nature structural & molecular biology

Article

https://doi.org/10.1038/s41594-026-01754-2

Structure and mechanism of antiphage retron Eco2

Received: 5 August 2025

Accepted: 21 January 2026

Published online: 18 February 2026

Check for updates

M. Jasnauskaitė[1], J. Juozapaitis[2], T. Liegutė[2], R. Grigaitis[1], A. Skorupskaitė[1], W. Steinchen[3], A. Mikšys[1,8], L. Truncaitė[4], K. Kazlauskaitė[1], M. F. Torres Jiménez[5], S. Khochare[1], G. Dudas[2], G. Bange[3,6], L. Malinauskaitė[1,9], I. Songailienė[2,7] & P. Pausch[1]

Retrons are prokaryotic reverse transcriptase systems that produce multicopy single-stranded DNA (msDNA), yet the principles by which they mediate antiviral defense remain largely unresolved. Here we investigate the mechanism of *Escherichia coli* Eco2, a minimal retron composed of a single reverse transcriptase–nuclease fusion protein. Cryogenic electron microscopy and hydrogen/deuterium exchange mass spectrometry reveal the structures and dynamics of a trimeric nucleoprotein complex assembled within a branched msDNA scaffold, which cages the TOPRIM nucleases. We show that the phage-encoded endonuclease DenB initiates msDNA degradation, thereby unblocking the nuclease active sites. Activated Eco2 cuts transfer RNAs, resulting in translational shutdown for antiphage defense. We further identify ribosomal protein S1 as a putative RNA chaperone that associates with the msDNA precursor. These findings provide insights into the molecular mechanisms of minimal retrons and establish a structural basis for engineering of Eco2.

Reverse transcriptases (RTs), which synthesize DNA from RNA templates, are found across all domains of life. In prokaryotes, RTs are typically associated with mobile genetic elements, either promoting or interfering with their propagation[1]. The ongoing conflict between viruses and their hosts has driven the evolution of diverse reverse-transcriptase-based immune systems, including retrons, abortive phage infection systems, CRISPR–Cas RTs and unknown group systems[2–4].

Although retrons were first described in the 1980s[5–7], their role in antiphage defense was only recently established[8–10]. Retrons produce multicopy single-stranded DNA (msDNA), an activity discovered in *Myxococcus xanthus*, in which msDNA molecules were found at levels up to ~700 copies per cell[6]. Later studies revealed that the msDNA was covalently linked to a noncoding RNA (ncRNA) via a unique 2′–5′ phosphodiester bond[11,12]. In addition to the *rt* and *ncRNA* genes, retron loci encode various accessory genes, which initially remained uncharacterized[5,13]. These accessory genes were later shown to encode diverse effector proteins that mediate antiphage immunity[8–10], likely activated in response to phage-encoded DNA-interacting factors such as nucleases, methyltransferases and single-stranded DNA binding proteins[10,14,15].

Retron systems are highly diverse and have been classified into types I–XIII and several subtypes based on reverse transcriptase phylogeny and associated effector identities[16]. Structural data are available for retron Eco1 (type II-A) from *Escherichia coli*, which forms a filamentous complex composed of multiple RT, msDNA

[1]LSC-EMBL Partnership Institute for Genome Editing Technologies, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [2]Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [3]Center for Synthetic Microbiology (SYNMIKRO) & Department of Chemistry, Philipps-University Marburg, Marburg, Germany. [4]Institute of Biochemistry, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [5]Institute of Biosciences, Life Sciences Center, Vilnius University, Vilnius, Lithuania. [6]Max-Planck-Institute for Terrestrial Microbiology, Molecular Physiology of Microbes, Marburg, Germany. [7]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [8]Present address: ATEM Structural Discovery GmbH, Remscheid, Germany. [9]Present address: BioNTech UK Ltd, MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK. e-mail: patrick.pausch@gmc.vu.lt

https://doi.org/10.1038/s41594-026-01754-2

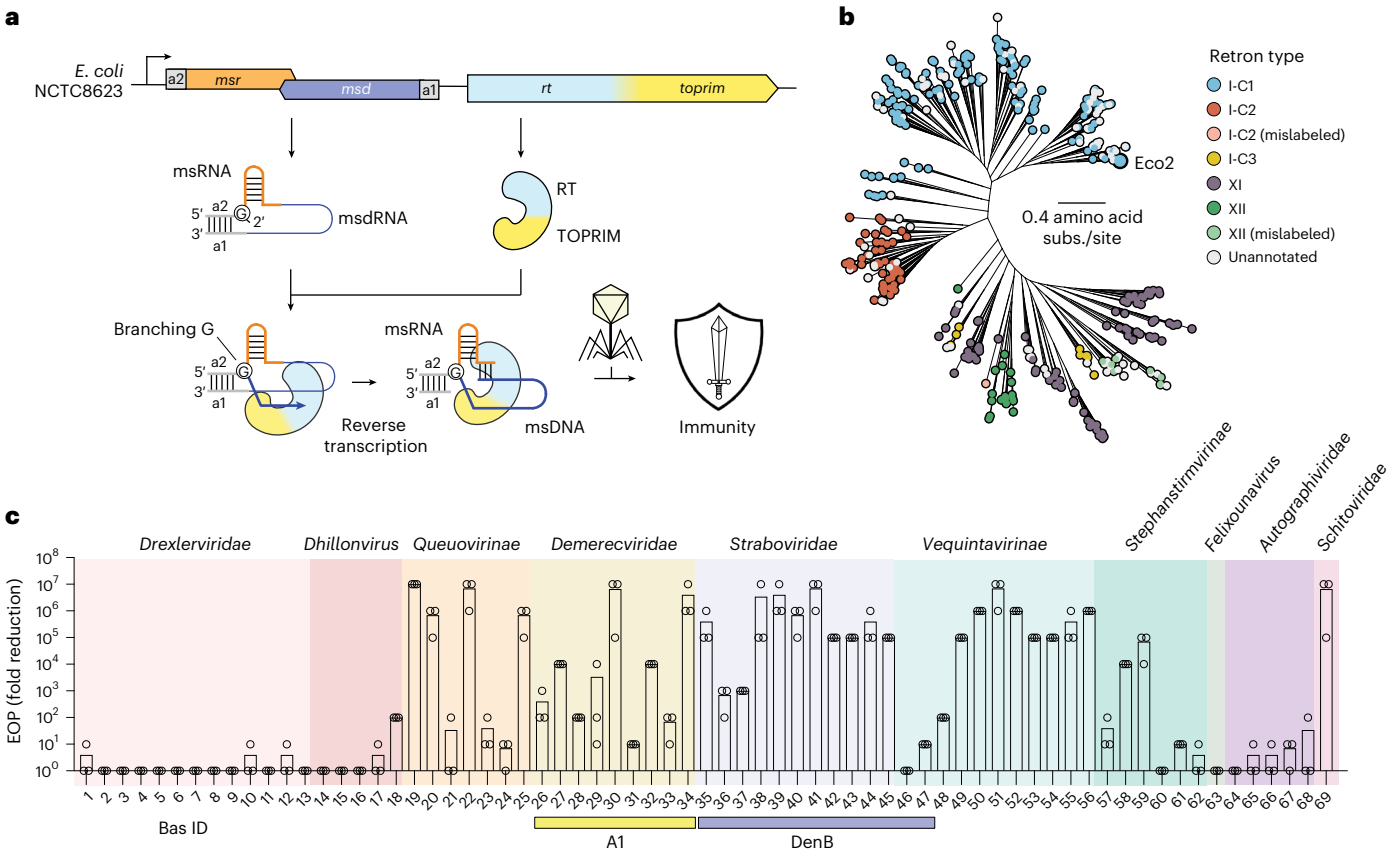


Fig. 1 | Eco2 provides immunity against diverse phages. a, Scheme illustrating the *eco2* locus and antiphage defense function. **b**, Phylogenetic tree showing RT relationships among types I-C, XI and XII retrons. Subs., substitutions. **c**, Effectiveness of Eco2 against BASEL phage collection, measured as efficiency of plating (EOP). $n = 3$ technical replicates; mean. Phages encoding A1 and DenB homologs are indicated by yellow and blue bars, respectively.

