZorAB complex in an inactive state, with the ZorB dimer interfaced highlighted. **b**, Topology diagrams of ZorB PGBDs and isolated crystal structures of the flagellar stator unit MotB and PomB PGBDs, indicating a conserved folding architecture. **c**, The two disulfide bonds identified from ZorB PGBDs, with the EM map overlapped. **d**, Structural comparison of PGBD of *Ec*ZorB with that of ProE that in complex with PG fragment, with the zoom in view highlighting the conserved residues from *Ec*ZorB that are likely involved in PG binding. **e**, Structural comparison of *Ec*ZorB PGBD and AlphaFold3 predicted *Ec*MotB PGBD, highlighting *Ec*ZorB PGBD is fused without a linker to the ZorB TM. **f**, In vitro pull-down assay of isolated *Ec*ZorAB complex with purified *E. coli* PG. **g**, In vitro pull-down assay of isolated *Ec*ZorAB complex and *Ec*ZorAB complex with mutations in the ZorB PGBD (ZorAB$^{Y151A/N152A/L155A/R159A}$) with purified *E. coli* PG. **h**, In vitro pull-down assay of isolated ZorB PGBD, mutant ZorB PGBD (ZorB$^{Y151A/N152A/L155A/R159A}$ PGBD), MotB PGBD (positive control) and ZorE (negative control) with purified *E. coli* PG. **i**, Cross-section view of the *Ec*ZorAB TMD, showing the surrounding residues of the two Asp26 from ZorB. **j**, Cartoon representation of the cryo-EM structure of the proton-driven flagellar stator unit MotAB from *Campylobacter jejuni* (*Cj*MotAB) in its inactive state, with the MotB plug motif highlighted. **k**, Cross-section view of the *Cj*MotAB TMD, showing the surrounding residues of the two Asp22 from MotB. **l**, Cartoon representation of the cryo-EM structure of the sodium-driven flagellar stator unit PomAB from *Vibrio alginolyticus* (*Va*PomAB) in its inactive state. **m**, Cross-section view of *Va*PomAB TMD, showing the surrounding residues of the two Asp24 from PomB. The absence of the strictly conserved threonine residue on ZorA TM3 (**k**) required for sodium ion binding, indicates that *Ec*ZorAB is a proton-driven stator unit. **n**, A representative of an SDS gel of the purified *Ec*ZorAB 'linker mutant' complex (with ZorB residues 46–52 replaced by a GGGSGGS linker). **o**, A representative cryo-EM image of *Ec*ZorAB 'linker mutant' sample. **p**, Representative 2D classes of the *Ec*ZorAB 'linker mutant' in comparison with that of the *Ec*ZorAB wild type, highlighting the flexibility of the ZorB PGBDs in the mutant. **q**, A representative negative stain EM image of *Ec*ZorAB 'PG-binding mutant' sample. **r**, Low pass filter of the cryo-EM density map of the *Ec*ZorAB linker mutant after non-uniform refinement. **s**, Transmembrane helix density of the *Ec*ZorAB 'linker mutant' and that in the wild type *Ec*ZorAB. Images in **f**, **g**, **h**, **n**, **o**, **q** are representatives of at least 3 replicates with similar results.



