

A DNA-gated molecular guard controls bacterial Hailong anti-phage defence

<https://doi.org/10.1038/s41586-025-09058-z>

Received: 26 August 2024

Accepted: 23 April 2025

Published online: 30 April 2025

 Check for updates

Joel M. J. Tan^{1,2}, Sarah Melamed³, Joshua C. Cofsky⁴, Deepsing Syangtan¹, Samuel J. Hobbs^{1,2}, Josefina del Marmol^{4,5}, Marco Jost¹, Andrew C. Kruse⁴, Rotem Sorek³ & Philip J. Kranzusch^{1,2,6}✉

Animal and bacterial cells use nucleotidyltransferase (NTase) enzymes to respond to viral infection and control major forms of immune signalling including cGAS-STING innate immunity and CBASS anti-phage defence^{1–4}. Here we discover a family of bacterial defence systems, which we name Hailong, that use NTase enzymes to constitutively synthesize DNA signals and guard against phage infection. Hailong protein B (HalB) is an NTase that converts deoxy-ATP into single-stranded DNA oligomers. A series of X-ray crystal structures define a stepwise mechanism of HalB DNA synthesis initiated by a C-terminal tyrosine residue that enables de novo enzymatic priming. We show that HalB DNA signals bind to and repress activation of a partnering Hailong protein A (HalA) effector complex. A 2.0-Å cryo-electron microscopy structure of the HalA–DNA complex reveals a membrane protein with a conserved ion channel domain and a unique crown domain that binds the DNA signal and gates activation. Analysing Hailong defence in vivo, we demonstrate that viral DNA exonucleases required for phage replication trigger release of the primed HalA complex and induce protective host cell growth arrest. Our results explain how inhibitory nucleotide immune signals can serve as molecular guards against phage infection and expand the mechanisms NTase enzymes use to control antiviral immunity.

We analysed protein-coding genes occurring in bacterial defence islands and phage competition loci—regions in bacterial and prophage genomes known to encode defence systems^{5–12}—for putative nucleotidyltransferase (NTase) proteins and identified a cluster of conserved proteins predicted to contain an approximately 25-kDa pol-β superfamily NTase enzymatic domain. The NTase protein is encoded as part of a widespread system that we named ‘Hailong’ after a Chinese ocean dragon of protection (Fig. 1a–c and Supplementary Table 1). Hailong is a two-gene operon that encodes Hailong protein A (HalA), a transmembrane protein of unknown function, and Hailong protein B (HalB), the predicted NTase enzyme (Fig. 1b). Other example Hailong operons include a WYL domain-containing protein designated Hailong protein C that may regulate transcription¹³, and operons with further ancillary genes of unknown function designated Hailong protein D and Hailong protein E (Fig. 1a). Hailong operons are widely conserved in defence islands across more than 70 genera of gram-positive and gram-negative bacteria, and are also present in diverse prophage competition elements, suggesting a role in protecting cells against phage infection (Fig. 1d and Supplementary Table 2). To test this hypothesis, we cloned four divergent Hailong operons and assessed the ability of these systems to defend *Escherichia coli* against phage infection. Each Hailong system broadly defended against phage infection, with Hailong from the bacterium *E. coli* STEC 1178 providing greater than 10,000-fold protection against double-stranded DNA (dsDNA) phages SECφ4 and

SECφ6, and Hailong from *Rhodobacteraceae bacterium* QY30 providing greater than 1,000-fold protection against phages T4 and T6 (Fig. 1e,f). These results define Hailong as a widespread defence system in bacteria that protects cells from phage infection.

HalB synthesizes oligodeoxyadenylate

We observed Hailong encoded next to anti-phage defence systems that synthesize nucleotide immune signals including CBASS and Pycsar^{3,4,14–17} (Fig. 1a). We hypothesized that similar to CBASS and Pycsar, HalB NTase function in Hailong defence may be required to initiate nucleotide immune signalling as part of the response to phage infection. We introduced mutations to the NTase active site of HalB and surprisingly observed that expression of mutant Hailong operons induced potent cellular toxicity in *E. coli*, resulting in near complete inhibition of bacterial growth (Fig. 2a). Cellular toxicity was also observed when HalA was expressed alone in the absence of HalB, suggesting that HalB NTase function is required to negatively regulate HalA activity before phage infection (Fig. 2a).

To define the function of HalB NTase activity, we next determined a 1.9-Å crystal structure of HalB from the *R. bacterium* QY30 Hailong defence system (Supplementary Table 3). The structure of HalB reveals an N-terminal Rossmann-like fold with five β-strands that form a mixed parallel, anti-parallel β-sheet as the minimal NTase core, and a

¹Department of Microbiology, Harvard Medical School, Boston, MA, USA. ²Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA, USA. ³Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel. ⁴Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. ⁵Howard Hughes Medical Institute, Boston, MA, USA. ⁶Parker Institute for Cancer Immunotherapy at Dana-Farber Cancer Institute, Boston, MA, USA. ✉e-mail: philip_kranzusch@dfci.harvard.edu

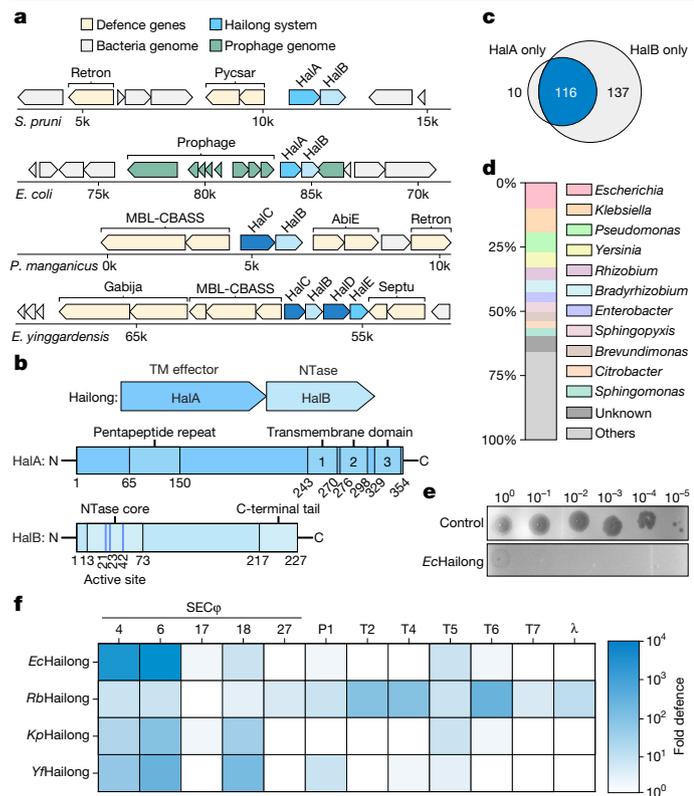


Fig. 1 | Diverse Hailong systems protect bacteria from phage infection. **a**, Representative Hailong operons and gene neighbourhoods from *Spingomonas pruni* NBRC 15498, *E. coli* 300073, *Pseudaminobacter manganicus* JH-7 and *Ensifer yinggardensis* WSM1721. Defence systems were annotated using Defence Finder¹⁰. **b**, Schematic of HalA and HalB proteins. **c**, Frequency of HalA and HalB found in all Hailong systems. **d**, Genera of bacteria encoding Hailong. **e**, Representative plaque assays of *E. coli* expressing empty vector (control) or Hailong from *E. coli* STEC 1178 and challenged with phage SECφ6. Serial dilution factors represent phage dilution to visualize plaque forming units (PFUs). **f**, Heatmap illustrating fold defence of *E. coli* expressing Hailong from gammaproteobacteria (*E. coli*, *Klebsiella*, *Yersinia*) or alphaproteobacteria (*Rhodobacteraceae*) species and challenged with indicated phages. Data shown are representative of three independent experiments. TM, transmembrane.

C-terminal helix bundle connected by a short linker region (Fig. 2b and Extended Data Fig. 1a,d–f). Structural homology analysis confirmed HalB as a pol-β superfamily NTase with related structures including kanamycin resistance and HEPN-MNT antitoxin enzymes that use ATP as a substrate to modify and inactivate small molecule and protein toxin targets^{12,18,19} (Extended Data Fig. 1b,c). The HalB structure reveals a tightly packed homotetrameric assembly. Extensive hydrophobic interactions along the back of the HalB protein and between the short linker regions that brace the NTase enzymatic core result in more than 1,690 Å² of buried surface area (Extended Data Fig. 1f,h). We confirmed using size exclusion chromatography with multi-angle static light scattering (SEC-MALS) that HalB is a tetramer in solution and observed that mutations to the tetramerization interface result in loss of function and Hailong operon toxicity in vivo (Extended Data Fig. 1f–k).

Surprisingly, we observed exceptionally strong electron density extending from each active site of the HalB tetrameric complex (Fig. 2b,c). Analysis of a Polder omit map allowed unambiguous assignment of this electron density as single-stranded nucleic acid chains formed by deoxyadenosine residues (Fig. 2c). We analysed the protein by denaturing SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and observed that purified HalB migrates as a ladder of high-order species, suggesting that the nucleic acid molecules are

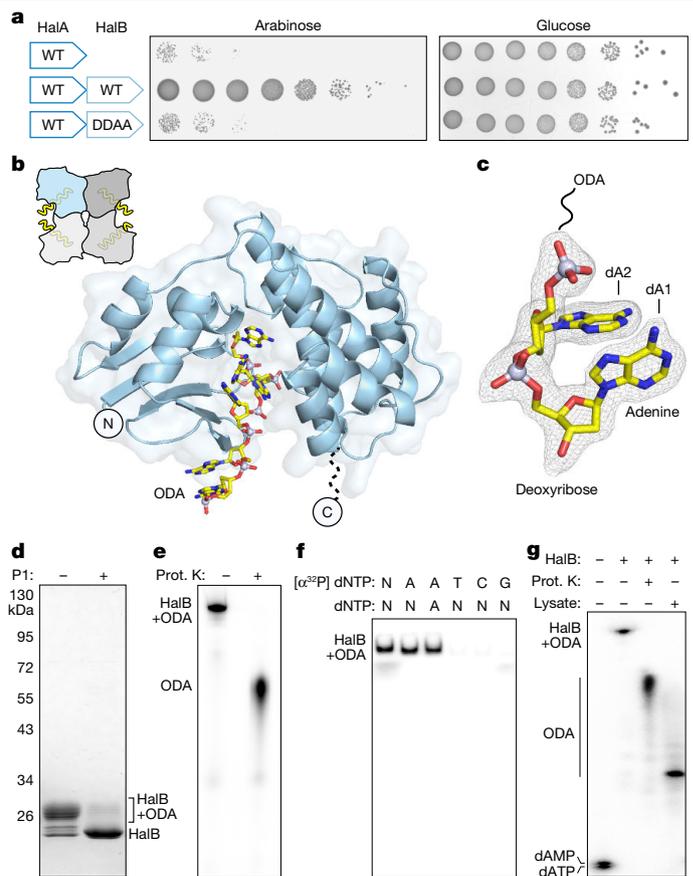


Fig. 2 | HalB synthesizes ODA. **a**, Bacterial growth assay of a tenfold dilution series of *E. coli* containing arabinose-inducible plasmids expressing HalA alone, HalA with HalB, or HalA with a HalB catalytically inactive mutant (DDAA, D21A/D23A). **b**, Crystal structure of HalB in complex with the nucleotide product ODA. Inset depicts cartoon representation of tetrameric structure. **c**, Polder omit map contoured at 3.4 σ of the first two deoxyadenosine molecules of ODA. **d**, Coomassie-stained SDS–PAGE analysis of purified HalB with and without nuclease P1 treatment. **e**, Urea–PAGE analysis of HalB ODA synthesis reactions labelled with α³²P-dATP and treated with and without proteinase K. **f**, Urea–PAGE analysis of HalB substrate specificity using dNTPs and α³²P-labelled dNTPs as indicated. **g**, Urea–PAGE analysis of ODA cleavage from the HalB–ODA complex using α³²P-labelled dATP treated with and without proteinase K or *E. coli* soluble lysate fraction. For further controls, see Extended Data Fig. 2d. Expression of HalB used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments. For gel source data, see Supplementary Fig. 1. Prot. K, proteinase K; WT, wild type.

covalently attached to the HalB protein (Fig. 2d). We treated HalB with nuclease P1 and confirmed that the ladder species collapsed to a single band corresponding to the molecular weight of the unmodified enzyme (Fig. 2d). The unexpected presence of nucleic acid attached to HalB suggested that the function of HalB NTase activity is to constitutively synthesize the product oligodeoxyadenylate (ODA). To test this hypothesis, we incubated HalB with α³²P-radiolabelled dNTP or NTP substrate nucleotides and analysed the reaction products by denaturing urea–polyacrylamide gel electrophoresis (urea–PAGE). HalB incorporates α³²P-dATP as part of a covalently attached product resulting in migration of a HalB–ODA complex at the top of the gel (Fig. 2e). Following ODA synthesis, degradation of HalB with proteinase K released the ODA product as a single-stranded DNA (ssDNA) species that migrated into the gel (Fig. 2e). HalB activity is metal-dependent and specifically requires dATP, with only weak residual activity observed using alternative nucleotide substrates (Fig. 2f and Extended Data Figs. 1l–n and 2a). We digested ODA with nuclease P1 and used liquid

chromatography–mass spectrometry (LC–MS) to confirm that ODA is composed specifically of deoxyadenosine residues (Extended Data Fig. 2b,c). Together, these structural and biochemical results demonstrate that HalB functions as a constitutively active NTase that converts dATP into ODA nucleic acid polymers.

HalB is a processive enzyme capable of synthesizing long uninterrupted chains of ODA in vitro (Extended Data Fig. 1m,n). We observed that treatment of the HalB–ODA complex with *E. coli* cell lysate resulted in release of ODA fragments (Fig. 2g and Extended Data Fig. 2d), suggesting that extensive polymers of ODA are unlikely to exist in vivo and that the HalB–ODA complex is probably processed by cellular housekeeping enzymes into shorter ODA species. To begin to define host factors involved in this process, we used the release of radiolabelled ODA as an assay to enable activity-guided fractionation (Extended Data Fig. 2e,f). Enrichment and mass spectrometry analysis of *E. coli* lysates identified two protein candidates, H-NS and StpA, known to interact with DNA, and we observed that purified H-NS and StpA each enhance reconstitution of ODA release from HalB in vitro (Extended Data Fig. 2g–j). These results support a potential role for host factors in HalB–ODA processing during ODA synthesis in vivo, although the mechanism of ODA release and specific length of ODA products remain important questions for future studies.

Mechanism of HalB ODA synthesis

Our results suggest that HalB catalyses de novo synthesis of ODA in the absence of a template or initiating nucleic acid primer. To define the mechanism of HalB ODA synthesis, we disrupted HalB constitutive NTase activity with mutations to the metal-coordinating residues D21 and D23 and determined a 2.0-Å crystal structure of the mutant enzyme (Supplementary Table 3). As expected, no ODA products are observed in the HalB D21A/D23A mutant enzyme structure (Fig. 3a and Extended Data Fig. 2k–n). In the mutant enzyme structure, we instead observed that residues 218–227 from the HalB C-terminal tail reach into the NTase core and position residue Y227 within the active site in a conformation poised to serve as an initiating priming hydroxyl for ODA synthesis (Fig. 3a–c). Sequence analysis of diverse HalB proteins in Hailong defence systems confirmed strict conservation of a C-terminal tyrosine, further supporting a direct role for this residue in priming nucleotide incorporation (Extended Data Fig. 1a). Comparison of the HalB wild-type and D21A/D23A structures reveals an overlapping network of residues including R13, R122 and R128 that explain how the HalB NTase domain coordinates both initial placement of the C-terminal protein primer and subsequent stabilization of the elongating ODA product (Fig. 3a–d).

We next used structural analysis of HalB active site mutant R164A (2.4 Å) combined with a dAp(c)pp (non-hydrolysable dATP) substrate analogue to define the stepwise mechanism of ODA synthesis (Extended Data Fig. 2o,p and Supplementary Table 3). HalB ODA synthesis initiates with positioning of the C-terminal protein primer in the NTase active site (Fig. 3e). Coordination between conserved residues R13/R128 of the NTase core and residues E224/Y227 carboxy-terminus positions Y227 for nucleophilic attack (Fig. 3c–e). The tyrosine side-chain hydroxyl attacks the α -phosphate of the first incoming molecule of dATP, with conserved residues H64 and R164 from the NTase core providing extra stabilizing stacking and hydrogen-bond interactions with the ribose and nucleotide phosphates, respectively, along with a sequence-specific contact between residue T129 and the adenine nucleobase (Fig. 3e and Extended Data Fig. 2q,r). The reaction results in transfer of deoxyadenosine monophosphate and formation of a covalent HalB–dAMP attachment. The HalB Y227–dAMP complex is then coordinated with new stacking and hydrogen-bond interactions between conserved residues F61 and R128 with the nucleobase and nucleotide phosphates, respectively (Fig. 3e and Extended Data Fig. 2r), and the dAMP 3' hydroxyl repositions to attack the next incoming

molecule of dATP. Subsequent repeated steps of elongation displace the C-terminal priming peptide out of the active site pocket and allow continued synthesis of the growing ODA chain (Fig. 3e). Deletion of the HalB C-terminal protein primer or mutation of the initiating Y227 residue prevents formation of the covalent HalB–ODA complex and eliminates ODA synthesis activity (Fig. 3f,g). The key HalB coordinating residues R13, R128 and R164 are each also required for efficient ODA synthesis. (Fig. 3f,g). We confirmed that HalB mutations that disrupt ODA synthesis in vitro cause loss of function and induce Hailong operon toxicity in vivo (Fig. 3h and Extended Data Fig. 1k). These findings define how HalB uses a unique mechanism of protein-based priming to catalyse ODA synthesis.