and *N*-glycosidase effector subunits[17–19]. Upon msDNA methylation by phage-encoded DNA cytosine methyltransferase, the effector is activated for NAD(P)$^+$ hydrolysis[19]. Recent studies of retron type I-A have described a high-molecular-weight defense complex consisting of the RT, msDNA, two ATPase dimers and an HNH nuclease[20–23]. To defend against phages, the subunits may sense phage-encoded exonucleases to release the HNH effector, leading either to transfer RNA (tRNA) degradation[22], single-stranded DNA cleavage[23] or both[21]. Consistent with msDNA-dependent regulation of retrons, Eco8 (type I-B2) was shown to form a 4:4:4 (RT/msDNA/overcoming lysogenization defect (OLD)) complex, which activates the OLD DNA nuclease upon msDNA displacement by phage-encoded single-stranded DNA binding proteins[24–26].

Type I-C retrons are common, with ~11% of retrons belonging to this type[8], and are often encoded on prophages[5]. Unlike most other types, type I-C loci consist of a single protein-coding gene downstream of the *ncRNA* gene[16] (Fig. 1a). This single protein couples an amino-terminal RT domain fused to a carboxy-terminal RNase-H-like topoisomerase-primase (TOPRIM) domain, an architecture reminiscent of retroviral RT–RNase H fusions, such as those found in MMLV and HIV-1 RTs[27,28]. Type I-C retrons are further subclassified into I-C1 to I-C3 variants according to their phylogeny[16].

As demonstrated for the type I-C1 retron Eco2 from *E. coli* (alias Ec67)[29], both the RT and TOPRIM active sites are essential for the defense activity of Eco2 in response to phage infection[8,9]. The ncRNA of Eco2 comprises multicopy single-stranded RNA (msRNA) and msDNA RNA (msdRNA) segments, flanked by complementary sequences a1 and a2, with the msdRNA serving as the template for msDNA synthesis[29,30] (Fig. 1a). msDNA synthesis by the RT is primed at the 2′-OH of branching nucleotide G15 and terminates at A67[29]. It has been speculated that the msdRNA template is hydrolyzed by the RNase-H-like TOPRIM during maturation of the msDNA[30]. Purification of Eco2 revealed large nucleoprotein complexes composed of the RT–TOPRIM protein bound to msDNA[30], and a recent study reported a 3:3 (RT–TOPRIM/msDNA) oligomeric structure and RNase activity of Eco2[31]. Isolation of phage escapers led to identification of mutations in nuclease-encoding genes, implicating phage nucleases A1 and DenB as triggers of Eco2[14,15,31]. However, the activity of Eco2 that leads to defense and the molecular details and regulation of this single protein retron system have remained unknown.

In this study, we explored the molecular mechanism of Eco2. Screening against a library of diverse phages showed that Eco2 protects across multiple phage families. We identified T2 phage-encoded nuclease DenB as an activator of Eco2. Cryo-electron microscopy (cryo-EM) revealed an intricate trimeric nucleoprotein structure and regulation of Eco2, and trapping of assembly intermediates resulted in identification of ribosomal protein S1 as a putative ncRNA chaperone. Hydrogen/deuterium exchange mass spectrometry (HDX-MS) highlighted pronounced msDNA-dependent structural dynamics of Eco2. Finally, we showed that DenB-mediated msDNA decay activates tRNA cleavage by Eco2, effectively halting protein synthesis for antiphage defense. Collectively, these findings provide mechanistic insights into the biogenesis, enzymatic regulation and antiphage defense of the single protein retron system Eco2.

Eco2 provides broad defense against phages

Immune systems of bacteria defend against phages either by clearing the cell of the virus or by inducing cell death upon infection, thereby limiting viral spread[32,33]. At the molecular level, these systems first detect phage-specific proteins or nucleic acids, before triggering an

effector-mediated immune response. Notably, phylogenetic analysis of type I-C retrons showed that the current classification may not accurately reflect the biological complexity of their relationship and suggested that some type I-C sequences are mislabeled as type XII[16] (Fig. 1b and Supplementary Fig. 1). We therefore focused our investigation on the distinct type I-C1 retron Eco2 from *E. coli* NCTC8623 (National Collection of Type Cultures 8623)[9].

Eco2 has previously been shown to protect *E. coli* against phages T2, T4, T5 and SP15[8,9,14,34]. In support of the defense activity described for *E. coli* during T5 infection[8], we observed that Eco2-expressing cultures continued growth at an estimated multiplicity of infection (MOI) of 0.1 and 1, or entered stasis at an MOI of 10, following infection with phages T2 and T5 (Supplementary Fig. 2). Notably, Eco2-expressing cultures grew more slowly even in the absence of infection compared to an empty vector control, suggesting that Eco2 expression may be toxic to the cells (Supplementary Fig. 2).

To systematically explore the defense specificity of Eco2, we challenged Eco2-expressing *E. coli* with a panel of diverse phages from the BASEL collection[35]. This screen revealed broad defense against phages from the families *Demerecviridae*, *Straboviridae*, *Queuovirinae*, *Vequintavirinae*, *Stephanstirmvirinae* and *Schitoviridae* but not against representatives of *Drexlerviridae* or *Autographiviridae*, as well as most *Dhillonvirus* and *Felixounavirus* species (Fig. 1c and Supplementary Fig. 3).

Phage T2 escaped Eco2-mediated defense via mutations in the *denB* gene, which encodes a dC-specific DNA endonuclease[36,37] (Extended Data Fig. 1). All three escape mutations occurred in conserved residues near the active site (G71R) or the DNA binding interface (G180E and W181*) (Extended Data Fig. 1b–d), suggesting that T2 evades Eco2 through inactivation of DenB. This observation aligns with recent reports, which additionally identified escaper mutants within the A1-encoding gene (an early essential gene involved in second-step DNA transfer[38]) from phages T5 and ΦSP15m[14,15]. Both DenB and A1 have been implicated in host genome degradation during phage replication[15,39–41] and are conserved across *Straboviridae* and *Demerecviridae*, respectively (Fig. 1c). Although structurally unrelated to DenB, escaper mutants in A1 were found to occur in the predicted C-terminal endonuclease domain or to be enriched in the N-terminal DNA binding domain, suggesting that evasion also depends on inactivation[14].

We could not identify homologs of DenB and A1 in *Queuovirinae*, *Stephanstirmvirinae*, *Schitoviridae* or most *Vequintavirinae* (Fig. 1c), suggesting that Eco2 responds to other factors encoded by these phages. Moreover, despite carrying *denB* or *a1* genes, phages Bas31, Bas46 and Bas47 were not susceptible to Eco2-mediated defense (Fig. 1c). This could indicate that these DenB or A1 homologs do not activate Eco2, or it could point to the presence of phage-encoded anti-Eco2 factors that inhibit the immune response.

Collectively, these findings demonstrate that Eco2 enables a broad and robust defense against a variety of phages across different families and provide support for the idea that Eco2 is triggered by DNA endonucleases that are central to phage replication.