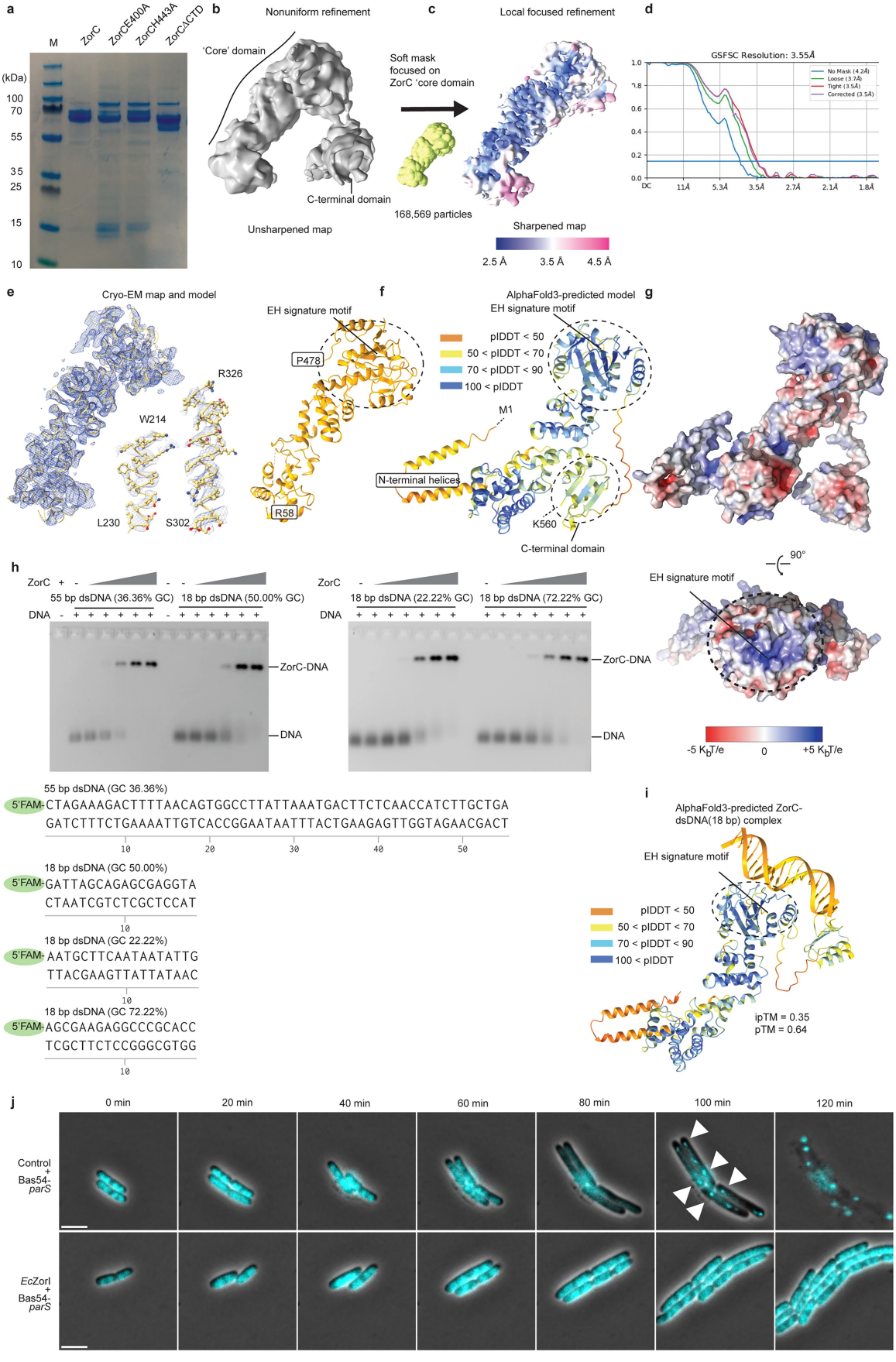
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | *Ec*ZorAB Ca^{2+} binding site and tail influence ZorAB motor assembly and function. a, ZorA tail truncations indicated in the composite model of *Ec*ZorAB complex. **b**, Cartoon representation of the *Ec*ZorAB ZorA single subunit. **c**, Interaction between the beginning of the ZorA tail and the β-hairpin motif. **d**, Extra density found inside the tail from cryo-EM map, which was modeled as a palmitic acid molecule, with the amino acids involved in the interactions indicated. **e**, Structural comparison of the ZorA wild type (cyan) and ZorA Ca^{2+} binding site mutation ($ZorA^{E86A/E89A}$, grey), the arrows highlight the changes from wild type to the mutant. **f**, Predicted ZorA lipid binding sites using PeSTo. **g**, An EM image of the *Ec*$ZorA^{L250G/L254G/L258G/L261G}$ZorB mutant under cryogenic condition and representative 2D classes. **h**, An EM image of the *Ec*$ZorA^{L250N/L254N/L258N/L261N}$ZorB mutant under cryogenic condition and representative 2D classes. **i**, An EM image of the $ZorA^{\Delta 223-729}$ZorB mutant under cryogenic condition and representative 2D classes. **j**, Negative staining images of the *Ec*ZorAB wild type, ZorA tail middle deletion ($ZorA^{\Delta 359-592}$), ZorA tail tip deletion ($ZorA^{\Delta 435-729}$). **k**, The tail lengths of the *Ec*ZorAB wild type, ZorA tail middle deletion ($ZorA^{\Delta 359-592}$), ZorA tail tip deletion ($ZorA^{\Delta 359-592}$) as measured in (**g**). Data represent the mean of at least eight measurements (data points indicate measurements), and error bars represent the standard error of the mean (SEM). **l**-**n**, Cryo-EM maps and resolutions of ZorA mutants with gold standard (0.143) Fourier Shell Correlation (GSFSC) curves.