HalA is a DNA-gated ion channel

We next sought to define the function of the Hailong membrane protein HalA and understand how HalB ODA signalling regulates anti-phage defence. HalA induces potent cellular toxicity in the absence of active HalB (Fig. 2a). We did not observe protein–protein interactions between HalB and HalA during purification of HalA (Extended Data Fig. 3a,b), and we therefore hypothesized that the HalB nucleotide product ODA directly binds HalA to regulate protein function. We co-expressed *R. bacterium* HalA in the presence of HalB, and purified HalA as a detergent-solubilized complex. Using this purified HalA species, we confirmed direct high-affinity recognition of ODA (K_d ~ 65 nM) (Fig. 4a and Extended Data Fig. 3a,b), although we note that this measurement probably underestimates the HalA–ODA complex dissociation constant owing to potential competition with ODA molecules already bound to HalA during expression in the presence of HalB. HalA–ODA complex formation is length- and nucleobase-specific, with shorter 5-mer ODA molecules exhibiting a more than tenfold higher affinity than longer ODA chains (Fig. 4a,b and Extended Data Fig. 3c), whereas no interaction was observed between HalA and other ssDNA or single-stranded RNA ligands (Fig. 4c). Together, these results highlight the strict specificity of HalA for recognition of the HalB ODA product.

We used cryo-electron microscopy (cryo-EM) to determine a 2.0-Å structure of the HalA–ODA complex (Fig. 4d, Extended Data Fig. 4 and Supplementary Table 3). The structure reveals a homotetrameric assembly with a C-terminal transmembrane domain and long central channel braced by a cytosolic domain resembling a crown (Fig. 4d and Extended Data Fig. 5a). The HalA crown domain is formed by N-terminal pentapeptide repeats from residues 68–153 that fold into a four-sided, right-handed β -helix (Fig. 4d and Extended Data Figs. 5a and 6a). HalA central α -helices $\alpha 4$ and $\alpha 5$ connect the crown to transmembrane domain helices $\alpha 6$ – $\alpha 10$ in the C-terminal domain (Fig. 4d and Extended Data Fig. 5a). HalA tetramerization is stabilized by conserved hydrophobic residues within the oligomerization interface along neighbouring helices $\alpha 8$ and $\alpha 10$ (Fig. 4d and Extended Data Fig. 5f). We observed an ODA molecule bound at each inter-subunit interface within the HalA crown domain resulting in a 4:4 HalA–ODA complex (Fig. 4d,e). In the HalA ODA binding pocket, highly conserved residues form extensive interactions with each ODA nucleobase and the DNA phosphate backbone (Fig. 4e,f). HalA residues Q178 (deoxyadenosine 1, dA1), W27 (dA2), E28/S141/N143 (dA3), E144 (dA4) and T83/T104/T124 (dA5) coordinate a network of adenine-specific interactions, and residues K24, K91, K109, R131 and R181 form hydrogen-bond interactions with the ODA phosphate backbone (Fig. 4e,f and Extended Data Fig. 5h). Predicted steric clashes between ODA and HalA side-chains, and the 2' OH position of RNA nucleotides, further explain selectivity of HalA for the DNA signal ODA (Fig. 4e and Extended Data Fig. 5h).

Structural analyses suggest that the HalA C-terminal domain is homologous to conserved ion channel proteins. Notably, HalA shares all major structural features of related well-characterized ion channel proteins, including the sodium channel Nav1.8, potassium channel Kir4.1 and non-selective cation channel NaK^{20–22} (Extended Data Fig. 5b–e).

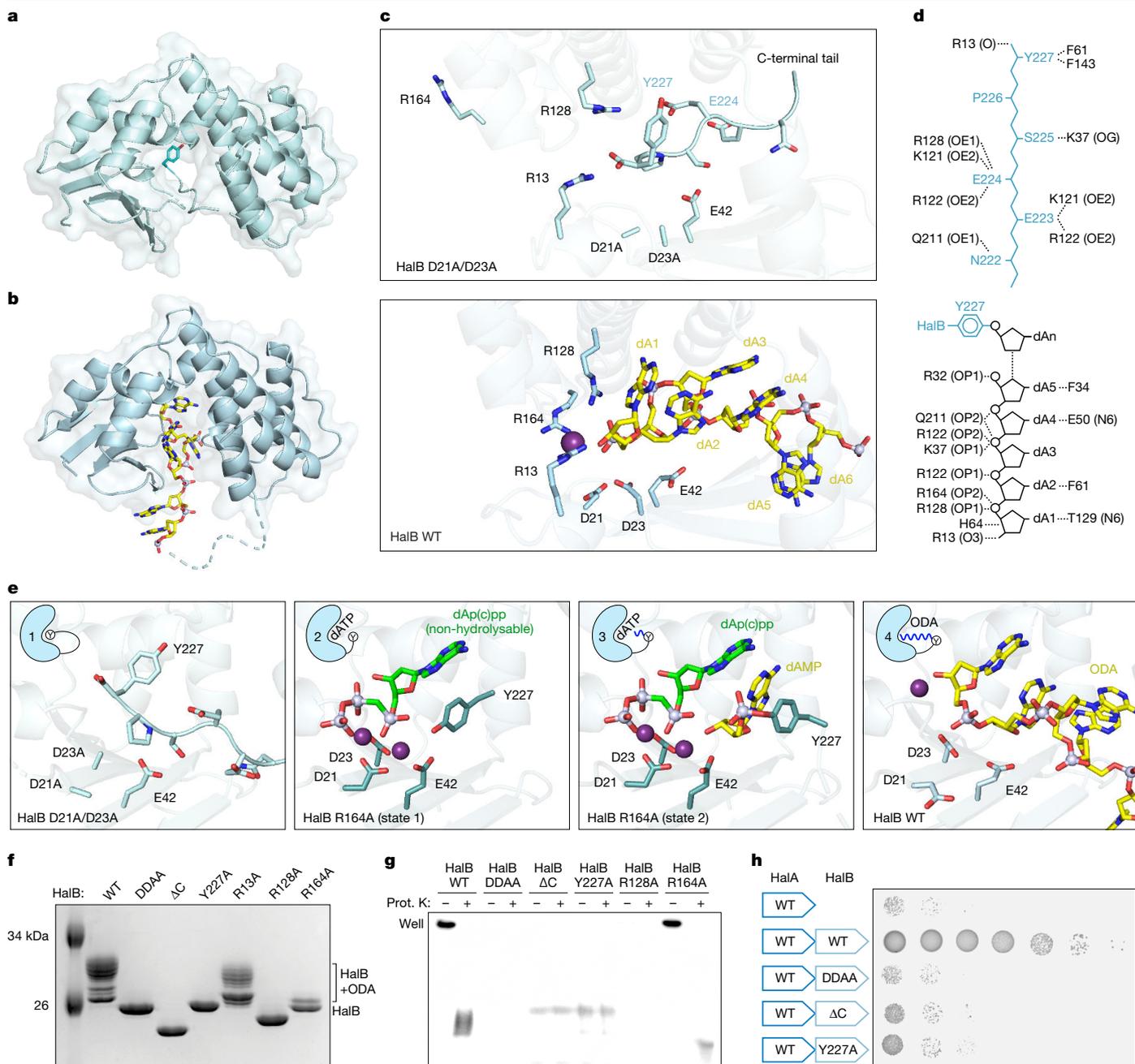


Fig. 3 | Mechanism of HalB ODA synthesis. **a**, Crystal structure of catalytically inactive HalB from *R. bacterium* QY30 revealing re-positioning of the flexible C-terminal tail within the NTase active site. **b**, Structure of HalB as in Fig. 2b with a dashed line indicating the projected position of the covalent attachment of ODA to the protein flexible C-terminal tail. **c**, Top, overview of HalB D21A/D23A active site and the C-terminal tail. Bottom, overview of WT HalB active site and a detailed view of residues interacting with ODA. **d**, Top, schematic of protein-protein contacts with HalB and its C-terminal tail. Bottom, schematic of protein-DNA contacts in the HalB-ODA complex. HalB residues are highlighted in blue. **e**, Schematic of the step-by-step mechanism of ODA synthesis: Step 1, a tyrosine in the C-terminal tail acts as a protein primer within the active site. Step 2, dATP binding is coordinated by metal ions and the tyrosine repositions to attack the

α -phosphate. Step 3, tyrosine forms a covalent bond with dAMP and the protein-DNA complex repositions for the next attack on dATP. Step 4, fully synthesized ODA. **f**, Coomassie-stained SDS-PAGE analysis of purified HalB mutants demonstrating the NTase active site, C-terminal tail and tyrosine priming residue are required for ODA synthesis and covalent attachment. **g**, Urea-PAGE analysis of mutant HalB ODA synthesis reactions labelled with $\alpha^{32}\text{P}$ -dATP and treated with and without proteinase K. **h**, Bacterial growth assay of *E. coli* expressing either HalA alone, or HalA with HalB mutants. Expression of HalB used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments. Manganese ions are depicted as purple spheres. For gel source data, see Supplementary Fig. 1.

In the HalA-ODA cryo-EM structure, ODA binding bridges inter-subunit interactions in the crown domain and compresses the channel into a state matching the closed conformation of the potassium channel MthK²³. AlphaFold modelling of HalA in the absence of ODA reveals a more open conformation of the ion conduction pathway, suggesting

that release of inter-subunit interactions in the crown domain allows channel opening (Fig. 4g, Extended Data Fig. 5g and Supplementary Video 1). These results support that ODA binding negatively regulates HalA channel function. Consistent with this model, we observed that mutations to residues K91E and K109E in the HalA ODA binding interface

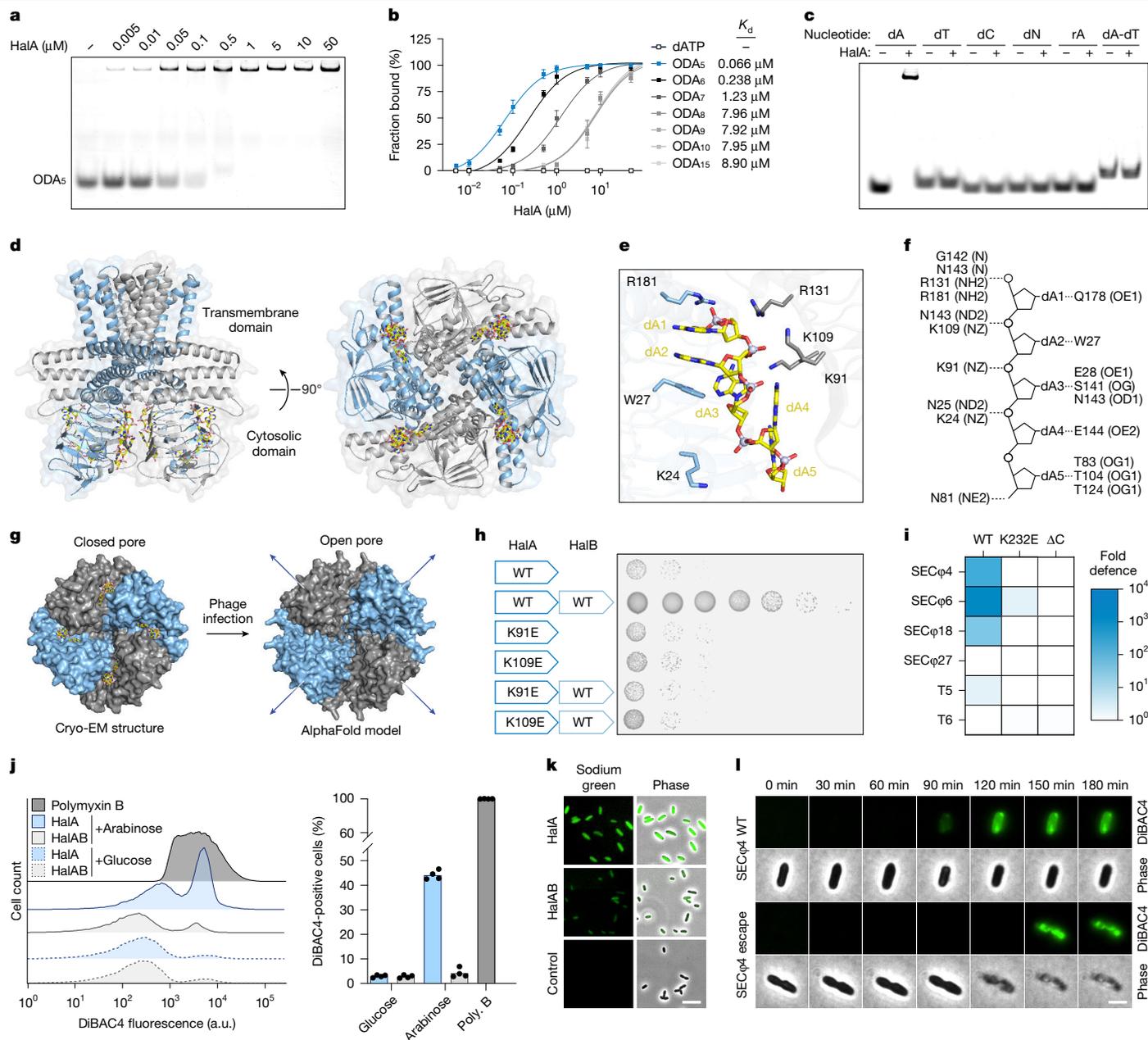


Fig. 4 | Structure and function of HalA. **a**, Electrophoretic mobility shift assay of ODA binding to HalA–ODA complex. **b**, Quantification of binding affinity of HalA–ODA complex with different lengths of ODA. Data are presented as mean \pm s.d. from $n = 3$ independent experiments. **c**, Electrophoretic mobility shift assay analysis of HalA–ODA complex binding to different 5-nucleotide (nt) single-stranded deoxy-nucleotide substrates (dA, dT, dC, dN), 5-nt single-stranded RNA (rA) or 20-base pair dsDNA (dA–dT). **d**, Cryo-EM structure of HalA from *R. bacterium* QY30 in complex with ODA. **e**, Magnified view of the ODA binding pocket at the interface between two HalA pentapeptide repeat domains and residues interacting with ODA. **f**, Schematic of protein–DNA contacts in the HalA–ODA complex. **g**, Structural model of HalA activation. Cryo-EM structure of ODA-bound HalA in a closed conformation and an AlphaFold modelled apo structure of HalA showing a wider diameter along the channel, suggesting a more open conformation. **h**, Bacterial growth assay of *E. coli* expressing HalA mutants with and without HalB. **i**, Heatmap illustrating fold defence of *E. coli* expressing *E. coli* STEC 1178 Hailong operons with HalA channel assembly mutants and challenged with indicated phages. K232E and ΔC mutants used for the phage

challenge assay are equivalent to *R. bacterium* QY30 K211E mutant and truncation of residues 352–359, respectively. **j**, Kernel density estimates of fluorescence distributions measured by flow cytometry (left) and quantification of membrane depolarization (right) of *E. coli* cells containing arabinose-inducible plasmids expressing HalA alone or HalAB, treated with glucose or arabinose for 8 h and incubated with the voltage-sensitive dye DiBAC4. Bar graph data are presented as means from four independent experiments. Polymyxin B was used as a positive control for membrane depolarization. **k**, Fluorescence microscopy analysis of ion flow using cell permeant fluorescent Na⁺ indicator Sodium Green Tetraacetate in *E. coli* containing arabinose-inducible plasmids expressing HalA alone or HalAB. Scale bar, 5 μm . **l**, Live cell imaging analysis of membrane depolarization using DiBAC4 in *E. coli* containing plasmids expressing Hailong defence system and infected with WT or escape mutant phage SEC ϕ 4. Scale bar, 1 μm . Hailong systems used in this figure were from *R. bacterium* QY30 (**a–h, j, k**) or *E. coli* STEC 1178 (**i, l**). Data shown are representative of at least three independent experiments. HalAB, HalA expressed with HalB.

result in constitutive Hailong operon toxicity in vivo irrespective of the presence of functional HalB (Fig. 4h). Furthermore, we observed that mutations to the HalA oligomerization interface that are expected to

inhibit channel assembly disrupt Hailong anti-phage defence (Fig. 4i and Extended Data Figs. 5f and 6a,b). Together, these data demonstrate that HalA probably functions as a DNA-gated ion channel that uses

release of ODA binding as a regulatory switch to control anti-phage defence.