Eco2 forms a unique trimeric nucleoprotein complex

To elucidate the mechanism of Eco2, we purified the msDNA-bound nucleoprotein complex and determined cryo-EM structures in the absence and presence of the catalytic Mg^{2+} cofactor, at average resolutions of ~2.7 Å and ~3.1 Å, respectively (Fig. 2a, Table 1 and Extended Data Figs. 2–4). This revealed a unique, shuriken-like C3-symmetric nucleoprotein complex composed of three RT–TOPRIM subunits interconnected by their msDNA products (Fig. 2a), which compared well in the presence and absence of Mg^{2+} (Extended Data Fig. 2). We locally refined a single subunit of Eco2 in the presence of Mg^{2+}, which resulted in an improved map for one of the TOPRIM domains (chain A) but a degraded map for the other two (Extended Data Fig. 4). Overall, our structures resembled a recently reported trimeric structure of Eco2[31] (Supplementary Fig. 4a,b), with a few discrepancies in the msDNA and TOPRIM active sites, as detailed below.

Importantly, the architecture of Eco2 was markedly different from those of retron types II-A, I-A and I-B2, with structural similarity limited to the conserved RT[17–19,21–26] (Supplementary Fig. 5). Each Eco2 RT–TOPRIM subunit adopted a claw-like fold, in which the RT thumb subdomain bridged the juxtaposed RT and TOPRIM domains (Fig. 2b). Between the domains, the msDNA–msRNA hybrid was tightly aligned to the RT thumb (Fig. 2b).

Inspection of the palm subdomain RT active site YADD motif (residues 199–202) revealed the bound magnesium ion cofactor unambiguously (Fig. 2c and Extended Data Fig. 2d). It also showed that msDNA product termination occurs at base pair G52:dC67, through flipping of the adjacent msRNA base U51, likely mediated by the finger subdomain β-hairpin and upstream msRNA (Fig. 2c).

We next examined the trimer contacts within the complex. This showed that the RT–TOPRIM monomers were linked in *trans* via msDNA-mediated interactions: the msRNA hairpin connects the palms of adjacent RT domains, whereas the msDNA loops around the fingers and extends across the TOPRIM domains toward the adjacent RT (Fig. 3a). Contrary to the proposed stem–loop secondary structure of the msDNA[29], the individual strands (nucleotides 29–50) did not form hairpins. Instead, they appeared to hybridize at the center of the TOPRIM domain triangle, forming a DNA three-way junction (3WJ) (Fig. 3b). These observations are in contrast to the previously determined Eco2–msDNA structure[31], in which no 3WJ was observed and the ncRNA was proposed to traverse the RT fingers in place of the msDNA (Supplementary Fig. 4c).

The only protein–protein interface within the complex was located adjacent to the 3WJ, where RT residues 325–347 of each monomer interconnected the three subunits (Fig. 3c). In proximity, a short single-stranded segment of the msDNA (nucleotides 25–27) interacted with α-helices 14 and 16 through stacking and polar interactions (Fig. 3d). The low local resolution around the 3WJ and adjacent msDNA suggested conformational flexibility, complicating unambiguous assignment of nucleotide identities and strand conformations (Extended Data Figs. 4 and 5).

Finally, comparison of the Mg^{2+}-bound TOPRIM active site in our experimental structure with a predicted cofactor-bound model of the RT–TOPRIM revealed that active site residue D462 adopted a conformation incompatible with binding of the catalytic Mg^{2+}_A cofactor (Fig. 3e,f and Supplementary Fig. 6). The conformation of D462 differed considerably from that in the previously reported Eco2 structure[31] (Supplementary Fig. 4d).

Together, these findings suggest that the TOPRIM resides in an inactive state. Thus, the msDNA not only mediates complex interactions but appears to also cage the TOPRIM domain, possibly functioning as a regulator to the effector for control over the activity of Eco2.

msDNA production and binding is a prerequisite for trimerization

Deactivation of either the RT or TOPRIM active site (dRT or dTOPRIM) rendered Eco2 incapable of defense against phages[9] (Fig. 4a). To explore the role of both active sites in the biogenesis of Eco2, we purified dRT (D201A/D202A) and dTOPRIM (E374A/D378A/D460A/D462A) variants (Fig. 4b,c and Supplementary Fig. 7). Notably, the msDNA product and trimeric state of the dTOPRIM variant were comparable to those of the wild type (WT) (Fig. 4b), suggesting that the RNase-H-like TOPRIM domain is not involved in msDNA maturation by template hydrolysis, as previously speculated[30].

By contrast, purification of the dRT variant primarily yielded ncRNA–dRT monomers (Fig. 4c,d), indicating that msDNA formation is a prerequisite for trimerization. To investigate the conformational changes of the RT–TOPRIM upon msDNA formation, we compared the

https://doi.org/10.1038/s41594-026-01754-2

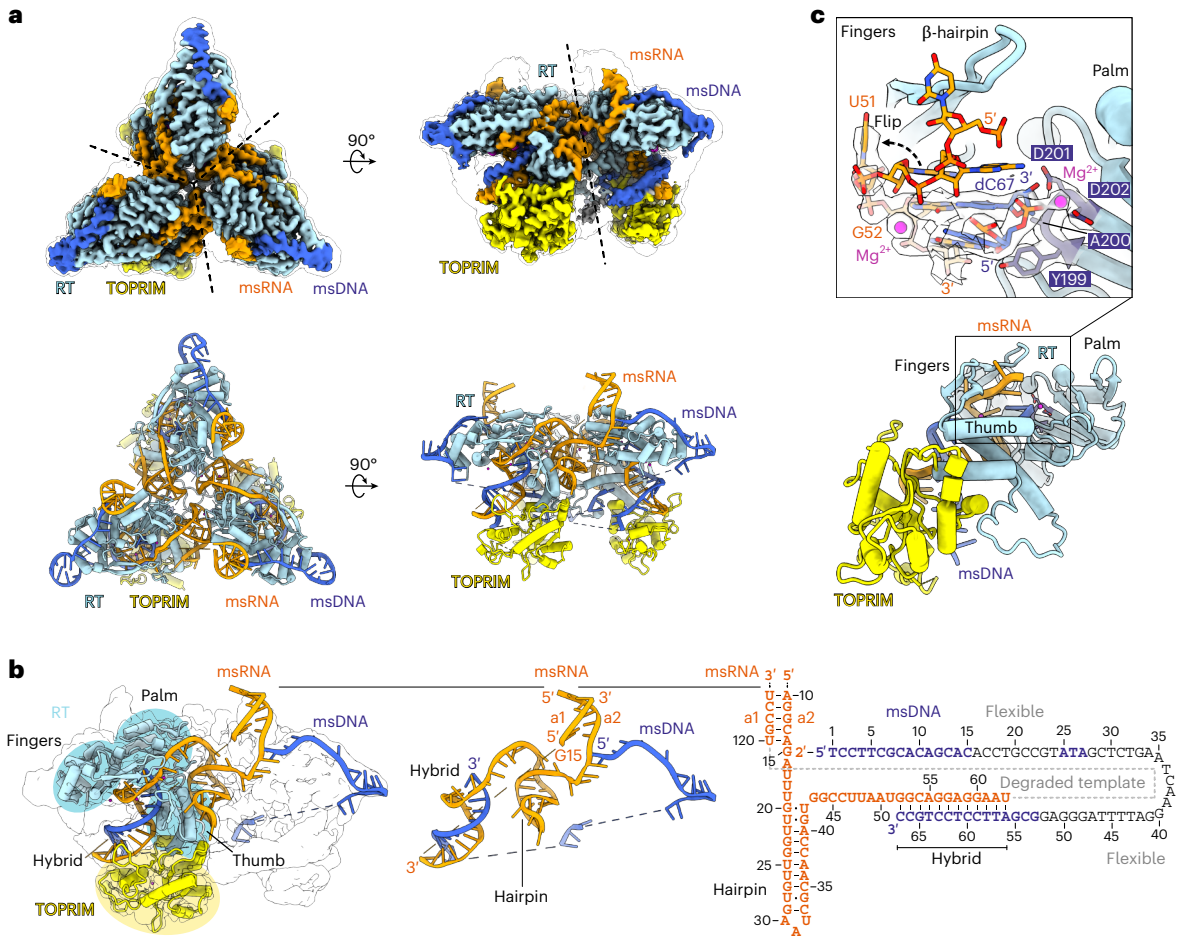


Fig. 2 | Eco2 forms a trimer. a, Cryo-EM maps (above) and corresponding model (below) of magnesium-bound Eco2, shown in two 90° rotations. Coloring follows the color scheme used in Fig. 1a. The unfiltered map is shown as a translucent surface. **b**, Left: one subunit of msDNA (orange and blue) and RT–TOPRIM (light blue and yellow cartoon) in the unfiltered cryo-EM map; center: msDNA model; right: msDNA scheme. **c**, Top: close-up of the RT active site. The sharpened map is shown as a translucent surface. Bottom: RT–TOPRIM overview.