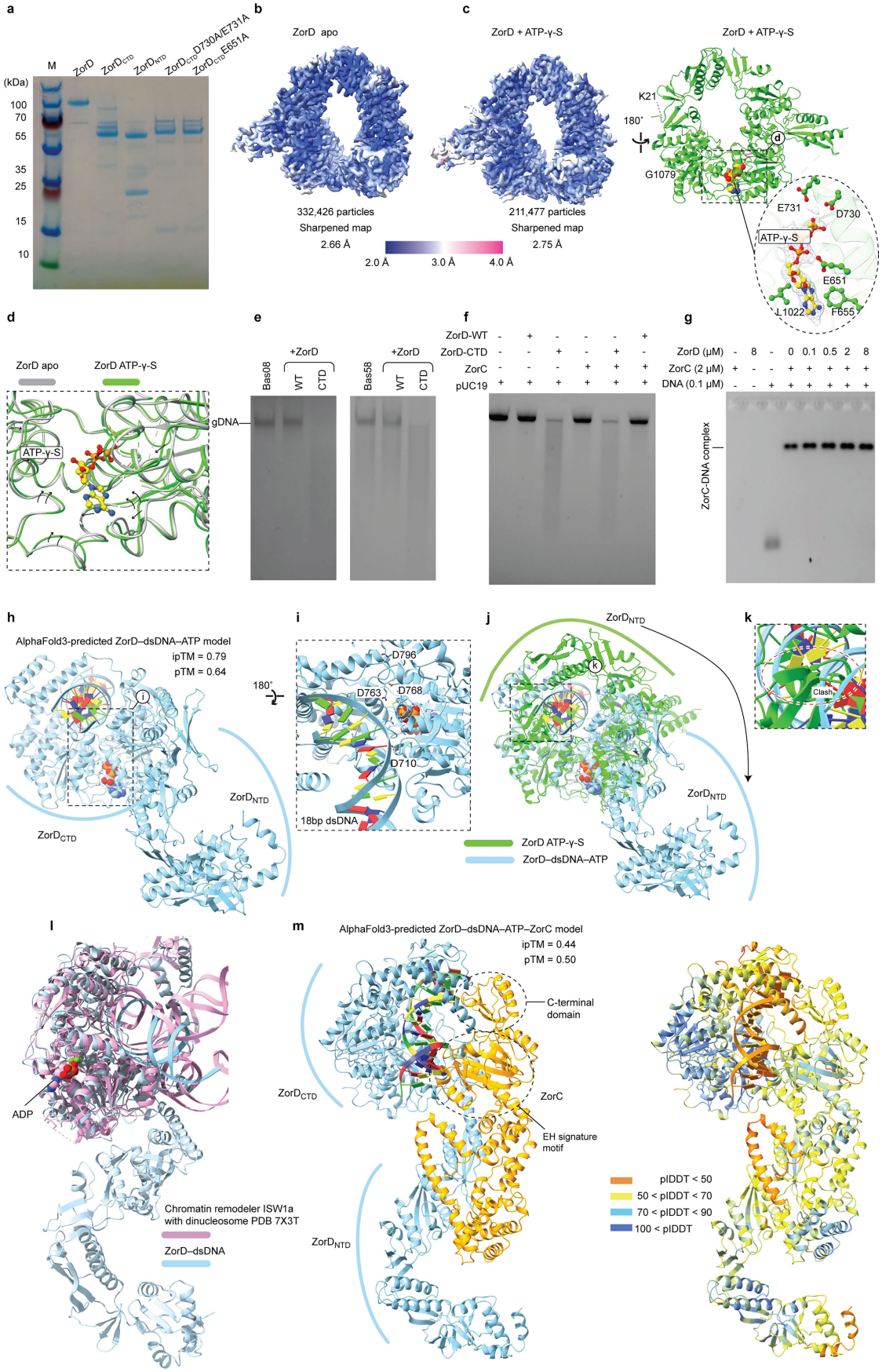
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | The effects of *Ec*Zorya mutations on *Ec*ZorI-mediated anti-phage defence and long ZorA tails are conserved amongst Zorya system types in diverse species. a, Effects of ZorA, ZorB, ZorC and ZorD mutations on *Ec*ZorI-mediated anti-phage defence, as measured using EOP assays with phages Bas02, Bas19 and Bas25. Data represent the mean of at least 3 replicates (data points indicate replicates) and are normalized to the control samples lacking *Ec*ZorI. **b**, The ZorA tail lengths found in different Zorya system types. Motor and tail lengths were determined by inspecting the predicted structures of several representative ZorA sequences, then inferring these lengths for the rest of the ZorA sequences through sequence alignment (methods). The reduce sequencing bias, unique Zorya systems encoded in RefSeq (v209) bacteria and archaea genomes were selected based on their distinct genomic context (methods). **c**, Time-lapse, phase contrast microscopy of *E. coli* cells expressing empty vector control, *Ec*ZorI wt, *Ec*ZorI ZorB^{D26N} and *Ec*ZorI ZorA$^{483–729}$ exposed to Bas24 at an MOI of 5. **d**, Quantitation of the time-lapse microscopy in (c), displaying the measured cell area relative to the first timepoint of the time-lapse. Data represent the means of three biological replicates and the shaded region indicate standard deviation. **e**, Quantitative Western blot of selected *Ec*ZorI-HaloTag translational fusions. Top: total protein stain of whole cell lysate. Bottom: Anti-HaloTag (mouse, Promega) Western blot against ZorB-HaloTag protein fusions. Mean ± standard deviation from four biological replicates.