To confirm the role of HalA as an effector protein that functions as an ion channel, we first analysed HalA activity *in vivo* using flow cytometry and DiBAC4, a voltage-sensitive dye that fluoresces upon entering bacterial cells that have lost membrane potential²⁴. HalA expression in *E. coli* results in a strong increase in DiBAC4 fluorescence consistent with membrane depolarization, whereas HalA co-expressed with HalB had no effect (Fig. 4j and Extended Data Fig. 7a,c). Cells expressing active HalA do not take up the membrane-impermeable dye propidium iodide, confirming that HalA induces membrane depolarization and not large-scale loss of membrane integrity or cell death²⁵ (Extended Data Fig. 7b,c). Using a cell permeant fluorescent Na⁺ indicator, we further observed that expression of HalA alone, but not HalA in the presence of HalB, induces increased fluorescence and cell swelling, supporting that HalA may mediate influx of sodium ions (Fig. 4k), although definitive confirmation of HalA function and identification of ion channel specificity will require future electrophysiology analysis. Finally, we monitored the *in vivo* impact of HalA activation on membrane depolarization in real time using live cell imaging of *E. coli* expressing Hailong anti-phage defence and infected with phage SECφ4. Upon phage infection, activation of Hailong induces initial membrane shrinkage followed by membrane swelling, increase in DiBAC4 fluorescence and cell growth arrest, consistent with depolarization of the cellular membrane^{24,26,27} (Fig. 4l and Extended Data Fig. 7d–f). Together, these results support that HalA activation causes membrane depolarization and influx of ions without disrupting overall membrane integrity, consistent with a role of HalA as an ion channel effector.

Phage DNA exonuclease activates Hailong

We reasoned that phage infection activates Hailong anti-phage defence by specifically disrupting ODA signalling. To determine the molecular cue responsible for defence activation *in vivo*, we infected cells encoding the *E. coli* STEC 1178 Hailong operon with phage SECφ4 and isolated escape mutants. In the presence of Hailong anti-phage defence, phage SECφ4 escape mutants no longer induce the membrane depolarization event observed with wild-type phage (Fig. 4l and Extended Data Fig. 7d,e). Instead, phage SECφ4 escape mutants cause cell lysis and replicate to high titres exhibiting more than 1,000-fold recovery and clear resistance to host defence (Figs. 4l and 5a and Extended Data Fig. 8a). Whole-genome sequencing of six independent SECφ4 escape phages revealed that all resistant phages contained missense mutations in a single uncharacterized phage gene, *gp43*, predicted to encode an exonuclease (InterPro IPR024432) (Fig. 5b and Extended Data Fig. 8a,b). AlphaFold modelling of SECφ4 *gp43* supports homology to many known ssDNA exonuclease proteins, including human EXO5 (ref. 28), and allowed identification of a conserved *gp43* E132, D159, K179 active site motif typical of PD(D/E)XK-like exonucleases that catalyse DNA degradation (Extended Data Fig. 8c,d). Gp43 is highly conserved in diverse phages and is frequently encoded adjacent to core viral DNA replication machinery, including phage DNA polymerase, primase, resolvase and ssDNA binding proteins (Fig. 5b, Extended Data Fig. 9a,b and Supplementary Table 4). No SECφ4 escape phages encoded nonsense or frameshift mutations in *gp43*, supporting that this exonuclease is probably required for efficient phage replication.

We hypothesized that SECφ4 *gp43* exonuclease activity may degrade ODA and directly trigger Hailong anti-phage defence. We purified SECφ4 *gp43* and confirmed that the protein is a promiscuous 5'→3' DNA exonuclease capable of degrading the HalB signal ODA as well as other ssDNA and dsDNA substrates (Fig. 5c and Extended Data Fig. 9c–e). Analysis of the AlphaFold model of SECφ4 *gp43* reveals that the escape mutations localize within predicted scaffolding regions of the core exonuclease domain, suggesting an indirect attenuation of enzyme efficiency or processivity (Extended Data Fig. 9f). We purified SECφ4

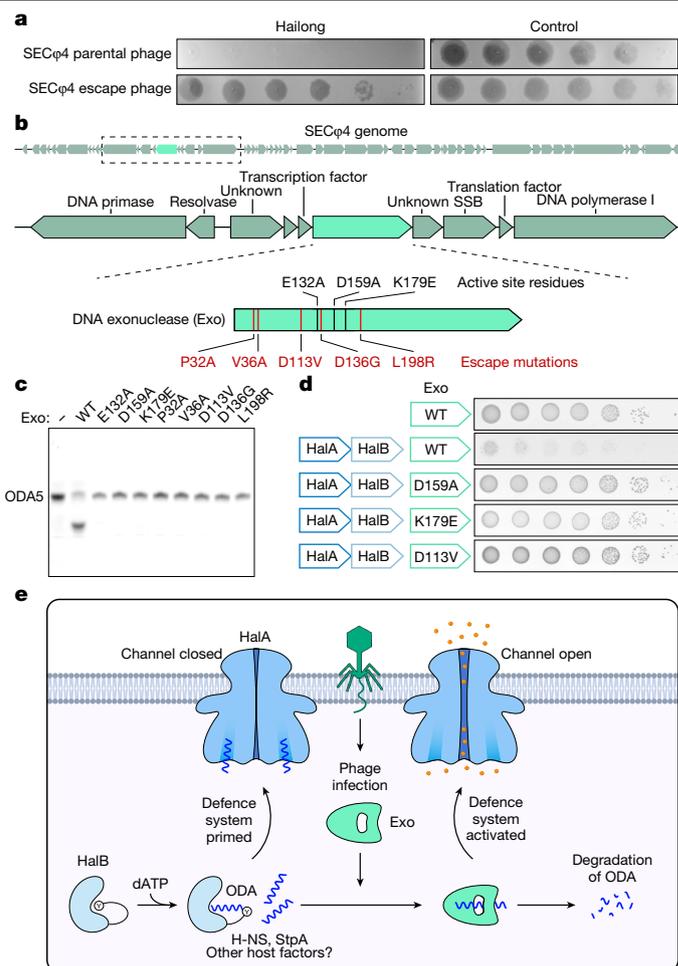


Fig. 5 | Phage DNA exonuclease activates Hailong. **a**, Representative plaque assays of *E. coli* containing empty vector (control) or Hailong from *E. coli* STEC 1178 plasmids and challenged with WT SECφ4 phage and SECφ4 escape mutant phage in tenfold serial dilution. **b**, Schematic depicting the genomic loci of the SECφ4 exonuclease *gp43* (Exo) and annotation of *gp43* escape mutations conferring resistance to Hailong anti-phage defence. **c**, DNA cleavage assay of ODA with purified WT and mutant Exo from phage SECφ4. **d**, Bacterial growth assay of *E. coli* transformed with two separate plasmids expressing Hailong from *R. bacterium* QY30 and SECφ4 WT or mutant Exo. **e**, Model of Hailong anti-phage defence system signalling and phage restriction. All data shown are representative of at least three independent experiments. SSB, ssDNA binding protein.

gp43 variants with each Hailong escape mutation and observed that these mutant proteins lost the ability to degrade the HalB signal ODA *in vitro* (Fig. 5c). Similarly, expression of the SECφ4 *gp43* wild-type protein, but not the escape mutant variants, was sufficient to induce Hailong toxicity *in vivo* in the absence of phage infection (Fig. 5d). Together, these results demonstrate that viral DNA exonuclease activity is a molecular cue that disrupts ODA signalling and activates Hailong anti-phage defence.

Our study defines Hailong as an anti-phage defence system that creates a DNA signal as a molecular guard to control antiviral immunity. In Hailong defence, the NTase enzyme HalB uses a self-priming mechanism to synthesize the ssDNA signal ODA which then functions as both a sensor for phage infection and a regulator for the cell growth arrest effector HalA (Fig. 5e). Hailong joins a prevalent set of antiviral defence systems in bacteria and animal cells that rely on specialized enzymes that synthesize nucleotide signals to control immune signalling^{1–4,17,29–36}. However, Hailong marks a departure from the canonical role of nucleotide signalling enzymes in bacterial and animal immunity

that are dormant in healthy cells and begin signal synthesis only following onset of viral infection. Rather than generating a nucleotide second messenger as a mechanism for signal amplification following infection, the Hailong enzyme HalB pre-emptively synthesizes ODA to function as a DNA molecular tripwire that guards the cell against phage replication. Likewise, the Hailong membrane effector protein HalA is distinct from conventional ligand-gated membrane channels that respond to activation by agonists such as cyclic nucleotides and neurotransmitters^{37–40}, and instead recognizes the ODA nucleotide signal as a reverse agonist that locks HalA in a primed, closed-channel state. In the absence of phage infection, HalB constitutively synthesizes ODA which is further processed by host factors such as H-NS and StpA to create an inhibitory nucleotide immune signal that is free to bind to HalA and inactivate effector function. Formation of a primed HalA–ODA complex before phage infection allows the host cell to rapidly respond to viral DNA replication enzymes that degrade ODA and trigger release of HalA ion channel effector function (Fig. 5e). We show that Hailong defence can specifically respond to viral DNA exonuclease activity to detect phage infection, but future studies will be required to explain a complete mechanism of ODA transfer from HalB to the effector protein HalA and how phage nucleases disrupt this process. Nucleotide signals such as ODA that function as guard molecules in host cells may also be capable of sensing cellular changes associated with phage infection, including expression of foreign nucleic acid binding proteins or enzymes that alter the nucleotide pool^{41,42}. Overall, our results define the mechanism of Hailong anti-phage defence and explain how host cells can use inhibitory nucleotide immune signals to guard against viral infection.

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/s41586-025-09058-z>.

- Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 (2013).
- Li, Y. et al. cGRLs are a diverse family of pattern recognition receptors in innate immunity. *Cell* **186**, 3261–3276 e3220 (2023).
- Whiteley, A. T. et al. Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194–199 (2019).
- Cohen, D. et al. Cyclic GMP-AMP signalling protects bacteria against viral infection. *Nature* **574**, 691–695 (2019).
- Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
- Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
- Rousset, F. et al. Phages and their satellites encode hotspots of antiviral systems. *Cell Host Microbe* **30**, 740–753 e745 (2022).
- Millman, A. et al. An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556–1569 e1555 (2022).
- 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).
- Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
- Iyer, L. M., Burroughs, A. M., Anand, S., de Souza, R. F. & Aravind, L. Polyvalent proteins, a pervasive theme in the intergenomic biological conflicts of bacteriophages and conjugative elements. *J. Bacteriol.* **199**, e00245-17 (2017).
- Kuchta, K., Knizewski, L., Wyrwicz, L. S., Rychlewski, L. & Ginalski, K. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic Acids Res.* **37**, 7701–7714 (2009).
- Blankenchip, C. L. et al. Control of bacterial immune signaling by a WYL domain transcription factor. *Nucleic Acids Res.* **50**, 5239–5250 (2022).
- Lowey, B. et al. CBASS immunity uses CARF-related effectors to sense 3'-5'- and 2'-5'-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* **182**, 38–49 e17 (2020).
- Morehouse, B. R. et al. STING cyclic dinucleotide sensing originated in bacteria. *Nature* **586**, 429–433 (2020).
- Ledvina, H. E. et al. An E1-E2 fusion protein primes antiviral immune signalling in bacteria. *Nature* **616**, 319–325 (2023).
- Tal, N. et al. Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell* **184**, 5728–5739 e5716 (2021).
- Pedersen, L. C., Benning, M. M. & Holden, H. M. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* **34**, 13305–13311 (1995).
- Songailiene, I. et al. HEPN-MNT toxin-antitoxin system: the HEPN ribonuclease is neutralized by oligoAMPylation. *Mol. Cell* **80**, 955–970 e957 (2020).
- Shi, N., Ye, S., Alam, A., Chen, L. & Jiang, Y. Atomic structure of a Na⁺- and K⁺-conducting channel. *Nature* **440**, 570–574 (2006).
- Zhou, X. et al. Pharmacological inhibition of Kir4.1 evokes rapid-onset antidepressant responses. *Nat. Chem. Biol.* **20**, 857–866 (2024).
- Huang, X. et al. Structural basis for high-voltage activation and subtype-specific inhibition of human Na_v1.8. *Proc. Natl Acad. Sci. USA* **119**, e2208211119 (2022).
- Fan, C. et al. Ball-and-chain inactivation in a calcium-gated potassium channel. *Nature* **580**, 288–293 (2020).
- VanderWal, A. R. et al. Csx28 is a membrane pore that enhances CRISPR-Cas13b-dependent antiphage defense. *Science* **380**, 410–415 (2023).
- Wein, T. et al. CARD domains mediate anti-phage defence in bacterial gasdermin systems. *Nature* **639**, 727–734 (2025).
- Mariano, G. et al. A family of Type VI secretion system effector proteins that form ion-selective pores. *Nat. Commun.* **10**, 5484 (2019).
- Cushnie, T. P., O'Driscoll, N. H. & Lamb, A. J. Morphological and ultrastructural changes in bacterial cells as an indicator of antibacterial mechanism of action. *Cell. Mol. Life Sci.* **73**, 4471–4492 (2016).
- Hambarde, S. et al. EXO5-DNA structure and BLM interactions direct DNA resection critical for ATR-dependent replication restart. *Mol. Cell* **81**, 2989–3006 e2989 (2021).
- Niewoehner, O. et al. Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543–548 (2017).
- Kazlauskienė, M., Kostiuk, G., Venclovas, C., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605–609 (2017).
- Chi, H. et al. Antiviral type III CRISPR signalling via conjugation of ATP and SAM. *Nature* **622**, 826–833 (2023).
- Ofir, G. et al. Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature* **600**, 116–120 (2021).
- Leavitt, A. et al. Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. *Nature* **611**, 326–331 (2022).
- Manik, M. K. et al. Cyclic ADP ribose isomers: production, chemical structures, and immune signaling. *Science* **377**, eadc8969 (2022).
- Kristiansen, H., Gad, H. H., Eskildsen-Larsen, S., Despres, P. & Hartmann, R. The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J. Interferon Cytokine Res.* **31**, 41–47 (2011).
- Hobbs, S. J. & Kranzusch, P. J. Nucleotide immune signaling in CBASS, Pycsar, Thoeris, and CRISPR antiphage defense. *Annu. Rev. Microbiol.* **78**, 255–276 (2024).
- Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. *Nature* **313**, 310–313 (1985).
- Twomey, E. C., Yelshanskaya, M. V., Grassucci, R. A., Frank, J. & Sobolevsky, A. I. Channel opening and gating mechanism in AMPA-subtype glutamate receptors. *Nature* **549**, 60–65 (2017).
- Gruschow, S. et al. CRISPR antiphage defence mediated by the cyclic nucleotide-binding membrane protein Csx23. *Nucleic Acids Res.* **52**, 2761–2775 (2024).
- Tak, U., Walth, P. & Whiteley, A. T. Bacterial cGAS-like enzymes produce 2',3'-cGAMP to activate an ion channel that restricts phage replication. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.07.24.550367> (2023).
- Stokar-Avihail, A. et al. Discovery of phage determinants that confer sensitivity to bacterial immune systems. *Cell* **186**, 1863–1876 e1816 (2023).
- Tal, N. et al. Bacteria deplete deoxynucleotides to defend against bacteriophage infection. *Nat. Microbiol.* **7**, 1200–1209 (2022).

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

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

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Methods

Bacterial strains and phages

E. coli Top10 (Thermo Fisher), BL21 (DE3) RIL (Agilent), BL21 LOBSTR (Kerafast) and MG1655 (ATCC 47076) strains were used in this study. Whenever applicable, media were supplemented with ampicillin ($100 \mu\text{g ml}^{-1}$) and/or chloramphenicol ($34 \mu\text{g ml}^{-1}$) to ensure the maintenance of plasmids. Phage strains were isolated as previously described⁵.

Protein expression and purification

HalA, HalB, HalB mutants, and SEC ϕ 4 gp43 DNA exonuclease (Exo) wild type and mutants were purified from *E. coli* using standard protocols as previously described^{43,44}. HalB genes were cloned from synthetic DNA fragments (Integrated DNA Technologies) into custom pET expression vectors containing amino-terminal 6 \times His-SUMO2 (pETS2) tags by Gibson assembly using HiFi DNA Assembly Master Mix (NEB). HalB plasmids were subsequently subcloned into a pETS2 vector containing 5 \times or 1 \times glycine-serine linkers using site-directed mutagenesis to ensure efficient cleavage of the SUMO2 tag. A pET expression vector containing an N-terminal 6 \times His tag encoding HalA co-expressed with untagged HalB was obtained from Twist Bioscience. A pETS2 expressing SEC ϕ 4 Exo was also obtained from Twist Bioscience.