nucleic-acid-free apo and msDNA-bound states using HDX-MS (Fig. 4e and Extended Data Fig. 6). This revealed a pronounced reduction in HDX across the RT, including palm, finger and thumb subdomains, in addition to the subunit-interconnecting trimerization loop (residues 325–347), when msDNA was present (Fig. 4e and Extended Data Fig. 6). Unlike the RT, the TOPRIM domain showed only minor HDX changes (Fig. 4e and Extended Data Fig. 6). This showed that primarily the RT, not the TOPRIM, becomes stabilized upon msDNA binding.

To understand the roles of the features of the RT that mediate trimer contacts and msDNA binding for the defense activity of Eco2, we introduced alanine substitutions in residues from the subunit-interconnecting loop and adjacent helices. Alanine substitution of K318 and K325/K328 in α-helix 14, and Y421 in α-helix 12, impaired Eco2-mediated defense against phage T2 in vivo, whereas substitutions of loop residues Y333 and V334 had no effect (Figs. 3c,d and 4f). In small-scale purification of these variants, no K318A, K325A/K328A or Y421A variants were recovered after elution, suggesting poor expression or low solubility, possibly due to toxicity or instability (Supplementary Fig. 8).

Collectively, these findings show that the RT becomes stabilized upon msDNA binding, which may facilitate complex formation and regulation of the activity of Eco2.

Ribosomal protein S1 associates with the msdRNA

The ncRNA–dRT was also copurified as a heterodimer with a protein of ~70 kDa (Fig. 4c and Supplementary Fig. 7). Mass spectrometry identified this protein as ribosomal protein S1 (Supplementary Fig. 9), a multifunctional protein involved in translation initiation and mRNA chaperoning[42–45].

Urea–PAGE analysis and sequencing of the ncRNA bound to the dRT monomer and dRT–S1 heterodimer suggested that S1 bound to the msdRNA part of the ncRNA, with this region exhibiting higher sequence coverage relative to the msRNA segment in the dRT–S1 fraction (Fig. 4c,g). Attempts to determine the cryo-EM structure of the dRT–S1 complex were hindered by its structural flexibility, prompting us to probe the interaction in vivo. Bacterial two-hybrid analysis showed no interaction between WT, apo or dRT variants of Eco2 and either full length S1 or its truncation constructs (Supplementary Fig. 10). In vitro reconstitution of the RT–TOPRIM–ncRNA–S1 complex revealed that S1 bound only in the presence of the ncRNA (Extended Data Fig. 7).

These findings demonstrate that S1 indirectly interacts with the RT–TOPRIM protein via the msdRNA, potentially acting as an RNA chaperone that stabilizes the msDNA template before reverse transcription in vivo.

DenB initiates msDNA decay to activate Eco2 TOPRIM

Our structures and HDX-MS analysis lent support to the idea that Eco2 is regulated by the msDNA. Although the substrate identity of Eco2 was unknown, the RNase-H-like TOPRIM domain suggested a metal-ion-dependent nuclease activity[46,47]. In support of this, a ΔRT

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

	Eco2 in absence of magnesium (EMD-52583; PDB 9I2F)	Eco2 in presence of magnesium (EMD-52584; PDB 9I2G)	Activated Eco2 in presence of magnesium (EMD-54448; PDB 9S1F)
Data collection and processing			
Magnification	×92,000	×92,000	×92,000
Voltage (kV)	200	200	200
Electron exposure ($e^-/Å^2$)	30	30	30
Defocus range (μm)	−2.0 to −1.0	−2.0 to −1.0	−2.0 to −1.0
Pixel size (Å)	1.1	1.1	1.1
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	2,914,940	1,037,763	2,271,864
Final particle images (no.)	504,525	1,017,678	617,364
Map resolution (Å)	2.76	3.04	2.90
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.5 to 7.0	2.5 to 6.5	2.5 to 6.5
Refinement			
Initial model used (PDB code)	AlphaFold2	9I2F	9I2G
Model resolution (Å)	3.2	3.3	3.2
FSC threshold	2.7	3.1	2.9
Map sharpening *B* factor ($Å^2$)	121.6	119.2	130.3
Model composition			
Nonhydrogen atoms	19,765	19,650	18,414
Protein residues	1,738	1,734	1,710
Nucleic acids	273	268	217
Ligands	0	MG: 21	MG: 7
B factors ($Å^2$)			
Protein	146.65	126.81	136.98
Nucleic acids	162.64	140.78	125.40
Ligand		86.91	102.11
Root mean square deviations			
Bond lengths (Å)	0.003 (0)	0.003 (0)	0.003 (0)
Bond angles (°)	0.545 (0)	0.519 (1)	0.547 (0)
Validation			
MolProbity score	1.15	1.14	1.22
Clashscore	3.55	3.49	4.43
Poor rotamers (%)	0.00	0.00	0.00
Ramachandran plot			
Favored (%)	99.36	99.13	99.35
Allowed (%)	0.52	0.87	0.47
Disallowed (%)	0.12	0.00	0.18
Rama-Z score, whole (root mean square Rama-Z)	−1.37 (0.20)	−0.08 (0.20)	−1.91 (0.20)
CC (mask)	0.80	0.77	0.73
CC (box)	0.82	0.77	0.75

CC, correlation coefficient; FSC, Fourier shell correlation.

truncation of the RT–TOPRIM was recently found to cleave RNA indiscriminately[31]. Efforts to clone and express the putative activator *denB* from phage T2 in *E. coli* were hindered by its toxicity, consistent with previous reports for *denB* from phage T4[37].

We treated Eco2 with DNase I as a functional proxy for DenB, which resulted in cleavage of both the msDNA and a fluorophore-quencher reporter RNA (Fig. 5a and Supplementary Fig. 11a). Alanine substitutions of the TOPRIM catalytic residues (dTOPRIM, E374A/D378A/D460A/D462A) prevented reporter turnover (Fig. 5a), indicating that msDNA degradation activates the TOPRIM domain for RNA cleavage.