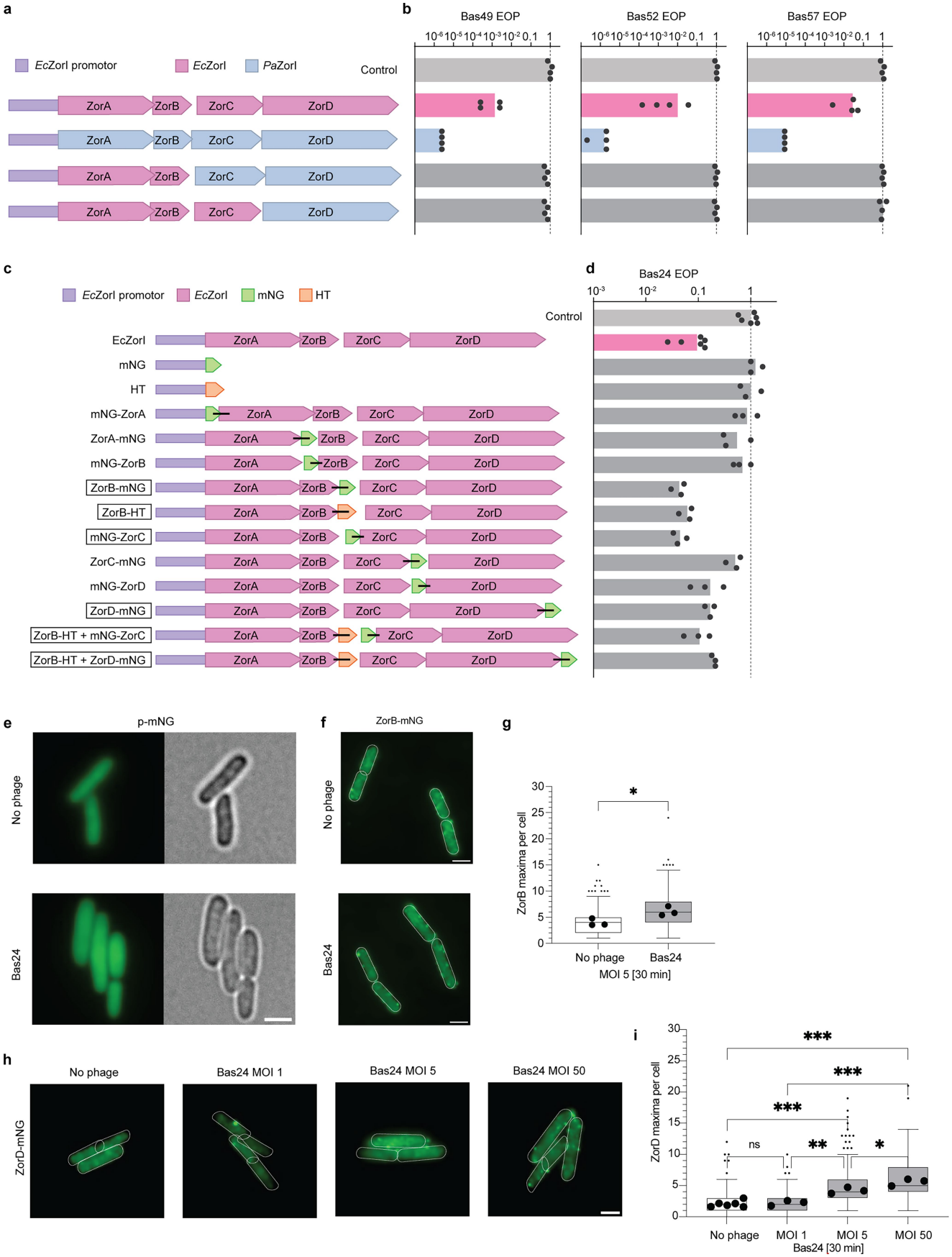
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Structural and functional investigation of *Ec*ZorC and in vivo DNA degradation. a, SDS gel of purified ZorC wild type, ZorC^{E400A}, ZorC^{H443A}, ZorC$^{\Delta CTD}$ (deletion residues 487–560). Gel is representative of at least 3 replicates. **b**, Unsharpened Cryo-EM map of *Ec*ZorC. **c**, Local refinement of the *Ec*ZorC core domain with a soft mask, with the local resolution (in Å) estimated in cryoSPARC. **d**, Gold standard (0.143) Fourier Shell Correlation (GSFSC) curves of the local refined of the *Ec*ZorC core domain. **e**, Representative of a model and segments of the ZorC fitted into EM density map. The right panel is the final model of *Ec*ZorC built from a cryo-EM map. **f**, AlphaFold3-predicted ZorC model. **g**, Electrostatic distribution of *Ec*ZorC calculated from AlphaFold3-predicted model. **h**, In vitro interaction of *Ec*ZorC with 55 bp dsDNA (36.36% GC), 18 bp dsDNA (50.00% GC), 18 bp dsDNA (22.22% GC), 18 bp dsDNA (72.22% GC). Image is representative of at least 3 replicates. DNA sequences are shown below. **i**, AlphaFold3-predicted model of ZorC in complex with 18 bp dsDNA. The colour code (per-atom confidence estimate on a 0–100 scale) in **f** and **i** are same. **j**, Representative time-lapse images of *E. coli* cells expressing ParB-mSc in the presence or absence of EcZorI, exposed to Bas54-*parS* phage. In *Ec*ZorI-null cells, ParB foci are observed prior to cell lysis, whereas *Ec*ZorI-expressing cells lack ParB focus formation and survive phage infection. Scale bar is set to 2 µm.