For HalB, HalB mutant, and SEC ϕ 4 Exo wild-type and mutant protein purifications, expression plasmids were transformed into BL21(DE3) RIL cells and plated onto MDG media (1.5% Bacto agar, 0.5% glucose, 25 mM Na₂HPO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 0.25% aspartic acid, 2–50 μM trace metals, 100 $\mu\text{g ml}^{-1}$ ampicillin, 34 $\mu\text{g ml}^{-1}$ chloramphenicol). After overnight incubation at 37 °C, three colonies were used to inoculate a 30-ml MDG starter culture for 16 h (37 °C, 230 rpm). M9ZB expression cultures of 1 l in volume (47.8 mM Na₂HPO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 85.6 mM NaCl, 1% casamino acids, 0.5% glycerol, 2 mM MgSO₄, 2–50 μM trace metals, 100 $\mu\text{g ml}^{-1}$ ampicillin, 34 $\mu\text{g ml}^{-1}$ chloramphenicol) were then inoculated with 15 ml of MDG starter culture and grown (37 °C, 230 rpm) to an optical density (OD)₆₀₀ of 2.5 before induction with 0.5 mM isopropyl- β -D-thiogalactoside (IPTG) for 16 h (16 °C, 230 rpm). After overnight expression, cell pellets were collected by centrifugation and then resuspended and lysed by sonication in 60 ml of lysis buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 10% glycerol, 30 mM imidazole, 1 mM DTT). Lysate was clarified by centrifugation at 50,000g for 30 min, supernatant was poured over 8 ml of Ni-NTA resin (Qiagen), resin was washed with 35 ml of lysis buffer supplemented with 1 M NaCl and protein was eluted with 20 ml of lysis buffer supplemented with 300 mM imidazole. Samples were then dialysed overnight in dialysis tubing with a 14-kDa molecular weight cut-off (Ward's Science), and SUMO2 tag cleavage was carried out with recombinant human SENP2 protease as previously described⁴³. Proteins were dialysed overnight at 4 °C in dialysis buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM DTT), and then purified further by size exclusion chromatography using a 16/600 Superdex 75 or 16/600 Superdex 200 column (Cytiva) in gel filtration buffer (20 mM HEPES-KOH pH 7.5, 250 mM KCl, 1 mM TCEP). Purified proteins were concentrated to greater than 15 mg ml⁻¹ using 30-kDa molecular weight cutoff (MWCO) centrifugal filter units (Millipore Sigma), aliquoted, flash-frozen in liquid nitrogen and stored at -80 °C. HalB wild-type and mutant proteins were subsequently treated with or without P1 nuclease (Sigma catalogue no. N8630) and analysed using a 15% Coomassie-stained SDS-PAGE gel. See Supplementary Fig. 1 for uncropped SDS-PAGE gels.

For HalA protein purification, co-expression with HalB was required to neutralize HalA cellular toxicity. Expression plasmid was transformed into BL21(DE3) LOBSTR cells to reduce co-purification of ArnA⁴⁵ and plated onto MDG media. After overnight incubation at 37 °C, six colonies were used to inoculate a 30-ml MDG starter culture supplemented with 1% glucose for 16 h (37 °C, 230 rpm). M9ZB supplemented with 1% glucose expression cultures of 1 l in volume were then inoculated with 15 ml of MDG starter culture and grown (37 °C, 230 rpm) to an OD₆₀₀

of 3.5 before induction with 0.5 mM IPTG for 16 h (16 °C, 230 rpm). After overnight expression, cell pellets were collected by centrifugation and then resuspended and lysed by sonication in 60 ml of lysis buffer. Lysate was clarified by centrifugation at 50,000g for 60 min, and the pellet containing the membrane fraction was homogenized in lysis buffer supplemented with 1% solgrade DDM detergent (Anatrace) using a T25 homogenizer (IKA) for 3 min at 10,000 rpm and incubated with stirring at 4 °C for 2 h. The mixture was centrifuged at 50,000g for 45 min and the supernatant was poured over 4 ml of Ni-NTA resin. Resin was washed with 35 ml of lysis buffer supplemented with 1 M NaCl, and protein was eluted with 20 ml of lysis buffer supplemented with 300 mM imidazole. Buffers used for washing and elution were supplemented with 0.01% anagrade DDM (Anatrace) and 0.04% GDN detergent (Anatrace), respectively. Proteins were dialysed overnight at 4 °C in dialysis buffer supplemented with 0.01% DDM, and then purified further by size exclusion chromatography using a 16/600 Superdex 200 column in gel filtration buffer supplemented with 0.02% GDN. Purified protein used for biochemical assays was concentrated to greater than 5 mg ml⁻¹ using 100-kDa MWCO centrifugal filter units, aliquoted, flash-frozen in liquid nitrogen and stored at -80 °C. For cryo-EM, a single 1-ml fraction collected from the peak A280 absorbance was concentrated to 8.4 mg ml⁻¹ using 100-kDa MWCO centrifugal filter units. Endogenously synthesized ODA made by HalB is bound to purified HalA protein to form a HalA-ODA complex that is used for both biochemical assays and structure determination.

Bioinformatic analysis of Hailong and SEC ϕ 4 Exo

The genomic sequences of HalA and HalB and closely related family members were aligned using MAFFT as previously described^{2,46}. Briefly, sequences were identified using IMG BLAST searches⁴⁷ performed with an E value cut-off of 0.005 and HalA and HalB from *R. bacterium* QY30 as query proteins. Hailong genetic neighbourhood was manually curated from IMG and annotated using Defence Finder¹⁰. Phylogenetic analysis of SEC ϕ 4 Exo was performed as previously described². Sequences were identified from PSI-BLAST and MMSeqs2 was used to remove protein redundancies (minimum sequence identity = 0.95, minimum alignment coverage = 1) and the final aligned 1,713 sequences were used to construct a phylogenetic tree in Geneious Prime (v.2024.0.5) using FastTree with no outgroup, and iTOL was used for tree visualization and annotation⁴⁸. AlphaFold modelling of apo HalA and SEC ϕ 4 Exo was performed using the AlphaFold 3 server⁴⁹. Structural comparisons between HalA and ion channels and between SEC ϕ 4 Exo and human EXO5 were performed using DALI and the FoldSeek server^{50,51}.

Phage challenge plaque assay

To measure defence, plaque assays were performed on bacterial strains expressing the defence system from gammaproteobacteria (*E. coli*, *Klebsiella*, *Yersinia*) or alphaproteobacteria (*Rhodobacteraceae*) species and a negative control strain lacking the system as previously described⁴¹. Briefly, 300 μl of an overnight culture of *E. coli* MG1655 was mixed with 30 ml of pre-melted MMB lysogeny broth (LB) agar with 0.1 mM MnCl₂, 5 mM MgCl₂, 5 mM CaCl₂ and 0.5% agar with the addition of 0.1 mM IPTG. The plates were left to solidify at room temperature for 1 h and then 10- μl drops from tenfold serial dilutions of the phage lysate in MMB were pipetted on top of the solidified layer containing the bacteria. After the drops dried, plates were incubated overnight at 25 °C and the PFUs per ml were counted. Fold defence was calculated as the ratio between the PFU per ml values retrieved using the same phage lysate on control bacteria and bacteria containing the candidate defence system. When individual plaques could not be differentiated, a faint lysis zone across the drop area was counted as ten plaques.

Bacterial growth assay

HalA effector function was measured in *E. coli* using conditions that result in activation of the Hailong defence system. *E. coli* Top10

competent cells were transformed with HalA wild type or mutants alone (pBAD03), HalA wild type or mutants expressed with HalB wild type or mutants (pBAD03), or HalA and HalB (pBAD03) co-expressed with SECφ4 Exo wild type or mutants (pCA24N). Transformations were plated onto LB plates supplemented with 1% glucose, and three colonies were picked and grown for 16 h (37 °C, 230 rpm) in 5-ml LB starter cultures. Cells were pelleted and then resuspended in PBS, and 200 µl was added to wells of a 96-well plate. Serial dilutions were spotted (5 µl) on LB plates supplemented with 100 µg ml⁻¹ ampicillin and 0.2% L-arabinose for pBAD03-carrying plasmids, and/or 34 µg ml⁻¹ chloramphenicol and 5 µM IPTG for pCA24N-carrying plasmids. Plates were allowed to dry for 1 h and incubated overnight at 37 °C. A reduction in colony forming units indicated HalA activation resulting from cell growth arrest.

Crystallization and structure determination

Crystals were grown in hanging-drop format using EasyXtal 15-well trays (NeXtal). Crystals of HalB wild type and mutants were grown at 18 °C in 2-µl drops mixed 1:1 with purified protein (4 mg ml⁻¹, 20 mM HEPES-KOH pH 7.5, 80 mM KCl, 1 mM TCEP) and reservoir solution (0.01 M KCl, 0.1 M HEPES pH 7.0, 10–15% PEG-4000). Crystals were grown for 1–3 d before being cryo-protected with reservoir solution supplemented with 30% ethylene glycol and collected by freezing in liquid nitrogen. Crystals of HalB R164A mutant were additionally grown in the presence of 5 mM MgCl₂, 1 mM MgCl₂, and 50 µM non-hydrolysable dATP (dAp(c)pp, Jena Bioscience). HalB R164A mutant ODA synthesis is severely attenuated rather than completely inhibited. Purified HalB R164A protein used for crystallography is capable of efficiently attaching only one molecule of dAMP onto Y227, and individual protomers in the crystal asymmetric unit occur as unmodified Y227 with no dAMP attached, or Y227 with a covalent bond to a single molecule of dAMP that co-purified with the protein. X-ray diffraction data were collected at NSLS2 (beamlines AMX and FMX) and Diamond Light Source (beamline I03), and data were processed manually using the SSRL autoxds script (A. Gonzalez, Stanford SSRL) and automatically using xia2.multiplex (ref. 52), respectively. Model building was performed using Coot, and then refined in Phenix. Statistics were analysed as described in Supplementary Table 3. See Supplementary Table 3 and the Data availability statement for the deposited PDB codes. All structure figures were generated with Pymol 2.5.5.

Nucleotide synthesis assay

To measure nucleotide synthesis, 50 µM of purified HalB wild-type and mutant proteins were incubated with 50 µM each of unlabelled ATP, CTP, GTP, UTP (New England Biolabs), dATP, dCTP, dGTP, dTTP (Thermo Fisher), or 500 µM each of ddATP, ddCTP, ddGTP, ddTTP (Jena Bioscience), and 0.5 µl of α³²P-labelled NTPs or dNTPs (Perkin Elmer/Revvity, approximately 0.5 µCi each of NTP or dNTP) in reaction buffer (50 mM Tris-HCl pH 9.4, 100 mM KCl, 10 mM MgCl₂, 1 mM MnCl₂, 1 mM TCEP) at 37 °C for 4 h. α³²P-labelled dAMP was generated by treating α³²P-labelled dATP with apyrase (New England Biolabs) at 37 °C for 30 min. α³²P-labelled dGTP is supplied as a solution premixed with a proprietary EasyTide loading dye (Perkin Elmer/Revvity) and this loading dye was therefore included within these individual reactions. Synthesis reactions using increasing nucleotide concentration were performed using 0.5, 5, 10, 50, 100, 500, 1,000 and 2,000 µM dNTP concentrations and incubated for 24 h to allow dNTP incorporation to reach completion. Reactions were terminated with addition of 0.5 µl of Quick CIP phosphatase (New England Biolabs) to remove terminal phosphate groups from unreacted nucleotides. Reactions were further treated with proteinase K (GoldBio) to release ODA product from the protein. Each reaction was analysed on a 20% urea-PAGE gel and exposed to a phosphor-screen before imaging with a Typhoon Trio Variable Mode Imager (GE Healthcare). Images were exported to Adobe Photoshop, and brightness, gamma and contrast adjustments were applied to the

entire image. Reactions that were used for ODA composition analysis were separated from unreacted nucleotides by centrifugation through a 3-kDa MWCO filter unit and further treated with nuclease P1 and Quick CIP phosphatase at 37 °C for 30 min before LC-MS analysis. Samples were analysed on an Agilent 1260 HPLC equipped with a diode array detector and an Agilent 6125 single quadrupole mass spectrometer in negative ion mode using a reverse-phase Agilent InfinityLab Poroshell SB-Aq column (2.7-µm particle size, 2.1-mm inner diameter, 150-mm length) at a flow rate of 0.45 ml min⁻¹ with a gradient from 100% solvent A (0.1% ammonium formate) to 100% solvent B (methanol). Formate (45 amu from ammonium formate LC-MS buffer) and chloride (35 amu from KCl in reaction buffer) ions formed the major adducts [dA+formate]⁻ and [dA+Cl]⁻, respectively, observed with deoxyadenosine. See Supplementary Fig. 1 for uncropped urea-PAGE gels.

ODA cleavage activity-guided fractionation and mass spectrometry

E. coli BL21 cells were grown and lysed as described above. Lysate fractionation was performed using a general workflow as previously described^{53,54}. Briefly, soluble cell lysates were subjected to 50% (NH₄)₂SO₄ cut at 4 °C for 1 h with continuous stirring, then centrifuged at 50,000g for 30 min to remove precipitated proteins. The resulting clarified supernatant was passed through a 0.22-µm filter and fractionated by hydrophobic interaction chromatography using a 5-ml HiTrap Phenyl column (Cytiva) then eluted with a gradient of 2.00 M to 0.00 M (NH₄)₂SO₄. Alternatively, soluble cell lysates were fractionated by ion exchange using a 5-ml HiTrap Heparin HP column (Cytiva) then eluted with a gradient of 0.15 M to 2.00 M NaCl. Active hydrophobic interaction chromatography or ion exchange fractions were pooled, concentrated and further fractionated by size exclusion chromatography with a Superdex 200 Increase 10/300 GL column (Cytiva). After each purification step, 1 µl of each fraction was tested for ODA cleavage activity by incubating with HalB-ODA labelled with α³²P-dATP as described above at 30 °C for 90 min. Active fractions from both HIC and IEX samples were selected for analysis by label-free mass spectrometry to identify host factor candidates. See Supplementary Fig. 1 for uncropped urea-PAGE gels.

Electrophoretic mobility shift assay

HalA interaction with nucleotide ligands (Integrated DNA Technologies) was monitored by electrophoretic mobility shift assay as previously described^{2,15}. Briefly, 20 nM of fluorescein-labelled (6-FAM) ODA nucleotide of varying lengths (5, 6, 7, 8, 9, 10, 15 nucleotides) was incubated with increasing concentrations of purified HalA-ODA complex (0.005–50 µM, protein concentration calculated as HalA-ODA tetramer with an extinction coefficient of 45,380 M⁻¹ cm⁻¹) for 15 min at room temperature to measure binding of unoccupied sites in the HalA-ODA complex. Binding to alternative nucleotide ligands (20 nM of dC, dT, dN, rA or dsDNA) was performed with 10 µM HalA-ODA complex. Reactions were analysed using a 5% non-denaturing PAGE gel and imaged on a ChemiDoc imaging system (BioRad).

DNA cleavage assay

To measure nuclease activity, 40 µM of purified SECφ4 Exo wild-type and mutant proteins were incubated with 50 nM of nucleotide substrates (Integrated DNA Technologies) in reaction buffer (50 mM Tris-HCl pH 7.0, 10 mM KCl, 5 mM MgCl₂, 1 mM MnCl₂) at room temperature for 90 min. After incubation, reactions were stopped with DNA loading buffer containing 95% deionized formamide and 20 mM EDTA, and analysed on a 20% denaturing urea-PAGE gel and imaged on a ChemiDoc imaging system. To measure cleavage of ODA from the HalB-ODA complex, soluble lysate fraction was obtained from *E. coli* BL21 (DE3) cells lysed by sonication in lysis buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 10% glycerol, 1 mM DTT) and clarified by centrifugation at 50,000g for 30 min. Lysates were then incubated with CIP-treated and

Article

$\alpha^{32}\text{P}$ -labelled HalB–ODA at 37 °C for 90 min. Reactions were analysed on a 20% urea–PAGE gel and exposed to a phosphor-screen before imaging with a Typhoon Trio Variable Mode Imager (GE Healthcare). See Supplementary Fig. 1 for uncropped urea–PAGE gels.

Cryo-EM sample preparation and data collection

Grids (Quantifoil R 1.2/1.3 400 Mesh, Cu, Electron Microscopy Sciences) were glow-discharged in air using an easiGlow (Pelco) (negative, 0.38 mBar, 15 mA, 30-s glow, 10-s hold). In a Vitrobot Mark IV (Thermo Fisher) held at 8 °C and 100% humidity, 4 μl of GDN-solubilized HalA (8.4 mg ml⁻¹) was applied to a grid, which was then blotted (6 s, force 15), plunged into liquid ethane and stored in liquid nitrogen. We collected 10,951 videos on a Titan Krios G3i microscope operated at 300 kV with a Selectris energy filter (10-eV slit width) and a Falcon 4i direct electron detector (Thermo Fisher) operated in counting mode (55 frames, 4,096² pixels, $\times 165,000$ magnification, pixel size 0.736 Å, total dose 53.8 electrons per Å²). Data were collected using EPU v.3.6. Three shots per hole were recorded, with a defocus range of –0.6 to –2.0 μm .