We next incubated purified Eco2 in *E. coli* cell-free extracts (CFEs) expressing *denB*. Purification and urea–PAGE analysis of DenB-activated Eco2 revealed near complete msDNA degradation (Fig. 5b and Supplementary Fig. 11b). To map the DenB cleavage pattern, we incubated purified, nonactivated Eco2 with either purified WT DenB or the G71R DenB escaper variant (Fig. 5c and Supplementary Fig. 11c,d).

https://doi.org/10.1038/s41594-026-01754-2

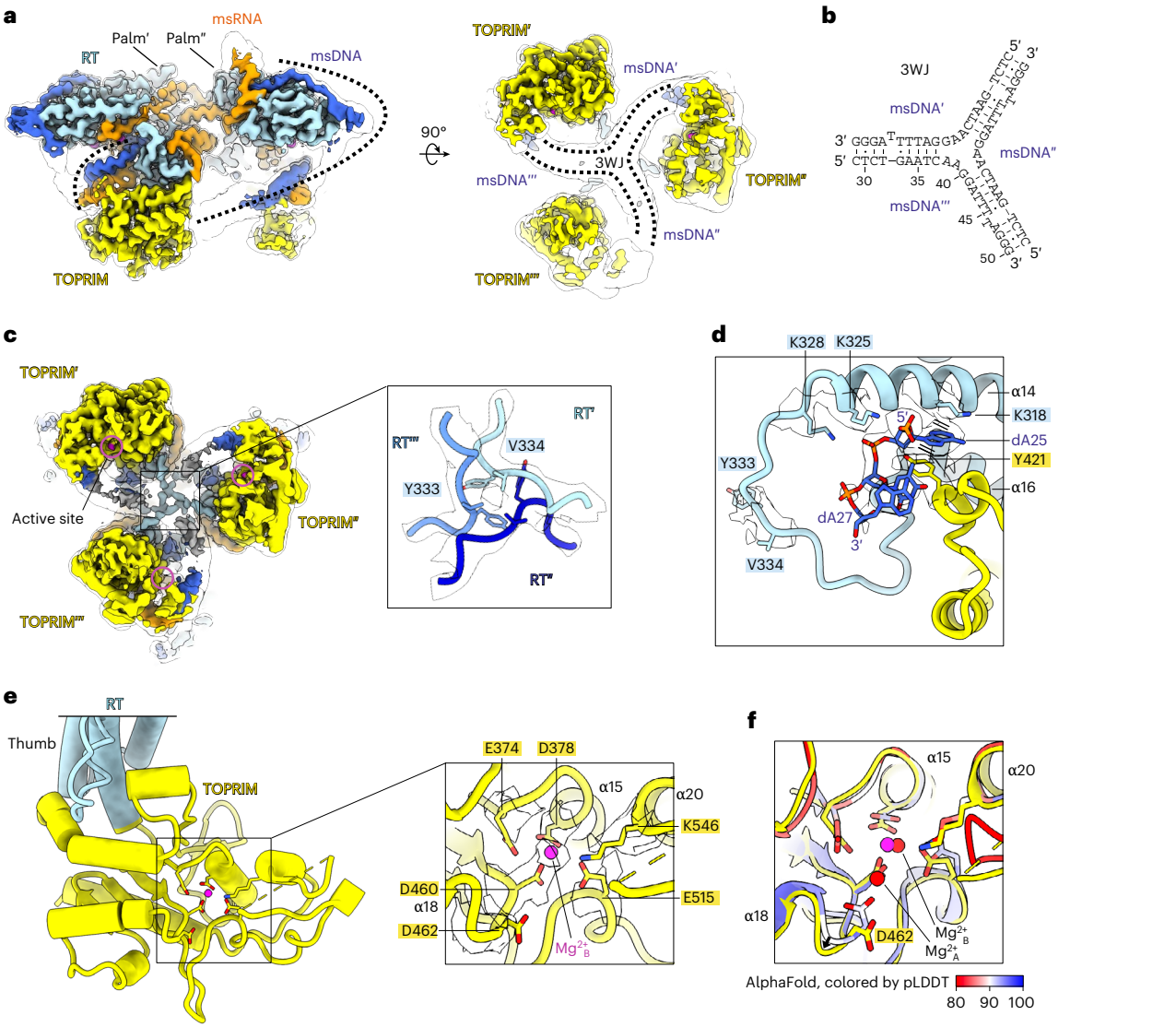


Fig. 3 | msDNA mediates complex interaction and cages TOPRIM. a, Sharpened (colored) and unfiltered (translucent surface) cryo-EM maps of Mg^{2+}-bound Eco2 in two 90° rotations. The dashed line indicates the msDNA path. **b**, Scheme of the 3WJ, illustrating predicted base-pairing. **c**, Left: bottom view of sharpened (colored) and unfiltered (translucent surface) maps of Mg^{2+}-bound Eco2; right: close-up of the trimer contact loop. **d**, Close-up of the trimer-contact-adjacent msDNA binding site. The sharpened map is shown as a translucent surface. **e**, Overview (left) and close-up (right) of the TOPRIM active site. The sharpened map is shown as a translucent surface. **f**, Superimposition of an AlphaFold3-predicted (colored red/blue according to predicted local distance difference test (pLDDT)) and the experimental Mg^{2+}-bound model. The arrow indicates different conformations.

Urea–PAGE analysis revealed endonucleolytic cleavage events upstream of dC nucleotides in positions 15, 17, 31 and 37 of the msDNA, in the presence of WT DenB but not the inactive G71R variant (Fig. 5c). These sites were within msDNA segments that loop around the fingers (positions 15, 17), lie adjacent to the trimer contact loop (position 31) or participate in formation of the 3WJ (position 37) (Fig. 5c). These results suggest that DenB cleaves the msDNA endonucleolytically, initiating msDNA decay that is likely completed by additional factors present in the CFEs.

To investigate the structural consequences of msDNA decay, we determined the cryo-EM structure of Eco2 in its active state at ~2.9 Å, including a local refinement of one subunit (chain A) to improve its map quality (Fig. 5d, Table 1 and Extended Data Fig. 8). The structure showed that removal of the msDNA exposes the TOPRIM domains, granting substrate access to the solvent-exposed active site (Fig. 5e).

Elimination of the msDNA cage did not result in large-scale structural rearrangements of the TOPRIM domain, as evidenced by superimposition with the msDNA-bound inhibited state (Fig. 5f). However, close inspection of the TOPRIM active site revealed a distinct reconfiguration of the catalytic pocket (Fig. 5g). Specifically, rearrangement of the loop connecting α-helix 18 and the active site placed residue D462 in a position more compatible with magnesium cofactor binding, compared to the inactive state structure (Fig. 5h). This conformation was reminiscent of the AlphaFold-predicted configuration (Fig. 3f and Supplementary Fig. 6). Concomitantly, the D460-coordinated Mg^{2+}_{B} cofactor was repositioned by ~1.5 Å, compared to the msDNA-inhibited inactive state (Fig. 5h). It is plausible that the proximity of the msDNA 3WJ to α-helices 17 and 18 facilitates this regulation allosterically (Extended Data Fig. 9).

Notably, the structure also revealed that only a single magnesium (D460-coordinated Mg^{2+}_{B}) was bound in the TOPRIM active site (Fig. 5g), suggesting that the second magnesium cofactor (D462-coordinated Mg^{2+}_{A}) may be recruited only upon substrate binding. This is consistent with previous observations for homologs of RNase H[48].