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

Extended Data Fig. 8 | *Ec*ZorD is autoinhibited nuclease. a, Representative of the SDS gel of the purified ZorD wild type, $ZorD_{CTD}$ (residues 503–1080), $ZorD_{NTD}$ (residues 1–502), $ZorD_{CTD}^{D730A/E731A}$ and $ZorD_{CTD}^{E651A}$. Gel is representative of at least 3 replicates. **b**, Cryo-EM map of the *Ec*ZorD apo form. **c**, Cryo-EM map and structure of *Ec*ZorD in complex with ATP-γ-S. A zoomed-in view of the ATP-γ-S binding site is depicted, with the cryo-EM map overlayed on the ATP-γ-S molecule. **d**, Structural comparison of the *Ec*ZorD apo from (grey) and *Ec*ZorD in complex with ATP-γ-S (light purple); the arrows highlight the changes from apo form to the ligand-bound form. **e**, $ZorD_{CTD}$ degrades phage Bas08 and Bas58 genomic DNA (gDNA). Gel is representative of 3 replicates. **f**, *Ec*ZorD WT and its isolated C-terminal domain nuclease activity in the absence and presence of *Ec*ZorC. **g**, *Ec*ZorC dsDNA binding activity in the absence and presence of *Ec*ZorD. **h**, AlphaFold3 predicted model of *Ec*ZorD in complex with 18 bp dsDNA and ATP, showing an alternative, open conformation of ZorD. **i**, ZorD-DNA interaction in the AlphaFold3 predicted model; key residues are highlighted. **j**, Structural superimposition of the cryo-EM structure of *Ec*ZorD with the AlphaFold3 predicted *Ec*ZorD in complex with dsDNA model. The arrow indicates the possible conformational transition of the *Ec*ZorD NTD. **k**, Zoom in from **j** highlighting that the NTD of *Ec*ZorD in the DNA free state clashes with DNA in the ZorD–DNA complex model. **l**, Superimposition of the AlphaFold3 predicted model ZorD–DNA complex with the top hit (PDB 7X3T[76]) from Dali (Z-score = 26.6). **m**, AlphaFold3 predicted ZorD–ZorC–dsDNA-ATP-Mg^{2+} complex, with a confidence-coloured (per-atom confidence estimate on a 0–100 scale) model shown in the right panel.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | ZorAB recruit ZorC and ZorD during phage invasion. **a**, Complementation experiment between *E. coli* and *P. aeruginosa* (*Pa*) Zorya I. Schematic representation of *Ec*ZorI, *Pa*ZorI and the constructs for *Pa*ZorCD or *Pa*ZorD complementation of *Ec*ZorI gene deletions. **b**, Anti-phage defence provided by the constructs in (**a**), as measured using EOP assays for phages Bas49, Bas52 and Bas57. Data represent the mean of at least 3 replicates (data points indicate replicates) and are normalized to the control samples lacking Zorya. **c**, Strategy of fusing mNeonGreen (mNG) or HaloTag (HT) or both into *Ec*ZorI operon. **d**, The effects of the mNeongreen (mNG) fusions to *Ec*ZorI components on anti-phage defence, as measured using EOP assays for phage Bas24. Data represent the mean of at least 3 replicates (data points indicate replicates) and are normalized to the control samples lacking *Ec*ZorI. The boxed constructs (ZorB C-terminal mNG fusion: ZorB-mNG; ZorB C-terminal HT fusion: ZorB-HT; ZorC N-terminal mNG fusion: mNG-ZorC; ZorD C-terminal mNG fusion: ZorD-mNG; Dual-tagged constructs, ZorB C-terminal HT fusion and ZorC N-terminal mNG fusion: ZorB-HT + mNG-ZorC; ZorB C-terminal HT fusion and ZorD C-terminal mNG fusion: ZorB-HT + ZorD- mNG) were used for subsequent microscopy experiments. **e**, Exemplary denoised TIRF and brightfield microscopy pictures of mNG expression driven by the *Ec*ZorI native promoter (p-mNG) either untreated or exposed to Bas24 at an MOI of 5 for 30 min. Scale bar 2 µm. **f**, Exemplary denoised TIRF microscopy pictures of ZorB C-terminal mNG fusion either untreated or exposed to Bas24 at an MOI of 5 for 30 min. Scale bar 2 µm. **g**, Comparison of detected maxima of the ZorAB complex foci between untreated or exposed to Bas24 at an MOI of 5 for 30 min (n cells > 250 from n = 3 replicates), p-value: 0.030. Means are derived from three independent biological replicates. **h**, Exemplary denoised TIRF microscopy pictures of ZorD-mNG either untreated or exposed to increasing Bas24 at MOIs of 1, 5, or 50 for 30 min. **i**, Statistical comparison of ZorD-mNG maxima between untreated and conditions stated in **h**, p-values: 0.9978., 0.0009, 0.0258 and <0.0001. Means and exemplarily images in **e** and **h** derive from at least three independent biological replicates. For **g** and **i** data are presented as mean values and Tukey whiskers. Scale bar 2 µm.