Cryo-EM data processing

All data processing software was compiled and configured by SBGrid⁵⁵. CryoSPARC v.4.4.1 to v.4.6.0 (ref. 56) was used to perform all data processing unless indicated otherwise. Videos underwent patch-based motion correction with dose weighting and patch-based CTF estimation, followed by manual curation to exclude micrographs with CTF fit resolution greater than 5 Å and outliers in intensity, ice thickness, motion and defocus. Particles were picked using Topaz v.0.2.5 (ref. 57), implemented with a model that was trained on the particles yielded by a preliminary round of processing and volume reconstruction (which applied a strategy analogous to the one described here to 100 videos, except particle picking was done with Blob Picker on images processed by Micrograph Denoiser, and particles on carbon were manually excluded). Particles underwent extraction with a 384-pixel box size and two rounds of two-dimensional classification to exclude junk picks. Ab initio reconstruction and heterogeneous refinement (five classes) yielded two classes with high-resolution volume reconstructions. These particles underwent a second round of ab initio reconstruction and heterogeneous refinement (five classes).

Three classes produced high-resolution volume reconstructions. One class was a monomer of the tetrameric HalA complex. The other two classes were dimers of the tetramer. (For this methods section, we define the terms ‘monomer’ and ‘dimer’ considering the tetrameric complex within a single micelle to constitute an individual protomer). Both dimers were C_2 -symmetric, although reconstructed in C_1 at this stage, with the protomer closer to the centre of the box showing stronger density. However, the dimerization mode was distinct in each class, occurring through different protein surfaces. At the site of each dimerization interaction, a few side-chains were shifted to form protomer–protomer contacts, breaking symmetry with the C_4 -related surfaces uninvolved in dimerization, but in such a minor way that the dimer particle stacks remained useful in improving global resolution through C_4 reconstruction later in processing. These dimerization modes are unlikely to be biologically relevant, as they could not occur in near-planar biological membranes owing to the sharp angle between the two protomers’ C_4 axes (33° and 57°). In many cases, only one of the two protomers of each dimer was represented in the particle stack at this stage (some particles were eliminated by the proximity criterion during two-dimensional classification, and some were never picked by Topaz in the first place). To capture all protomers, dimeric particles were duplicated, and each duplicate particle’s pose was transformed across the C_2 axis using Align 3D Volumes. Particles from all three classes—monomers, 33° dimers and 57° dimers—were re-extracted with new centres (corresponding to the centre of an autogenerated mask from a C_4 refinement of the monomer class) and the poses assigned to them by refinement or by the C_2 transformation, where

applicable. Within each of the three classes, the particle stacks were then de-duplicated with a 20-Å separation criterion.

From this point on, all refinements were restricted with Gaussian priors on shift (standard deviation 5 Å) and rotation (standard deviation 10°) to prevent poses from wandering to symmetry-related positions (within the same protomer or in neighbouring protomers in the same particle micrograph) besides the one identified in the first refinement. In their three separate classes, particles underwent cycles of non-uniform refinement (C_4 for monomers, C_1 for dimers)⁵⁸, local CTF refinement and global CTF refinement until convergence. Particles then underwent reference-based motion correction, followed by similar rounds of refinement.

Particles from the three classes were then combined, de-duplicated and refined together (C_4 , gold-standard Fourier shell correlation (FSC) 2.07 Å). Performing this refinement with versions of the dimer particle micrographs in which the off-centre protomer had been subtracted yielded a reconstruction with equivalent resolution, so the original non-subtracted images were used. The per-particle scale factors, minimized during this refinement, exhibited a bimodal distribution (a similar result occurred for the monomer stack alone), and manual micrograph inspection revealed that the stack contained on-carbon particles, suggesting that low-quality particles were contributing to the reconstruction. Ten iterations of CryoSieve v.1.2.5 (ref. 59) were run on this particle stack, with a per-round retention ratio of 0.8 and a high-pass cut-off incrementing linearly (in frequency space) from 40 Å to 2.07 Å. Although the map did not change significantly across iterations, as observed previously, the optimal FSC, as determined by RELION v.4.0.1 (ref. 60) within the CryoSieve cycles, was achieved at iteration 4, at which point on-carbon particles had disappeared from the stack and scale factor minimization yielded a monomodal distribution. This particle stack (231,898 particles) was subjected to cycles of non-uniform refinement (C_4), local CTF refinement and global CTF refinement (fit for beam tilt, beam trefoil, spherical aberration and beam tetrafoil) until convergence, yielding a gold-standard FSC resolution of 1.98 Å after mask auto-tightening.

The resulting half-maps were then subjected to a single cycle of model-blind density modification and sharpening using resolve_cryo_em v.2.15 (ref. 61) as implemented in Phenix v.1.21.1–5286 (ref. 62). During subsequent model refinement, the same starting model achieved a superior $d_{\text{FSC, model}}(0.5)$ when refined into the density-modified map (1.9 Å) as compared with the unmodified map (2.1 Å), so the density-modified map was selected for final model refinement.

Cryo-EM model building

The starting model, an AlphaFold2-predicted structure of the HalA tetramer, was docked into the electron microscopy density in Coot. DNA was built manually in Coot. Waters were built using phenix.douse, followed by manual curation. Densities for amino acid residues 1–13, 232–241 and 357–359 were weak, and these residues were left unmodelled. Unidentified density penetrating from the micelle into the channel lumen, potentially contributed by lipids, was left unmodelled. Unidentified adducts on the thiols of Cys54 and Cys163 were left unmodelled. The model was iteratively refined in Phenix (with non-crystallographic symmetry constraints applied to protein and DNA chains) and manually adjusted in Coot. To generate the final model, chains A, E and W were cloned with perfect C_4 symmetry. Structure stereochemistry statistics are reported in Supplementary Table 3. Figures were prepared in PyMOL 2.5.5.

Flow cytometry assay

To measure membrane depolarization, *E. coli* Top10 cells containing arabinose-inducible plasmids expressing HalA or HalA and HalB were grown in LB culture at 37 °C for 16 h. Overnight starter culture was diluted 1:10 in LB cultures supplemented with 1% glucose or 0.5% arabinose and incubated for 1, 2, 3, 6 or 8 h at 37 °C. Polymyxin B (10 $\mu\text{g ml}^{-1}$,

Sigma catalogue no. P1004) treatment at 37 °C for 30 min was used as a positive control. Following incubation, bacterial suspensions were treated with DiBAC4 (5 µM, Thermo Fisher catalogue no. B438) and propidium iodide (1 µg ml⁻¹ Sigma) in PBS for 5 min. Cells were analysed using a BD FACSymphony A1 Cell Analyzer by measurement of 100,000 events (cells) per sample. Cells treated with polymyxin B were used as a positive control for gating of DiBAC4- and propidium iodide-positive cells, and cells treated with glucose and cells expressing both HalA and HalB were used as negative controls for gating of DiBAC4- and propidium iodide-negative cells. An example of the flow cytometry gating strategy is included in Extended Data Fig. 7c. Data were analysed using FlowJo v.10 software.

Fluorescence microscopy

To measure membrane depolarization during phage infection, *E. coli* MG1655 cells containing plasmids expressing Hailong anti-phage defence were incubated at 37 °C overnight in LB culture supplemented with 0.2 mM IPTG to induce protein expression. Where indicated, the bacteria were additionally transformed with a plasmid expressing EGFP-HalA and periplasmic mCherry as previously described⁶³, and subjected to further selection with 50 µg ml⁻¹ spectinomycin. Overnight culture was diluted in MMB culture supplemented with 5 µM DiBAC4 to an OD of 0.2 and infected with wild-type or D113V escape mutant phage SECφ4 at multiplicity of infection 10 or 40. Bacteria were then plated on agarose pads (LB supplemented with 2% agarose, 0.2 mM IPTG and 5 µM DiBAC4) and immediately covered with a coverslip and sealed with VALAP wax as previously described⁶⁴. Slides were imaged every 30 min for 6 h after initial infection; during time-lapse imaging, slides were maintained at 25 °C.

To measure ion permeability, *E. coli* Top10 cells containing arabinose-inducible plasmids expressing HalA or HalAB were grown in LB culture supplemented with 0.5% arabinose at 37 °C for 6 h. Cells were pelleted, washed with PBS and resuspended with PBS supplemented with 10 mM EDTA at room temperature for 10 min to facilitate the removal of LPS in outer membrane. Cells were then washed twice with PBS supplemented with 0.1 mM EDTA and diluted to an OD of 0.2 in PBS supplemented with 0.1 mM EDTA and 40 µM Sodium Green Tetraacetate (Thermo Fisher catalogue no. S6900) diluted 1:1 in Pluronic F127 (Thermo Fisher) at room temperature for 30 min. Cells were washed three times with PBS supplemented with 0.1 mM EDTA and plated on agarose pads (LB supplemented with 2% agarose) and immediately covered with a coverslip and sealed with VALAP wax. Snapshot images were captured for 10 min.

All images were acquired using a Nikon Ti microscope equipped with Plan Apo ×100/1.40 Oil Ph3 DM objective and Andor Zyla 4.2 Plus sCMOS camera and Nikon Elements v.5.22 acquisition software. Images were analysed using Fiji v.2.9.0 and Adobe Photoshop was used for brightness and contrast adjustments.

Escape phage isolation

To isolate mutant phages that overcome defence, a drop assay was performed as described above with Hailong using the bacterium *E. coli* STEC 1178 or a negative control strain lacking the defence system. Tenfold serial dilutions in MMB were made for the ancestor phage lysate and then 10-µl drops from the dilution series were dropped on top of the solidified layer containing the bacteria. After the drops dried, plates were incubated overnight at 25 °C. Single plaques were isolated on the defence-containing strain into 100 µl of phage buffer (50 mM Tris pH 7.4, 100 mM MgCl₂, 10 mM NaCl) and another drop assay was performed with the new phages and the ancestral phages as a comparison. PFUs observed after overnight incubation were counted and used to calculate the fold defence using the ratio between phage PFUs obtained on negative control cells and PFUs obtained on defence-containing cells. In cases for which the escape phage titre was low, the phages were propagated on the defence system. DNA was extracted using the

QIAGEN DNeasy Blood and Tissue kit and DNA libraries were prepared using a modified Nextera protocol⁶⁵ for Illumina sequencing.

Statistics and reproducibility

Experimental details regarding replicates and sample size are described in the figure legends.

Reporting summary

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

Data availability

Coordinates and structure factors of the Hailong HalB–ODA complex and HalB mutant enzymes have been deposited in PDB under the accession codes 9DBH, 9DBI and 9DBJ. Coordinates and density maps of the HalA–ODA complex have been deposited in PDB and EMDB under the accession codes 9NYI and EMD-49920. Source data for Fig. 4b,j and Extended Data Figs. 2h and 7a,b are provided with this paper. All other data are available in the manuscript or the Supplementary Information.

- Zhou, W. et al. Structure of the human cGAS–DNA complex reveals enhanced control of immune surveillance. *Cell* **174**, 300–311 e311 (2018).
- Pahl, K. S. et al. A new antibiotic traps lipopolysaccharide in its intermembrane transporter. *Nature* **625**, 572–577 (2024).
- Andersen, K. R., Lekska, N. C. & Schwartz, T. U. Optimized *E. coli* expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* **81**, 1857–1861 (2013).
- Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).
- Chen, I. A. et al. The IMG/M data management and analysis system v.7: content updates and new features. *Nucleic Acids Res.* **51**, D723–D732 (2023).
- Leticnic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
- Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
- van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).
- Holm, L., Laiho, A., Toronen, P. & Salgado, M. DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci.* **32**, e4519 (2023).
- Gildea, R. J. et al. xia2.multiplex: a multi-crystal data-analysis pipeline. *Acta Crystallogr. D Struct. Biol.* **78**, 752–769 (2022).
- Eaglesham, J. B., Pan, Y., Kupper, T. S. & Kranzusch, P. J. Viral and metazoan poxins are cGAMP-specific nucleases that restrict cGAS–STING signalling. *Nature* **566**, 259–263 (2019).
- Hobbs, S. J. et al. Phage anti-CBASS and anti-Pycsar nucleases subvert bacterial immunity. *Nature* **605**, 522–526 (2022).
- Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).
- 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).
- Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).
- Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
- Zhu, J. et al. A minority of final stacks yields superior amplitude in single-particle cryo-EM. *Nat. Commun.* **14**, 7822 (2023).
- Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
- Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *Nat. Methods* **17**, 923–927 (2020).
- Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
- Uehara, T., Parzych, K. R., Dinh, T. & Bernhardt, T. G. Daughter cell separation is controlled by cytolethal ring-activated cell wall hydrolysis. *EMBO J.* **29**, 1412–1422 (2010).
- Wang, X. & Montero Llopis, P. Visualizing *Bacillus subtilis* during vegetative growth and spore formation. *Methods Mol. Biol.* **1431**, 275–287 (2016).
- Baym, M. et al. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS ONE* **10**, e0128036 (2015).

Acknowledgements We thank members of the Kranzusch and Sorek laboratories for helpful comments and discussion. We thank D. Wasserman and S. Yamaguchi from the Kranzusch laboratory for assistance with LC–MS analysis, and M. Gilman from the Kruse laboratory for assistance with cryo-EM data collection and processing. Mass spectrometry was performed at the Harvard Taplin Mass Spectrometry Facility with assistance from R. Tomaino. The work was funded by grants to P.J.K. from the Pew Biomedical Scholars programme, the Burroughs Wellcome Fund PATH programme, the G. Harold and Leila Y. Mathers Charitable Foundation, the Mark Foundation for Cancer Research, the Cancer Research Institute, the Parker Institute for Cancer Immunotherapy and the National Institutes of Health (grant no. 1DP2GM146250-01); grants to R.S. from the European Research Council (grant no. ERC-AdG GA 101018520).

Article

the Israel Science Foundation (MAPATS grant no. 2720/22), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, the Deutsche Forschungsgemeinschaft (SPP 2330, grant no. 464312965), a research grant from the Estate of Marjorie Plesset and the Knell Family Center for Microbiology; a grant to A.C.K. from the National Institutes of Health (grant no. U19AI158028); grants to M.J. from the National Institutes of Health (grant no. 1DP2GM154152) and the G. Harold and Leila Y. Mathers Charitable Foundation; and a grant to J.d.M. from National Institutes of Health (grant no. 5R00DC019401) and the Howard Hughes Medical Institute. J.M.J.T. is supported by a Servier PhD Fellowship. S.J.H. is supported through a Cancer Research Institute Irvington Postdoctoral Fellowship (grant no. CRI3996), and J.C.C. is supported by a Helen Hay Whitney postdoctoral fellowship. X-ray data were collected through support from an agreement between the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357, and the Diamond Light Source, the UK's national synchrotron science facility, located at the Harwell Science and Innovation Campus in Oxfordshire. X-ray data were additionally collected at the Center for Bio-Molecular Structure (CBMS) which is primarily supported by the NIH-NIGMS through a Center Core P30 Grant (no. P30GM133893), and by the DOE Office of Biological and Environmental Research (grant no. KP1607011). NSLS2 is a US DOE Office of Science User Facility operated under Contract No. DE-SC0012704. This publication resulted from the data collected using the beamtime obtained through NECAT BAG proposal no. 311950. Cryo-EM data were collected at the Harvard Cryo-EM Center for Structural Biology at Harvard Medical School. Imaging experiments were supported by the Microscopy Resources on the North Quad (MicRoN) core at Harvard Medical School.

Author contributions The study was designed and conceived by J.M.J.T. and P.J.K. Crystallography, biochemical experiments, bacterial toxicity assays and live cell imaging microscopy were performed by J.M.J.T. Phage challenge assays and escape mutant generation and sequencing were performed by S.M. and R.S. Cryo-EM data collection, analysis and model building were performed by J.M.J.T., J.C.C. and A.C.K. Flow cytometry collection and analysis were performed by J.M.J.T., D.S. and M.J. SEC-MALS was performed by S.J.H. Bioinformatics analysis was performed by J.M.J.T. and J.d.M. The manuscript was written by J.M.J.T. and P.J.K. All authors contributed to editing the manuscript and support the conclusions.

Competing interests R.S. is a scientific cofounder and advisor of BiomX and Ecophage. A.C.K. is a cofounder and consultant for Tectonic Therapeutic and Seismic Therapeutic and for the Institute for Protein Innovation, a non-profit research institute. M.J. declares outside interest in Evozyne and DEM BioPharma. The other authors declare no competing interests.