https://doi.org/10.1038/s41594-026-01754-2

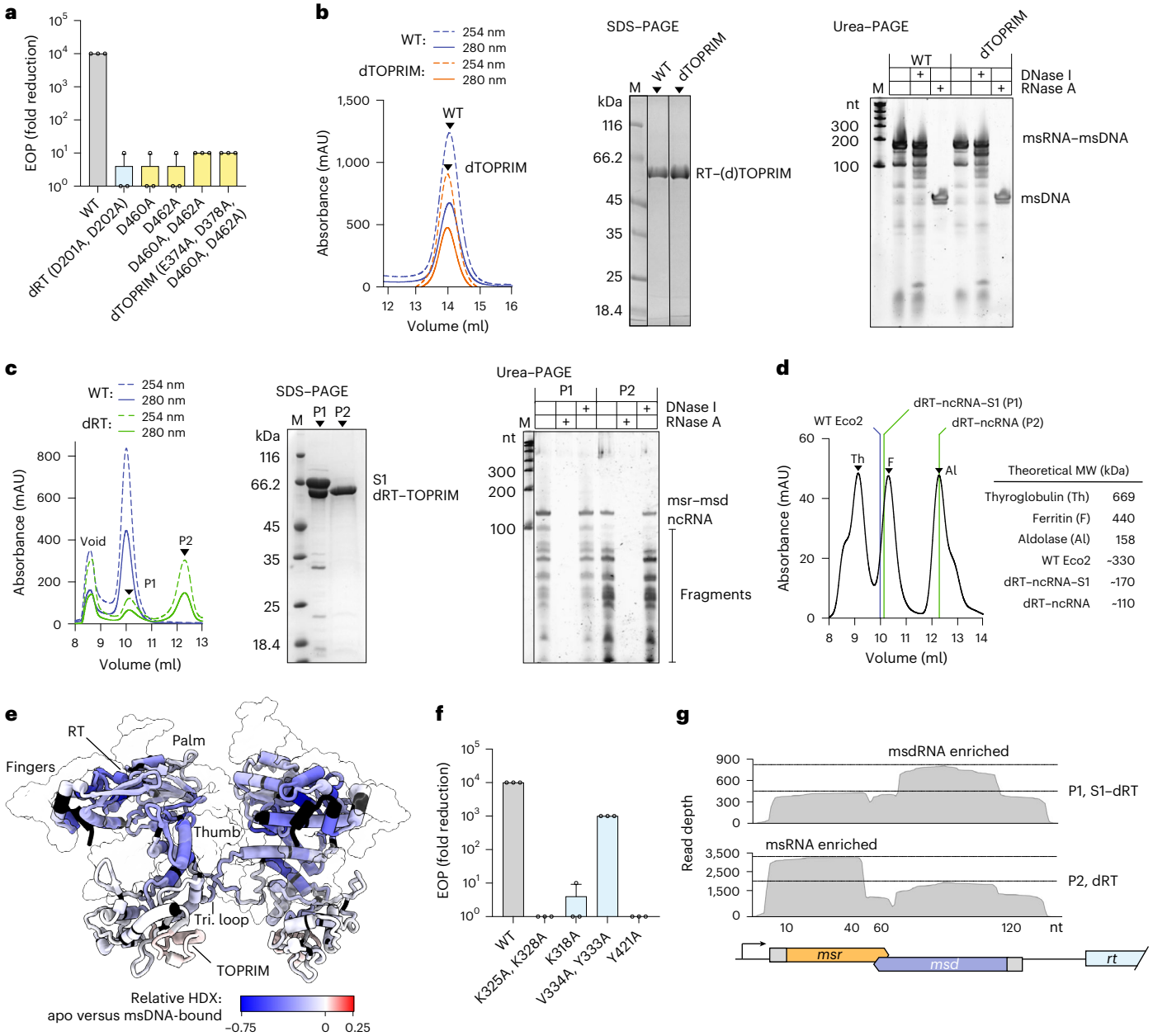


Fig. 4 | Biogenesis and structural dynamics of Eco2. a, Protection efficiency of Eco2 active site variants against T2, measured as EOP. $n = 3$ technical replicates; mean ± s.d. **b**,**c**, SEC traces for S6-column-purified (**b**) and S200-column-purified (**c**) Eco2 variants (left), and analytic SDS–PAGE (center) and urea–PAGE (right). Representative data; $n \geq 2$ biological replicates. **d**, Left: S200 calibration size-exclusion trace (Th, F and Al weight standards). Right: theoretical molecular weights. **e**, Changes in hydrogen-deuterium exchange (HDX) between apo and msDNA-bound RT–TOPRIM, mapped on the trimer structure. The msDNA is shown as a translucent cartoon surface. Color scale represents the highest difference in HDX observed at any of the time points. Mean values are plotted; $n = 3$ technical replicates. tri., trimerization. **f**, Protection efficiency of trimerization loop and msDNA binding site variants against T2. $n = 3$ technical replicates; mean ± s.d. **g**, Nanopore RNA sequencing results mapped to the *ncRNA–eco2* locus. $n = 1$. MW, molecular weight; nt, nucleotides.

Together, these data demonstrate that DenB-initiated msDNA degradation unblocks access to the active site and may allow stabilization of the catalytic pocket in an RNA-hydrolysis-competent conformation.

tRNA depletion interferes with protein production and phage propagation

The RNase activity of Eco2 suggested that the TOPRIM might cut cellular RNAs to prevent protein production for antiphage defense. To test this hypothesis, we employed a transcription–translation interference assay to measure GFP production in the presence of DenB and Eco2. The assay showed significant interference with GFP production in the presence of WT DenB and WT Eco2 but not their inactive variants G71R DenB (escaper variant) or dTOPRIM Eco2 (Fig. 6a). Nucleic acid urea–PAGE analysis of the CFEs revealed both msDNA degradation and tRNA cleavage in the presence of WT DenB and WT Eco2 (Fig. 6b). Incubation of isolated tRNAs with DenB-activated and subsequently purified WT Eco2, but not with the dTOPRIM variant, resulted in cutting of tRNAs (Fig. 6c).

We next analyzed the RNA cleavage preferences of Eco2 by sequencing RNA from the CFE reactions. Nanopore sequencing identified primarily tRNAs as substrates (tRNA-Gln, tRNA-Val, tRNA-Gly, tRNA-Asp, tRNA-Trp and tRNA-His) and revealed nicks in double-stranded tRNA acceptor stems, predominantly in C-rich sequences