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

	*Ec*ZorAB WT (EMDB-18751) (PDB 8QYD)	*Ec*ZorA$^{E86A/E89A}$ZorB (EMD-18754) (PDB 8QYH)	*Ec*ZorA$^{\Delta 359-592}$ZorB (EMD-18756) (PDB 8QYK)	*Ec*ZorA$^{\Delta 435-729}$ZorB (EMD-18766) (PDB 8QYY)	*Ec*ZorD apo form (EMD-18747) (PDB 8QY7)	*Ec*ZorD + ATP-γ-S (EMD-18750) (PDB 8QYC)	*Ec*ZorC Apo from (EMD-18948) (PDB 8R68)
Data collection and processing							
Microscope	Titan Krios G2						
Magnification (nominal)	96,000x						
Voltage (kV)	300						
Total exposure ($e^-/Å^2$)	40.01	38.00	41.00	41.00	40.00	38.00	39.50
Exposure fractions (no.)	40						
Pixel size (Å)	0.832	0.832	0.832	0.832	0.832	0.832	0.832
Symmetry imposed	C1						
Movies used (no.)	7,700	3,203	3,181	1008	1,750	1,213	5,209
Final particles (no.)	227,728	97,005	278,186	62,926	332,426	211,477	168,569
Box size (pixels)	500	500	500	500	400	400	256
Map resolution (Å) (FSC 0.143)	2.67	2.40	2.07	2.56	2.66	2.75	3.55
Refinement							
Refinement resolution (FSC map vs. model (masked)=0.143) (Å)	2.60	2.40	2.10	2.50	2.60	2.70	3.50
Model composition							
Non-hydrogen atoms	16,154	15,925	16,182	14,829	7,168	7,199	3,240
Protein residues	1,892	1,837	1,817	1,717	893	893	391
Water molecules	313	484	866	313	n/a	n/a	n/a
B-factors (mean; $Å^2$)							
Protein	89.70	90.46	64.60	84.84	27.26	56.79	64.60
Solvent	88.61	50.89	34.01	44.93	n/a	n/a	n/a
R.m.s. deviations							
Bond lengths (Å)	0.006	0.008	0.008	0.009	0.018	0.002	0.003
Bond angles (°)	0.874	1.025	1.373	1.339	1.706	0.441	0.648
CC (mask)	0.86	0.85	0.85	0.84	0.85	0.83	0.52
Validation							
MolProbity score	1.46	1.46	1.42	1.45	1.90	1.47	2.11
Poor rotamers (%)	0.12	0.12	0.19	0.00	0.00	1.29	0.30
Ramachandran plot							
Favored (%)	98.30	98.63	98.56	98.12	96.72	97.40	94.43
Allowed (%)	1.7	1.37	1.44	1.88	3.28	2.60	5.57
Disallowed (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00