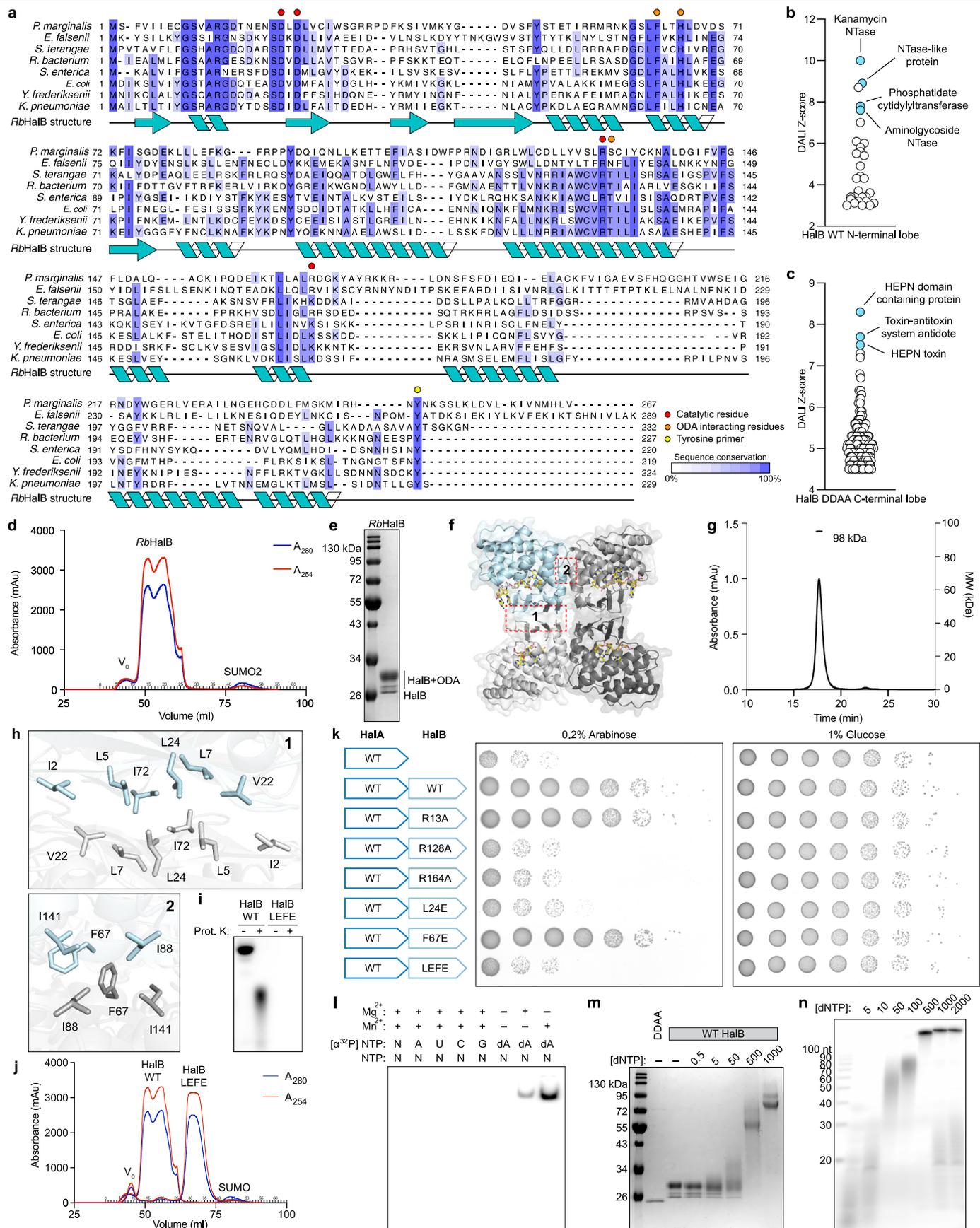
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-025-09058-z>.

Correspondence and requests for materials should be addressed to Philip J. Kranzusch.

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

Reprints and permissions information is available at <http://www.nature.com/reprints>.

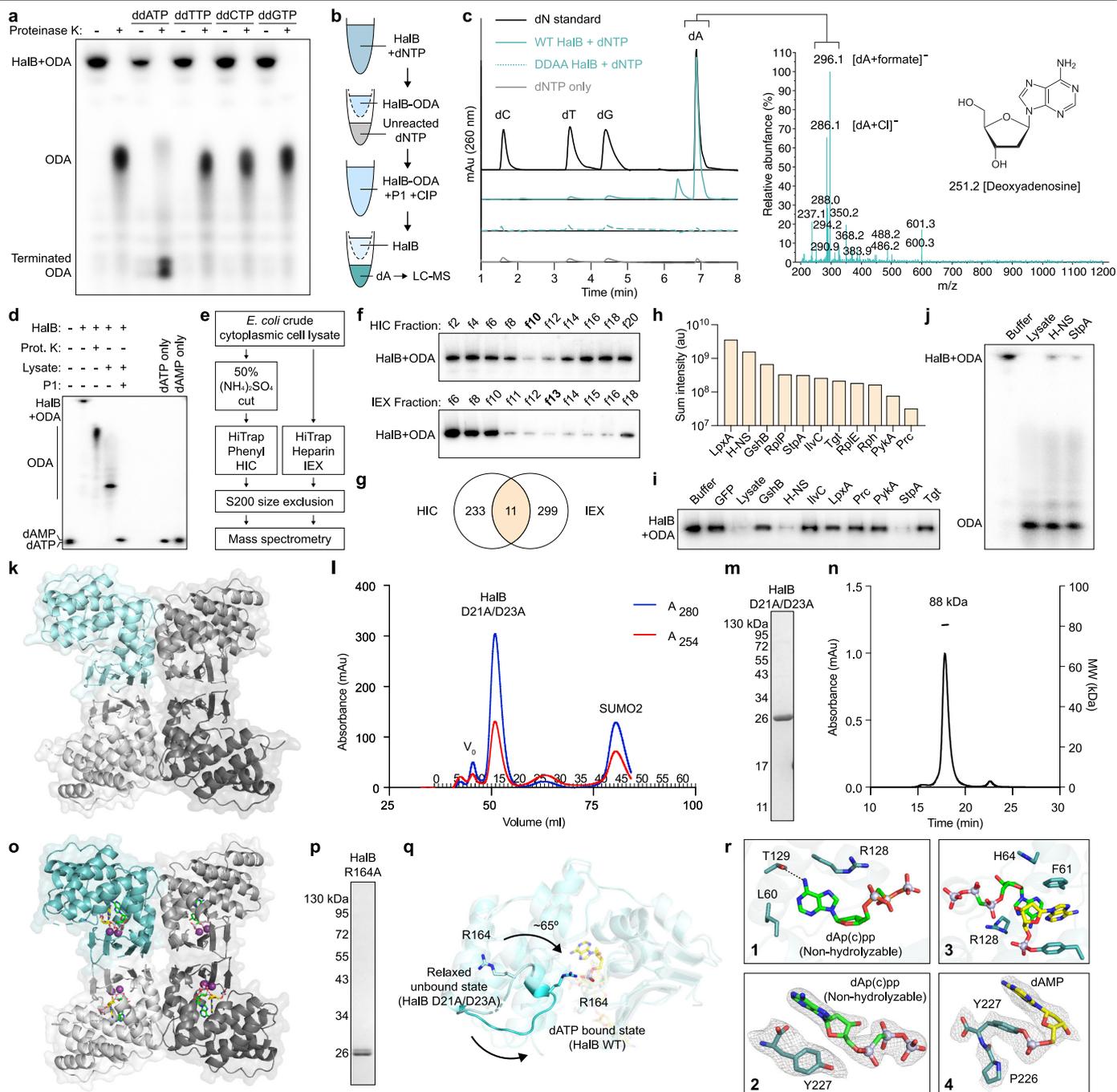


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Structural and biochemical analysis of HalB.

a, Structure guided multiple sequence alignment of HalB protein homologues from indicated bacterial species. Shading indicates degree of residue conservation. **b,c**, Protein structural homology of HalB WT N-terminal lobe (**b**) and HalB catalytically inactive mutant (DDAA) C-terminal lobe (**c**) against all entries in PDB showing the DALI Z score of the top 30 and 150 hits respectively. HalB DDAA was used in (**c**) as the structure of the mutant protein resolves additional residues in the C-terminal tail that are not visible in the wildtype structure. **d**, *R. bacterium* QY30 HalB was expressed as an N-terminal 6×His-SUMO2 fusion and purified by Ni-NTA and separated from His-SUMO2 by size exclusion chromatography. **e**, Coomassie-stained SDS-PAGE analysis of fully purified HalB from *R. bacterium* QY30. **f**, Overview of tetrameric HalB crystal structure with one protomer in blue and three protomers in different shades of grey. Numbered boxes highlight dimerization interface described in detail in (**h**). **g**, Size exclusion chromatography with multi-angle static light scattering

(SEC-MALS) analysis of purified HalB confirms HalB tetramerization. **h**, Detailed view of interacting residues as in (**f**) showing dimerization interface between HalB protomers. **i**, Urea-PAGE analysis of HalB WT and LEFE mutant ODA synthesis. **j**, Size exclusion chromatography overlay of HalB WT and LEFE mutant showing a rightward shift indicating a loss of tetrameric complex formation. **k**, Bacterial growth assay of *E. coli* expressing HalA with HalB active site and dimerization mutants. LEFE, L24E and F67E double mutant. **l**, Urea-PAGE analysis of HalB substrate and metal specificity using NTPs and $\alpha^{32}\text{P}$ -labeled dATP and NTPs as indicated. **m**, Coomassie-stained SDS-PAGE analysis of ODA synthesis with purified HalB WT and catalytically inactive mutant (DDAA) given indicated μM concentrations of dNTP. **n**, Urea-PAGE analysis of ODA synthesis over a 24 h period with purified HalB given $\alpha^{32}\text{P}$ -dNTP, the indicated μM concentrations of dNTP, and treated with proteinase K. Expression of HalB used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments.

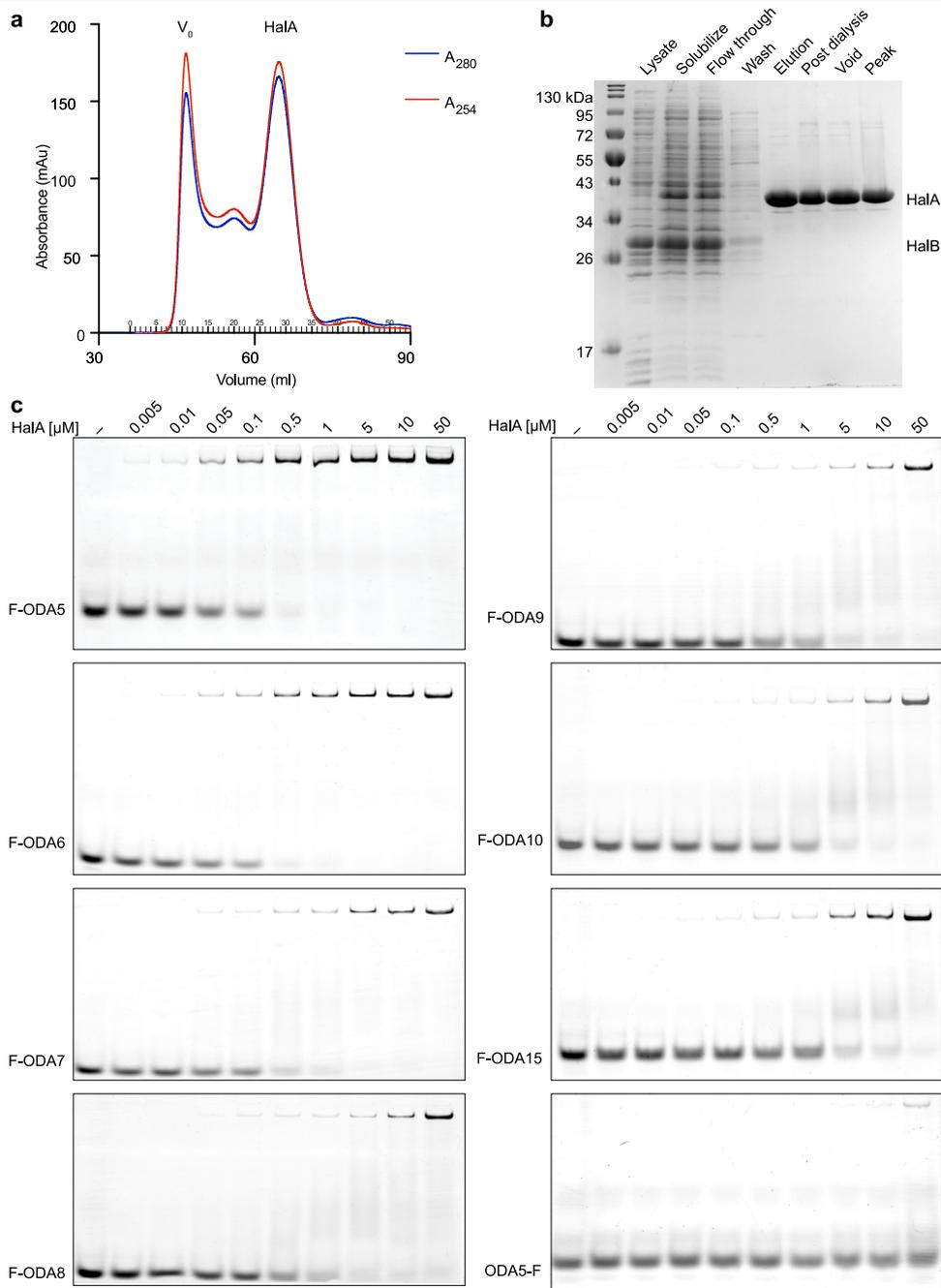


Extended Data Fig. 2 | See next page for caption.

Article

Extended Data Fig. 2 | Mechanism of ODA synthesis. **a**, Urea-PAGE analysis of ODA synthesis with purified HalB and $\alpha^{32}\text{P}$ -dATP supplemented with different ddNTP substrates to induce chain termination and determine specificity to adenine. **b**, Schematic of LC-MS workflow to determine the chemical composition of ODA. **c**, Left, LC-MS analysis of purified HalB WT or DDAA mutant incubated with dNTP. Right, Characterization of the extracted signal molecule by LC-MS in negative mode. Formate and chloride ions formed the major adducts $[\text{dA}+\text{formate}]^-$ and $[\text{dA}+\text{Cl}]^-$ respectively observed with deoxyadenosine. **d**, Uncropped Urea-PAGE analysis of ODA cleavage from the HalB-ODA complex using $\alpha^{32}\text{P}$ -labeled dATP treated with proteinase K and *E. coli* soluble lysate fraction as in Fig. 2. $\alpha^{32}\text{P}$ -labeled dAMP made with apyrase-treated dATP and nuclease P1 were used as controls to visualize single nucleotide product. **e**, Schematic depicting process of biochemical fraction of *E. coli* BL21 cell lysate to enrich for HalB-ODA cleavage activity using heparin IEX or ammonium sulfate precipitation followed by phenyl hydrophobic interaction. **f**, Urea-PAGE analysis of ODA cleavage from the HalB-ODA complex using fractions obtained after S200 size-exclusion chromatography. Active fractions used for mass spectrometry analysis are highlighted in bold. **g**, Summary of mass spectrometry results from both fractionation methods. HIC, hydrophobic-interaction chromatography; IEX, ion exchange. **h**, List of enriched candidates shared between both purification schemes. Data shown as mean for two independent experiments. **i**, Urea-PAGE analysis of ODA cleavage from the HalB-ODA complex using $\alpha^{32}\text{P}$ -labeled dATP treated with purified reconstitution of candidate proteins. Two common contaminating ribosomal proteins (RplP and RplE)

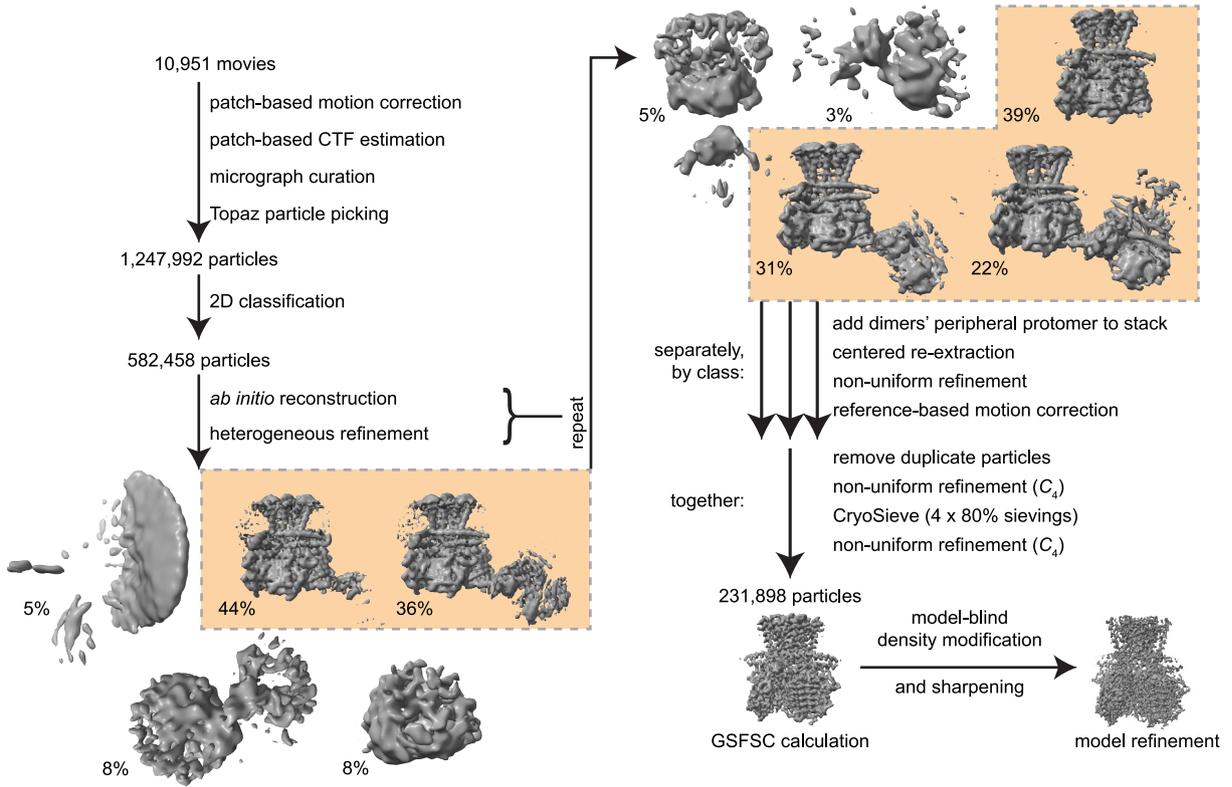
and one catalytically inactive protein due to a frameshift mutation (Rph) were excluded from further analysis. **j**, Urea-PAGE analysis of ODA cleavage from the HalB-ODA complex using $\alpha^{32}\text{P}$ -labeled dATP treated with purified H-NS and StpA. **k**, Overview of tetrameric HalB catalytically inactive mutant crystal structure with one protomer in light blue and three protomers in different shades of grey. **l**, *R. bacterium* QY30 HalB catalytically inactive mutant was expressed as an N-terminal 6 \times His-SUMO2 fusion and purified by Ni-NTA and separated from 6 \times His-SUMO2 by size exclusion chromatography. **m**, Coomassie-stained SDS-PAGE analysis of fully purified HalB catalytically inactive mutant. **n**, Size exclusion chromatography with multi-angle static light scattering analysis of purified HalB catalytically inactive mutant. **o**, Overview of tetrameric HalB R164A crystal structure with one monomer in dark blue and three monomers in different shades of grey. **p**, Coomassie-stained SDS-PAGE analysis of fully purified HalB R164A. **q**, Overview of HalB and residue R164 in an open conformation prior to ODA synthesis and an active confirmation during ODA synthesis. **r**, Detailed view of 1) HalB residues stabilizing incoming dATP substrate and 2) polder omit map contoured at 5.5 σ of Y227 and non-hydrolyzable dATP. HalB adenine discrimination occurs through residues T129 that coordinate sequence-specific contacts with the adenine nucleobase Watson-Crick edge, and L60 and R128 that restrict guanine and pyrimidine base recognition. Detailed view of 3) newly synthesized ODA bound to HalB C-terminal tyrosine residue and 4) polder omit map contoured at 3.5 σ of dAMP bound to Y227. Expression of HalB used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments.