https://doi.org/10.1038/s41594-026-01754-2

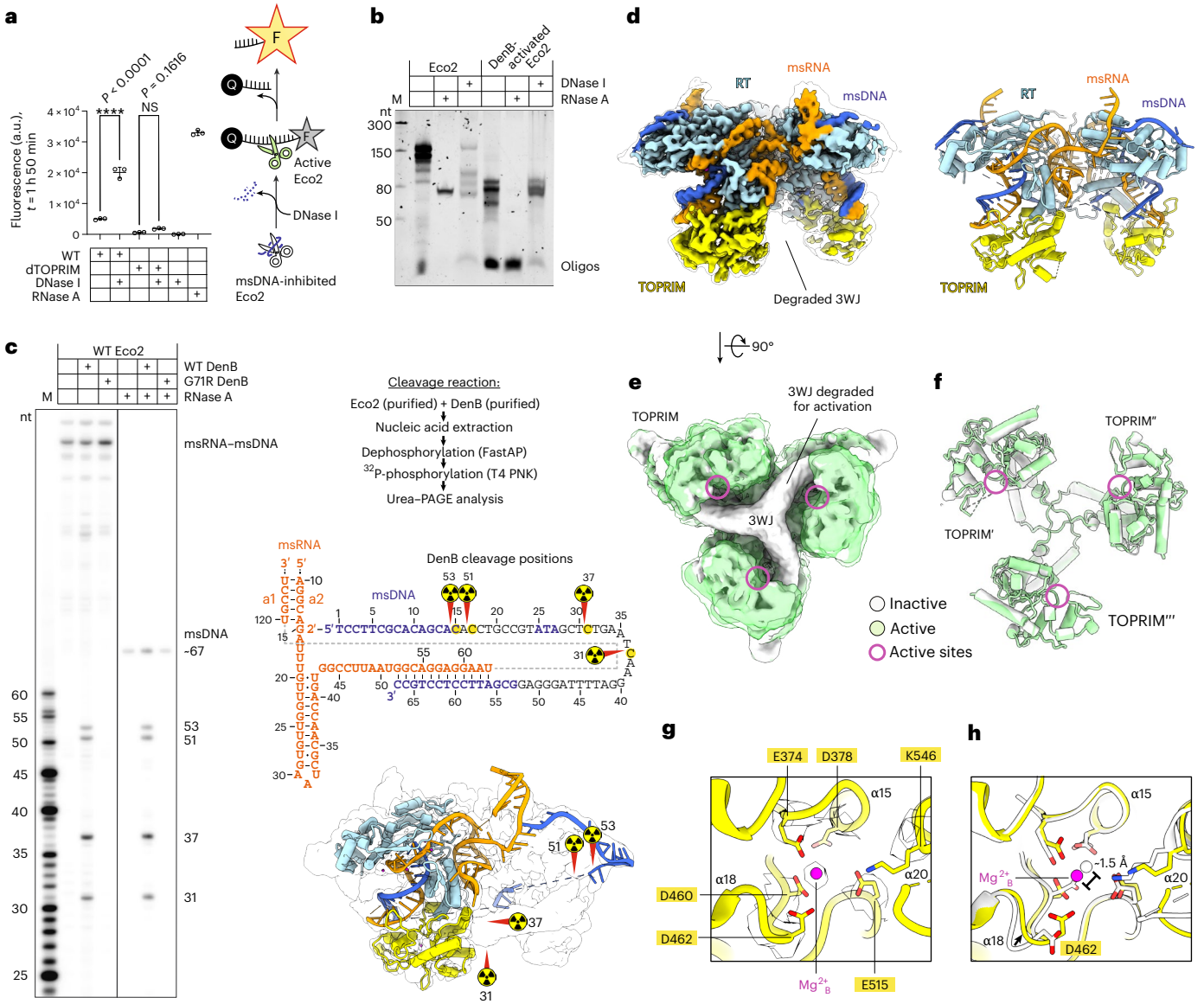


Fig. 5 | msDNA decay primes the TOPRIM for RNA cutting. a, Fluorophore-quencher (FQ) RNaseAlert assay probing FQ-reporter RNA cleavage by Eco2 in the presence of DNase I. The scheme illustrates the activation of Eco2 for FQ-reporter turnover. $n = 3$ technical replicates; mean ± s.d., one-way analysis of variance. NS, not significant. **b**, Urea–PAGE analysis of nucleic acids, RT–TOPRIM-copurified from *denB*-expressing CFEs. Representative data; $n \geq 2$ technical replicates. **c**, Left: urea–PAGE analysis of ^{32}P-labeled msDNA-cleavage products; right: msDNA scheme and inactive state structure (compare Fig. 2b) illustrating cleavage positions (red triangles). Representative data; $n \geq 2$ technical replicates. **d**, Cryo-EM map (left) and corresponding model (right) of activated and Mg^{2+}-bound Eco2. The unfiltered map is shown as a translucent surface. **e**, Unfiltered cryo-EM map superimposition of inactive state (solid white) and active state (translucent green) structures of Mg^{2+}-bound Eco2. **f**, TOPRIM domain superimposition of inactive state (white) and active state (green) structures of Mg^{2+}-bound Eco2. **g**, Close-up of the TOPRIM active site (active state). The sharpened map is shown as a translucent surface. **h**, Close-up of the superimposed TOPRIM active sites of the inactive (white) and active state (yellow). The arrow highlights different conformations.

located 6–7 nucleotides upstream of the CCA 3′ terminus (Fig. 6d and Supplementary Fig. 12).

To assess RNA turnover during phage infection, we infected *E. coli* strains expressing either WT Eco2 or the dTOPRIM variant with phages T2 or T5 at an estimated MOI of 10. RNA sequencing confirmed active infection by both T2 and T5 (Supplementary Fig. 13a). Notably, RNA reads mapping to T5 were more abundant than those mapping to T2, relative to reads from the *E. coli* host (Supplementary Fig. 13a). We next identified RNA species cut in the presence of the WT Eco2 but not the dTOPRIM variant. This revealed that Eco2 cuts host tRNAs (tRNA-Gly, tRNA-Arg, tRNA-Gln, tRNA-Val, tRNA-Met and tRNA-Pro) and phage tRNAs (tRNA-Asp, tRNA-His and tRNA-Gln) during the late stage of T5 infection (Fig. 6e). In addition, both phage and host transcripts were targeted by Eco2 (Fig. 6e and Supplementary Fig. 13b,d). Whereas the sequence specificity for tRNA aligned well with the in vitro preferences (Fig. 6d,f), Eco2 appeared to cut other RNAs promiscuously in vivo (Supplementary Fig. 13c). This observation was consistent with the results of our fluorophore-quencher-reporter assay, which showed nonspecific RNA turnover of the reporter (Fig. 5a).

These findings collectively show that Eco2 cuts various tRNAs and other RNAs in response to DenB, demonstrating that Eco2 is a broadly active RNase with a preference for selected tRNA species. In consequence, RNA degradation and tRNA depletion might lead to shutdown of translation for defense against phages, possibly by abortive infection.