nature portfolio

Corresponding author(s): Nicholas M.I. Taylor

Last updated by author(s): Nov 19, 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	For cryo-EM: EPU; for TIRF microscopy: NIS-Elements 5.42.04
Data analysis	For cryo-EM: CryoSPARC 4.x, Other: Coot 0.9-pre, Phenix 1.13, Chimera X 1.8, Pymol 2.5.4 and 3.0.2, Bioinformatics Toolkit server (MUSCLE alignment), , MMseqs2, Python 3, ImageJ 1.53t, MicrobeJ 5.13j, GraphPad Prism 9.4.1 and GraphPad Prisim 10. Python code used to generate Extended Data Videos 1 - 6 and plot graphs is available at: https://github.com/SalmoLab/Zorya_Nature2024

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

Atomic coordinates for ZorAB WT, ZorA_E86A/E89A_ZorB, ZorA_delta_359-592_ZorB and ZorA_delta_435-729_ZorB were deposited in the Protein Data Bank (PDB) under accession codes 8QYD, 8QYH, 8QYK, 8QYY, respectively. The corresponding electrostatic potential maps were deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD-18751, EMD-18754, EMD-18756, EMD-18766, respectively. The local refinement map of ZorB PGBD in ZorAB WT were deposited in the EMDB under accession codes EMD-18752. Atomic coordinates for ZorC were deposited in the PDB under accession codes PDB: 8R68. The corresponding electrostatic potential maps was deposited in the EMDB under accession codes EMDB: EMD-18848. Atomic coordinates for ZorD apo form and its complex with ATP-□-S were deposited in the PDB under accession codes PDB: 8QY7 and 8QYC, respectively. The corresponding electrostatic potential maps were deposited in the EMDB under accession codes EMD-18747 and EMD-18750. Official validation reports from wwPDB for all macromolecular structures studied in the paper have been provided. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE67 partner repository with the dataset identifier PXD047450.

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	For cryo-EM: No sample size was calculated. Sample size was based on experience to have a number of micrographs/particles that, if possible, would result in a cryo-EM reconstruction that would allow atomic model construction. For microscopy experiments, no a priori sample sizes were calculated. Sample size were chosen according to our experience concerning quantitative single cell microscopy (e.g. 10.1038/s41467-024-50278-0;10.1073/pnas.2310842120). Exact sample size are provided in the figure legends.
Data exclusions	Junk particles were removed during cryo-EM data processing in CryoSPARC. No microscopy data was excluded from the analysis.
Replication	TIRF: Data derived from at least three independent clones of bacteria and two sets of phage isolations. All attempts at replication were successful. Allocating experimental groups was not relevant for this study, as all bacterial cells of a particular strain are genetic clones. We performed 3 replicates under similar conditions for both the ZorAB full length and ZorB PGBD fragment PG pull-down experiments. The experimental results were consistent. For ZorB PGBD fragment experiments, we selected the experiment with the lowest non-specific binding (as indicated by amount of ZorE pulled down
Randomization	No experimental groups were formed/compared. Data was collected randomly in each set of experiments.
Blinding	Image and data analysis was automated whenever possible. Blinding was neither possible nor necessary for this study, as all bacterial cells of a particular strain are genetic clones and analyses were not sufficiently subjective to require researcher blinding.

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

Antibodies

Antibodies used	Promega Anti-Halo Tag Monoclonal Antibody #G921A 1:1,000
Validation	Mouse monoclonal antibody was raised against the Halo Tag protein and validated using E.coli cell lysates expressing Halo Tag as positive control and E.coli lysate not expressing any HaloTag protein as negative control.

Plants

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