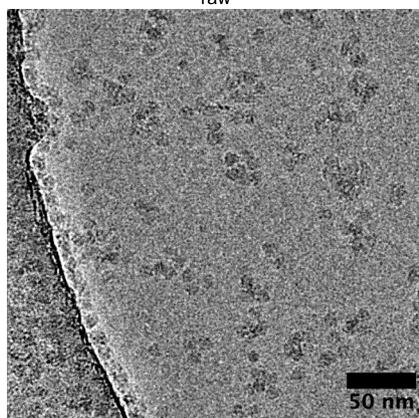
Extended Data Fig. 3 | HalA substrate binding specificity. *a*, *R. bacterium* QY30 HalA was expressed as an N-terminal 6×His fusion with untagged HalB and purified by Ni-NTA in the presence of DDM detergent and separated by size exclusion chromatography in the presence of GDN detergent. *b*, Coomassie-stained SDS-PAGE analysis of fully purified HalA. *c*, Electrophoretic mobility

shift assay of HalA-ODA complex formation with indicated FAM-labeled (abbreviated as F) DNA substrates. Expression of HalA used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments.

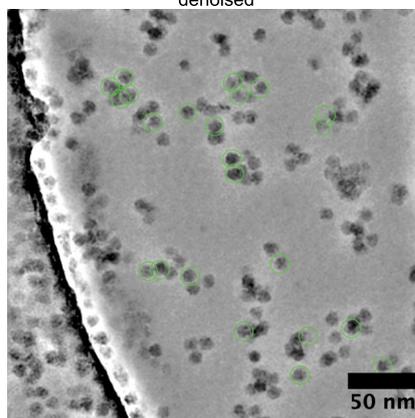
a



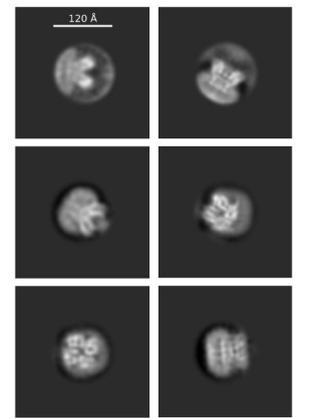
b



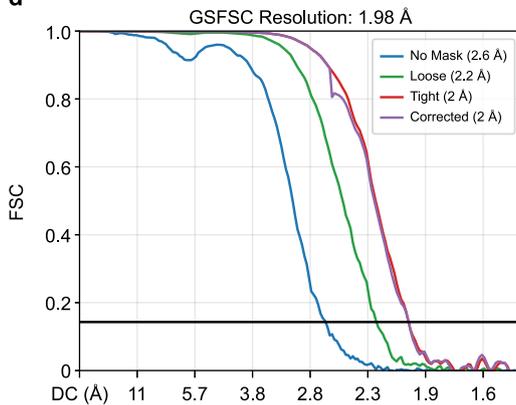
denoised



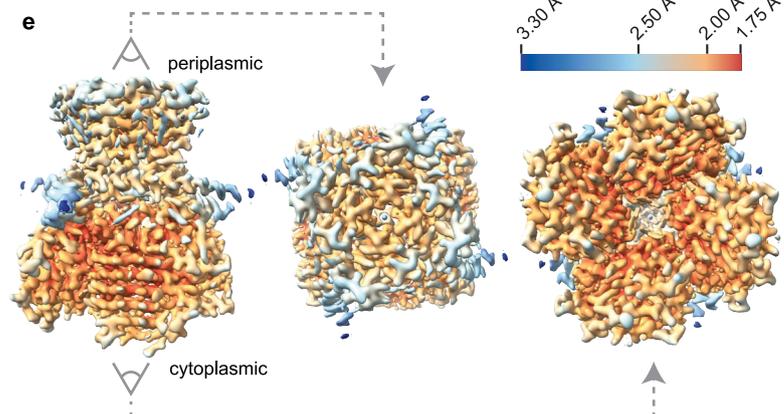
c



d



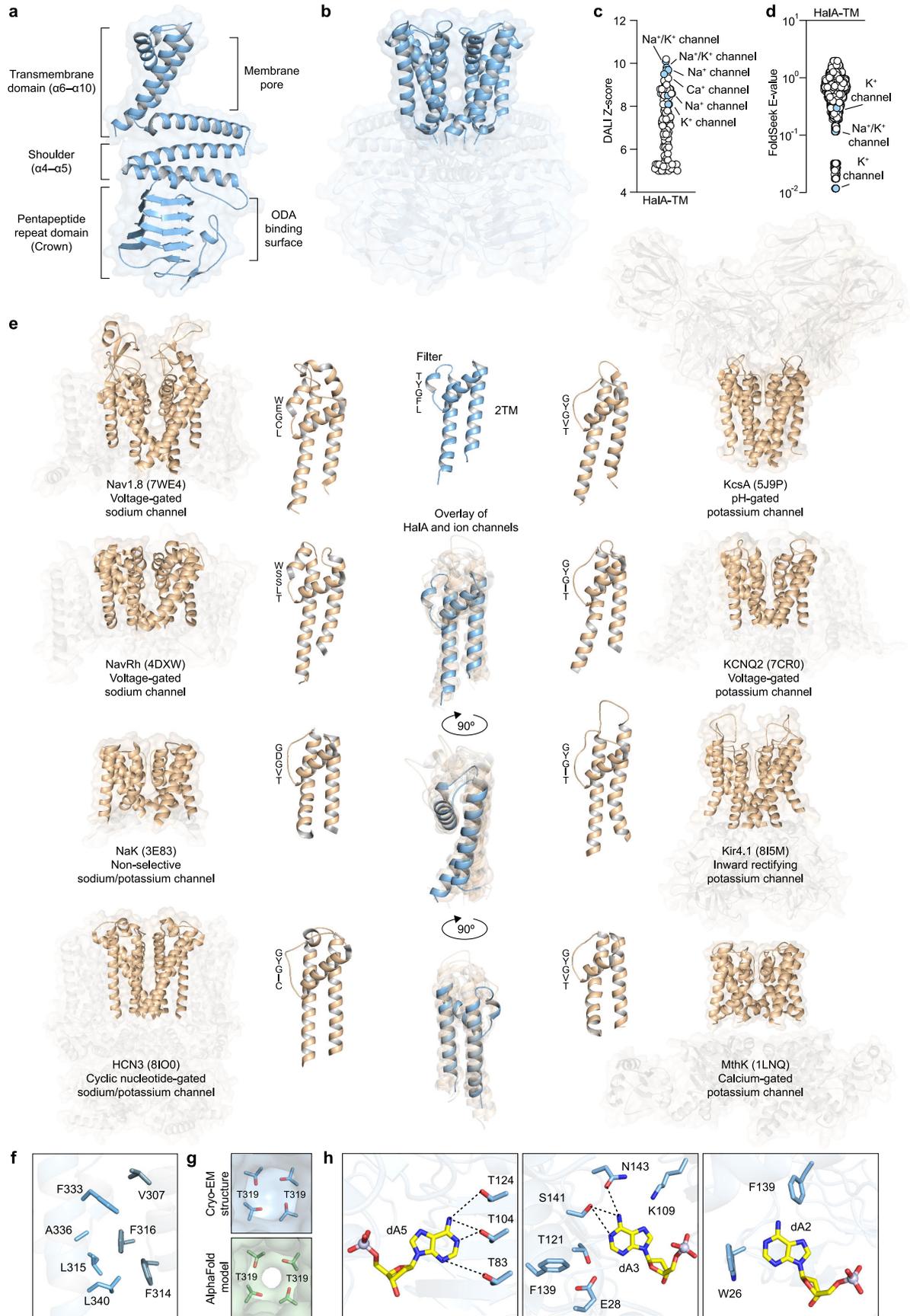
e



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

Extended Data Fig. 4 | Cryo-EM data processing for the HalA-ODA complex. **a**, Cryo-EM data processing scheme. To facilitate visual comparison, each volume's hand was flipped wherever necessary to match the hand of the final reconstruction. **b**, Left, example motion-corrected micrograph, subjected to a 5-Å long-pass filter. Right, the same image subjected to CryoSPARC's Micrograph Denoiser. Particles present in the final stack are circled in green. **c**, Example 2D

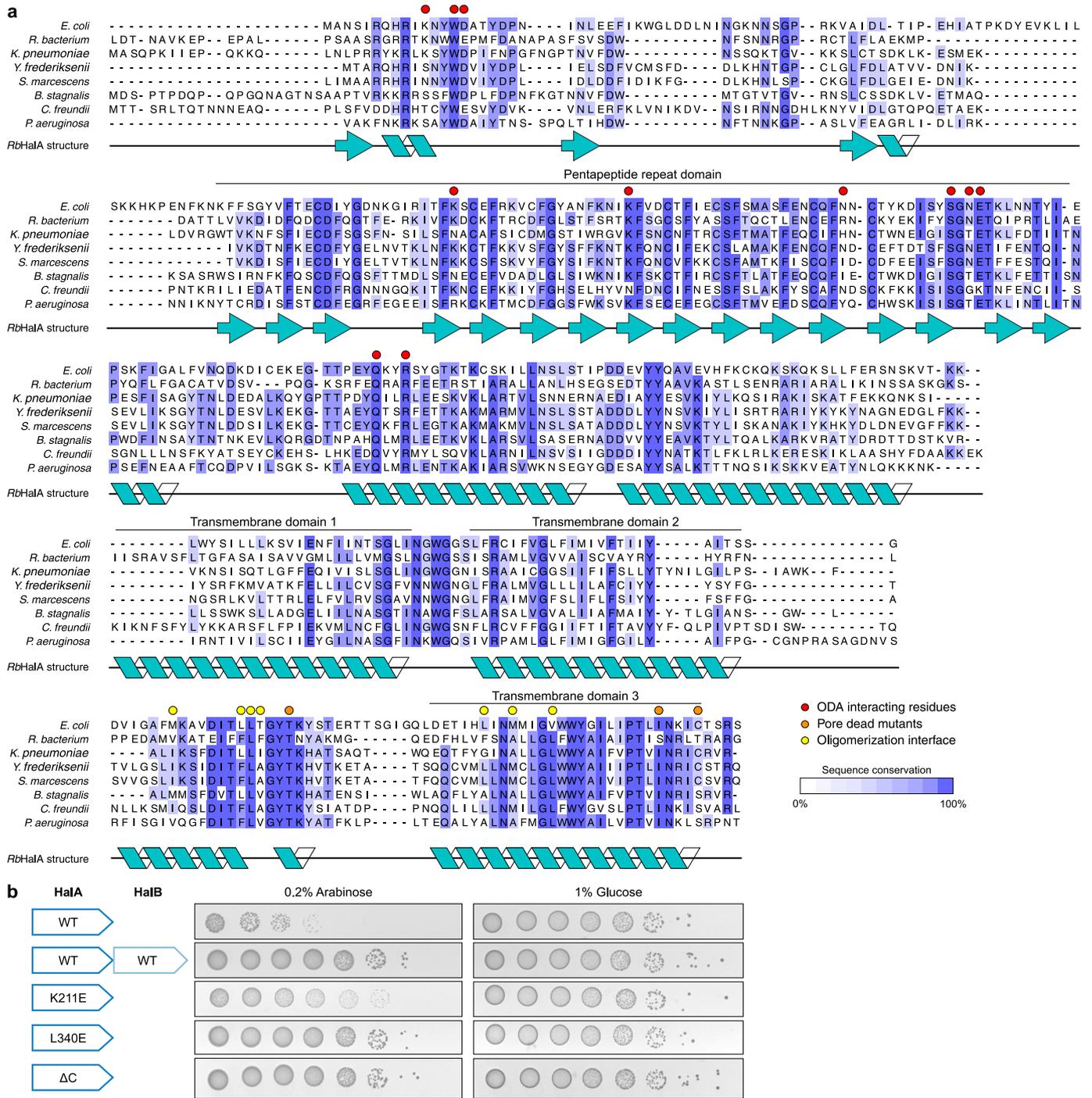
class averages from the particle curation stage. **d**, Gold-standard Fourier shell correlation (GSFSC) curves after FSC-mask auto-tightening, as produced by CryoSPARC. The horizontal black line represents the FSC = 0.143 threshold. **e**, Local resolution of the final map (unsharpened and without density modification).



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

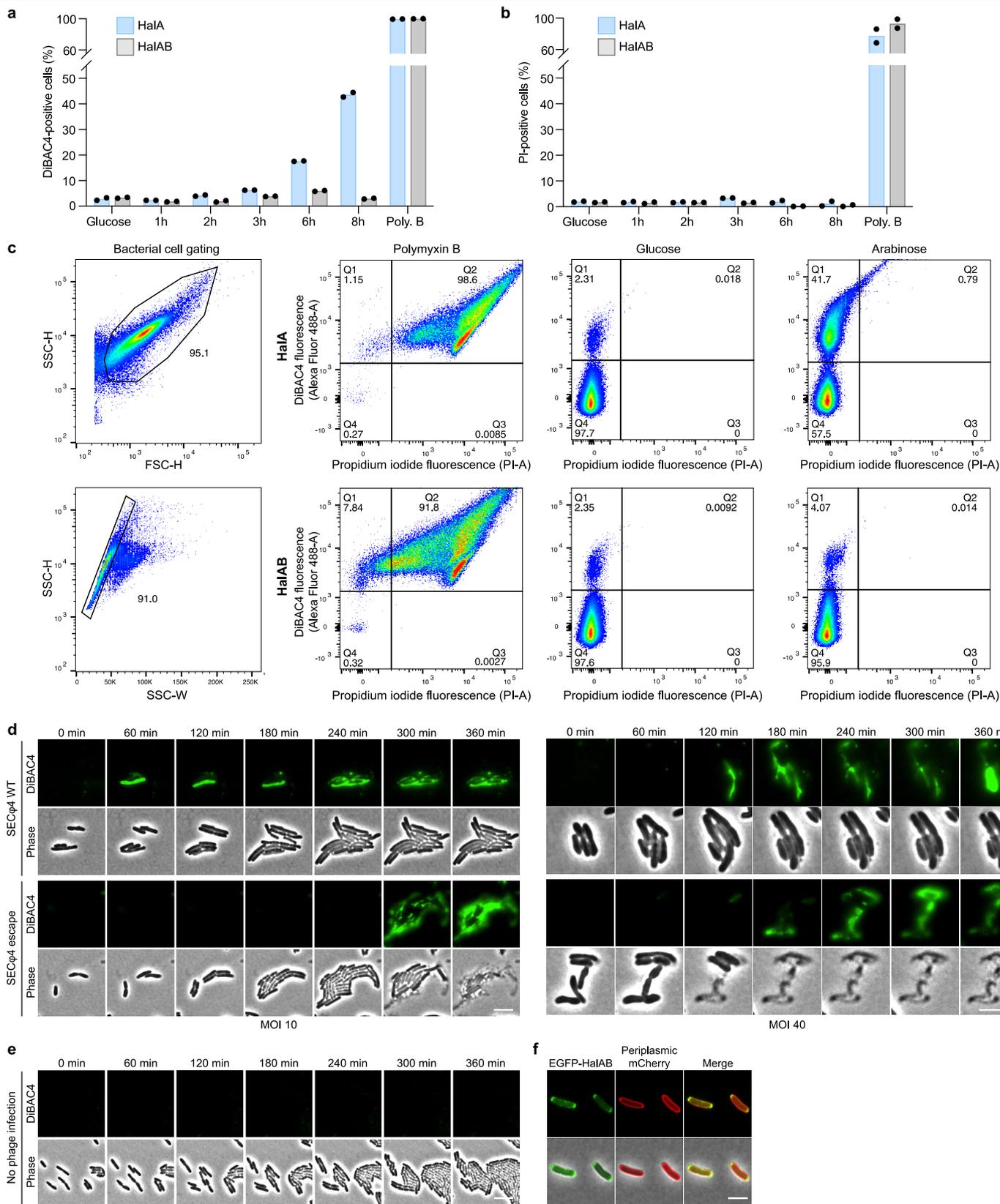
Extended Data Fig. 5 | Structural analysis of HalA. **a**, Isolated HalA protomer showing regions involved in membrane interaction and formation of the ion channel (transmembrane domain $\alpha 6$ - $\alpha 10$) and regions involved in ODA binding (pentapeptide repeat domain). **b**, Isolated HalA tetrameric ion channel region used for structural homolog comparison. **c,d**, Protein structural homology of HalA transmembrane domain against all entries in PDB showing the DALI Z score and FoldSeek E value of the top 130 and 1200 hits respectively. Entries with annotated ion channel activity are highlighted. **e**, HalA comparison to structurally related and well-characterized ion channel proteins, and corresponding PDB entry IDs are in parentheses. 2TM domain of HalA and ion channel proteins were used to highlight their structural similarities within the ion selectivity filter and transmembrane regions. Highlighted 2TM domains

are shown as individual proteins next to the full structure, or overlaid altogether with HalA. **f**, Detailed view of HalA interacting residues required for oligomerization. **g**, Detailed view of the ion channel comparing the closed state (cryo-EM structure) and AlphaFold model of HalA. The AlphaFold model shows a wider conformation of the ion conduction pathway, suggestive of conformational rearrangements that would lead to an open state in the absence of ODA. **h**, Detailed view of HalA residues involved in recognition of ODA binding specificity. Sequence-specific interactions occur between HalA residues T83, T104, and T124 with base dA5; S141 and N143 with base dA3; and side-chains W26, K109, and F139 facilitate additional interactions that control selective ODA recognition and restrict guanine and pyrimidine base recognition.



Extended Data Fig. 6 | HalA sequence analysis and growth assay. a, Structure guided multiple sequence alignment of HalA protein homologs from indicated bacterial species. Shading indicates degree of residue conservation. **b**, Bacterial growth assay of *E. coli* expressing HalA mutants required for channel assembly

with and without HalB. Expression of HalA used in this figure was from *R. bacterium* QY30. Data shown are representative of at least three independent experiments.

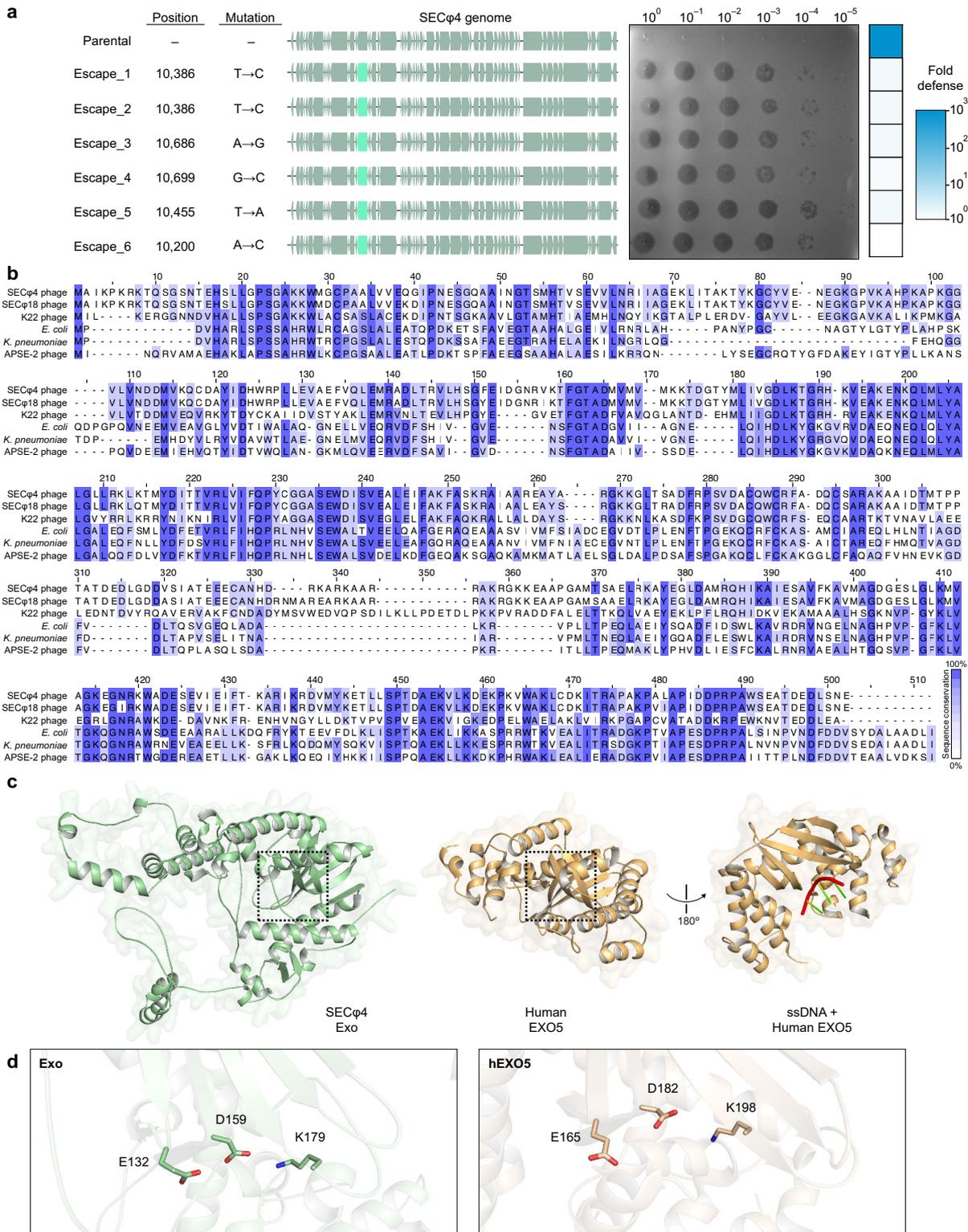


Extended Data Fig. 7 | See next page for caption.

Article

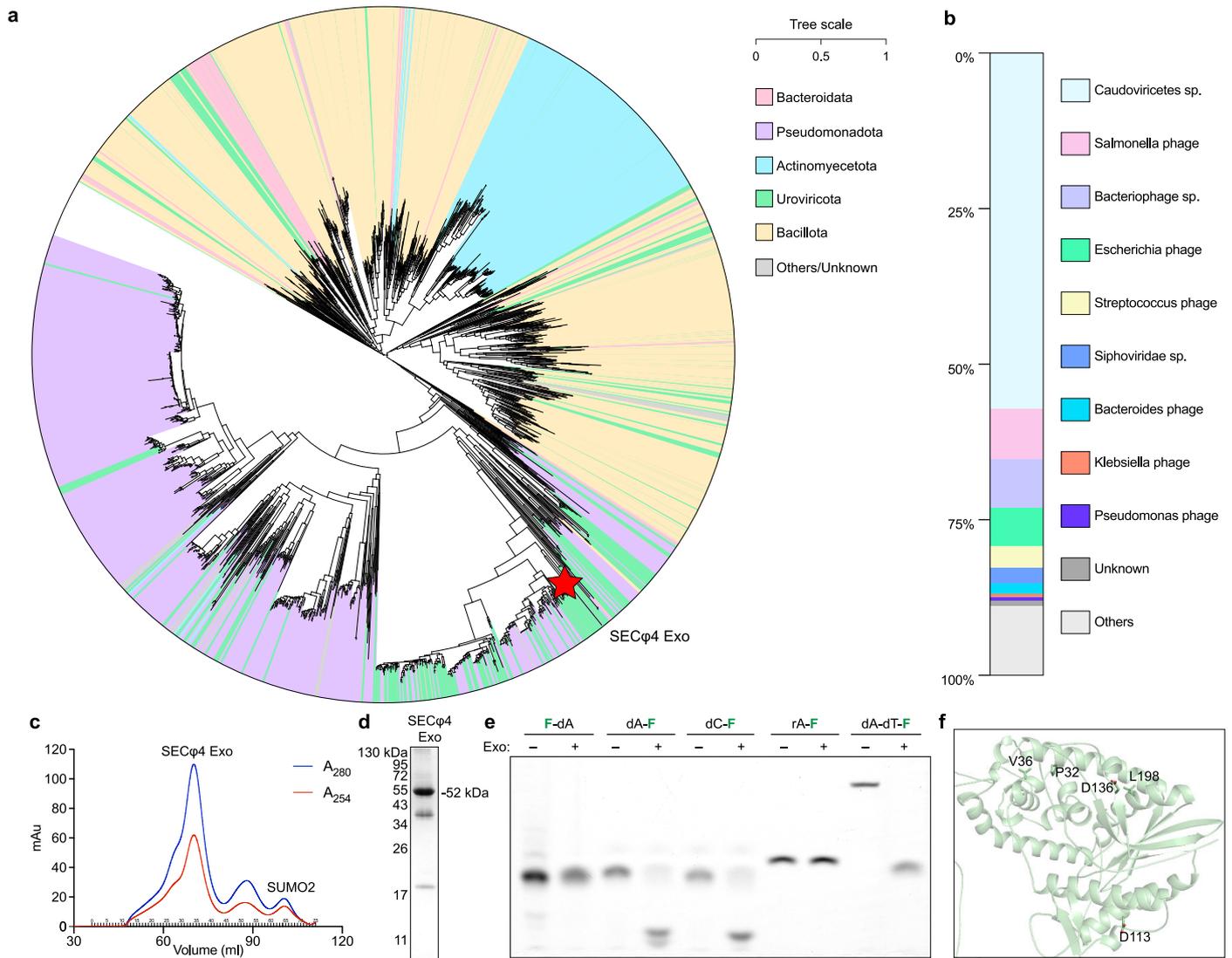
Extended Data Fig. 7 | Flow cytometry and live cell imaging analysis of HalA activation. a,b, Flow cytometry quantification of DiBAC4-positive cells (**a**) and PI-positive cells (**b**) from *E. coli* expressing HalA or HalAB treated with glucose, polymyxin B, or induced with arabinose for the indicated times. **c,** Left, gating strategy: bacterial cells were selected using side scatter height versus forward scatter height (SSC-H vs. FSC-H) and side scatter height versus width (SSC-H vs. SSC-W). Right, representative plots from cells treated with polymyxin B, or cells expressing arabinose-inducible plasmids containing HalA or HalAB. Flow cytometry data shown are collected from 100,000 events and are representative

of at least two independent experiments. **d,e,** Live cell imaging analysis of membrane depolarization using DiBAC4 in *E. coli* containing plasmids expressing Hailong defence system and infected with (**d**) WT or escape mutant phage SEC ϕ 4 at a calculated multiplicity of infection of 10 or 40, or (**e**) no phage infection. Scale bars, 5 μ m. **f,** Cellular localization of HalA visualized in *E. coli* expressing EGFP-HalA and periplasmic-mCherry. Scale bars, 1 μ m. Expression of HalA and HalAB used in this figure was from *R. bacterium* QY30. Live cell imaging data shown are representative of at least three independent experiments.



Extended Data Fig. 8 | Identification and structural characterization of Hailong phage escape mutants. **a**, Full analysis of isolated phage escape mutants. Left, sequenced genes containing the indicated point mutations are highlighted in light green within the SECφ4 genome. Right, representative plaque assays and heatmap illustrating fold defence of *E. coli* expressing Hailong from *E. coli* STEC1178 and challenged with WT SECφ4 phage and SECφ4

escape mutant phages. **b**, Multiple sequence alignment of SECφ4 escape mutant gp43 proteins from indicated phage and bacterial homologs. Shading indicates degree of residue conservation. **c**, AlphaFold modeled structure of SECφ4 Exo (left) and comparison with human EXO5 (right) with (PDB ID: 7LW9) and without DNA (PDB ID: 7LW7). Boxes highlight nuclease active site. **d**, Detailed view of nuclease active residues in SECφ4 Exo (left) and human EXO5 (right).



Extended Data Fig. 9 | Phylogenetic and biochemical analysis of SECφ4 Exo. **a**, Phylogenetic analysis of ~1,700 SECφ4 Exo sequence homologs obtained using NCBI BLAST. **b**, Genera of phages encoding SECφ4 Exo. **c**, SECφ4 gp43 DNA exonuclease (Exo) was expressed as an N-terminal 6xHis-SUMO2 fusion and purified by Ni-NTA and separated by size exclusion chromatography. **d**, Coomassie-stained SDS-PAGE analysis of fully purified SECφ4 Exo. **e**, Analysis of 5 nt fluorescein-labeled deoxynucleotide substrates (F-dA, dA-F, dC-F),

5 nt single-stranded RNA (rA-F), or 20 bp thymidine-labeled double-stranded DNA (dA-dT-F). Direction of fluorescein (labeled as green F) tagged to the oligonucleotide indicates either a 5'-tagged DNA (indicating a free 3' end) or a 3'-tagged DNA (indicating a free 5' end). Data shown are representative of at least three independent experiments. **f**, AlphaFold modeled structure of SECφ4 Exo with escape mutant residues within scaffolding regions.

Reporting Summary

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

Statistics

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

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

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

Software and code

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

Data collection	No software was used for data collection
Data analysis	Coot 0.8.9.3, Phenix 1.21, PyMol 2.5.5, SBGrid Package (cryoSPARC 4.1.1, CryoSieve v1.2.5, RELION v4.0.1, resolve_cryo_em v2.15), FlowJo 10, Nikon Elements 5.22, Fiji 2.9.0

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](#)

Data supporting the findings in this study are available within the article or in the Extended Data. Protein Data Bank (PDB) accessions are listed when relevant and the following accessions were used: PDB 9DBH, PDB 9DBI, PDB 9DBJ. Coordinates and density maps of the HaIA-ODA complex have been deposited in PDB and EMDB under the accession codes PDB 9NYI and EMD-49920.

Research involving human participants, their data, or biological material

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

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

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

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	No sample size calculations were performed for this study as this is typically not applicable for biochemical, flow cytometry, microscopy, and bacterial and phage growth experiments. Sample sizes were chosen by established field standards for each assay: Biochemical (Hobbs et al, 2022), microscopy and flow cytometry (Duncan-Lowe et al, 2021), bacterial and phage growth (Stokar-Avihail et al, 2023).
Data exclusions	No data were excluded from the analyses.
Replication	All data were performed with replicates as described and are consistent with commonly used field standards. Biochemical or growth-based assays were performed with 2–3 technical replicates and repeated on at least 3 independent biological replicates. Figures show one representative example of a biologically independent experiment and other experiments not shown yielded similar results within expected variation.
Randomization	No randomization was performed as it does not impact the interpretation of the data collected.
Blinding	Data were not blinded as blinding is not relevant to the experiments described in this work since there were no subjective analyses performed which would impact the interpretation of the data.

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

Methods

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

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

E. coli Top10 cells containing arabinose-inducible plasmids expressing HalA or HalA and HalB were grown in LB culture at 37 °C for 16 h. Overnight starter culture was diluted 1:10 in LB cultures supplemented with 1% glucose or 0.5% arabinose and incubated for 1, 2, 3, 6, or 8 h at 37 °C. Polymyxin B treatment at 37 °C for 30 min was used as a positive control. Following incubation, bacterial suspensions were treated with DiBAC4 and propidium iodide in PBS for 5 min, and immediately run on the flow cytometer without further processing.

Instrument

Cells were analyzed using a BD FACSymphony A1 Cell Analyzer.

Software

FACSDiva v8.0 software was used for data acquisition, and data were analyzed using FlowJo v10 software.

Cell population abundance

Samples were analyzed by flow cytometry but no cell sorting was performed.

Gating strategy

Preliminary FSC/SSC gates were applied followed by removal of doublets (FSC-H vs SSC-H and also SSC-W vs SSC-H). Cells treated with polymyxin B were used as a positive control for gating of DiBAC4- and propidium iodide-positive cells, and cells treated with glucose and cells expressing both HalA and HalB were used as negative controls for gating of DiBAC4- and propidium iodide-negative cells.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.