https://doi.org/10.1038/s41594-026-01754-2

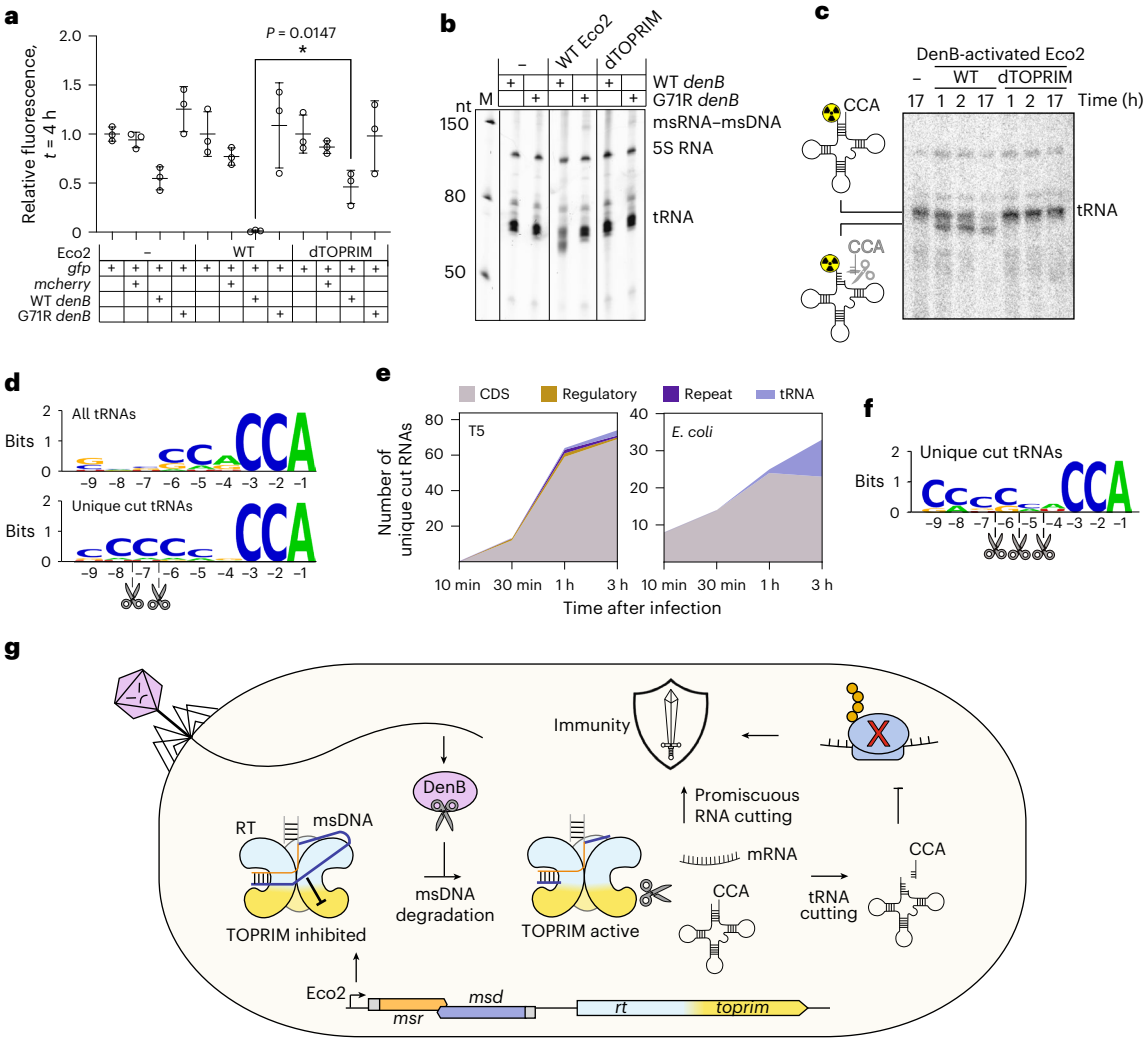


Fig. 6 | Eco2 cuts RNAs for antiphage defense. a, CFE transcription–translation interference assay measuring GFP fluorescence in the presence of Eco2 upon expression of *denB*. $n = 3$ technical replicates; mean ± s.d., one-way analysis of variance. Asterisk indicates statistical significance. **b**, Urea–PAGE analysis of CFE reactions producing DenB in the presence of Eco2. Representative data; $n \geq 2$ technical replicates. **c**, Urea–PAGE analysis of ^{32}P-labeled total tRNA turnover by DenB-activated and purified Eco2. Representative data; $n \geq 2$ technical replicates. **d**, tRNA sequence specificity of Eco2 (bottom) relative to all tRNA sequences (top). $n = 1$. **e**, Unique cut RNAs of T5 (left) and *E. coli* (right). $n = 2$ biological replicates. CDS, coding sequence. **f**, tRNA sequence specificity of Eco2 3 h after T5 infection. $n = 2$ biological replicates. **g**, Schematic of the defense mechanism of Eco2.

Discussion

Our study revealed the unique trimeric nucleoprotein structure of Eco2, its msDNA-dependent dynamics and regulation of the RT–TOPRIM fusion protein, and its RNA cutting activity with preference for tRNAs, which together enable the immune response to phage-encoded nucleases (Fig. 6g). Although there are currently no other documented examples of RT–effector fusion protein structures and molecular mechanisms beyond type I-C Eco2, it is plausible that other fusion protein retrons (that is, types VII-A1, XI and XII[16]) also utilize msDNA binding and dynamic rearrangements of the effector domain for activity regulation.

Recent studies on type II-A, type I-A and type I-B2 retrons have shown that the msDNAs produced by these multicomponent systems act as regulators of the separate effectors[17–19,21–26]. Together with our findings, these studies underscore the role of msDNA as a versatile sensor for DNA modifying and degrading activities that are central to phage biology. Our data suggest that in the case of Eco2, the msDNA might be further processed by host nucleases after initial endonucleolytic cleavage by DenB. It is plausible that phage-encoded DNA ligases could allow msDNA repair and thus Eco2 inactivation. Full msDNA degradation by the host may preclude such an escape mechanism and effectively set Eco2 to a state of no return for lasting protection.

We demonstrated that DenB-mediated msDNA degradation activates the TOPRIM domain of Eco2 for endonucleolytic nicking of double-stranded tRNA acceptor stems, resulting in a subsequent shutdown of protein production. This unique mechanism adds to the repertoire of immune systems that induce translational arrest by cleaving tRNAs. Unlike the acceptor stem cleavage of tRNAs by Eco2, other bona fide immune systems such as CRISPR–Cas13[49] and PARIS[50] target tRNA anticodon loops or stems, respectively. Structurally and mechanistically dissimilar type I-A retrons have been suggested to cut D-loops of tRNAs[22,34]. The mechanism of Eco2 appears to be akin to that of the HEPN–MNT toxin–antitoxin system, which clips off single-stranded tRNA CCA tails to interfere with translation[51], further illustrating the mechanistic parallels between tRNA-targeting toxin–antitoxin and immune systems[52–57]. How these different active sites bind to and hydrolyze their tRNA substrates is not well understood. Future studies will need to determine effector structures with trapped substrates in the active sites for a complete mechanistic understanding.

DenB and A1 are widely conserved among phages that are susceptible to Eco2. However, DenB- or A1-encoding phages Bas31, Bas46 and Bas47 evaded Eco2-mediated defense. This observation suggests that some phages employ strategies to overcome immunity, potentially through inhibitory factors[58] or by expression of degradation-resistant tRNAs[22,34,50,59].

Our Eco2 structures show that the msDNA forms a 3WJ. Formation of this structure necessitates template removal before DNA strand hybridization. Likewise, template degradation must occur before the msDNA can be recruited to RT DNA binding sites *in trans*. We demonstrated that the RNase-H-like TOPRIM of Eco2 does not participate in biogenesis. Given the independence of Eco2 from RNase H I[60], it will be interesting to see whether RNase H II[47,61] or another ribonuclease matures its msDNA.

We observed copurification of ribosomal protein S1 with the dRT–TOPRIM-bound ncRNA, indicating that S1 acts as a chaperone to stabilize the msDNA template. In addition to its function in translation initiation and mRNA chaperoning, S1 coordinates transcription and translation through RNAP-mediated delivery of mRNAs to ribosomes[42]. Thus, as well as chaperoning the msDNA template, S1 may facilitate coordination of transcription, translation and ncRNA recruitment for the nascent RT–TOPRIM protein. The orchestration of biogenesis, which must involve stringent regulation of effector components during complex assembly, remains poorly understood for many other microbial immune systems.

Retron-derived RTs have been widely used as genome editing tools, both independently and in combination with Cas nickases as prime editors[62–68]. Our study revealed the formation of an intricate trimeric nucleoprotein complex, which may have hindered applications of Eco2 for prime editing[65]. Our findings could inform structure-based engineering of Eco2-based editing tools to enable functional designs.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-026-01754-2.

References

1. González-Delgado, A., Mestre, M. R., Martínez-Abarca, F. & Toro, N. Prokaryotic reverse transcriptases: from retroelements to specialized defense systems. *FEMS Microbiol. Rev.* **45**, fuab025 (2021).
2. Mestre, M. R. et al. UG/Abi: a highly diverse family of prokaryotic reverse transcriptases associated with defense functions. *Nucleic Acids Res.* **50**, 6084–6101 (2022).
3. Toro, N., Martínez-Abarca, F., Mestre, M. R. & González-Delgado, A. Multiple origins of reverse transcriptases linked to CRISPR-Cas systems. *RNA Biol.* **16**, 1486–1493 (2019).
4. Simon, D. M. & Zimmerly, S. A diversity of uncharacterized reverse transcriptases in bacteria. *Nucleic Acids Res.* **36**, 7219–7229 (2008).
5. Lampson, B. C., Inouye, M. & Inouye, S. Retrons, msDNA, and the bacterial genome. *Cytogenet. Genome Res.* **110**, 491–499 (2005).
6. Yee, T., Furuichi, T., Inouye, S. & Inouye, M. Multicopy single-stranded DNA isolated from a gram-negative bacterium, *Myxococcus xanthus*. *Cell* **38**, 203–209 (1984).
7. Dhundale, A. R., Furuichi, T., Inouye, S. & Inouye, M. Distribution of multicopy single-stranded DNA among myxobacteria and related species. *J. Bacteriol.* **164**, 914–917 (1985).
8. Millman, A. et al. Bacterial retrons function in anti-phage defense. *Cell* **183**, 1551–1561 (2020).
9. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
10. Bobonis, J. et al. Bacterial retrons encode phage-defending tripartite toxin–antitoxin systems. *Nature* **609**, 144–150 (2022).
11. Dhundale, A., Lampson, B., Furuichi, T., Inouye, M. & Inouye, S. Structure of msDNA from Myxococcus xanthus: evidence for a long, self-annealing RNA precursor for the covalently linked, branched RNA. *Cell* **51**, 1105–1112 (1987).
12. Furuichi, T., Dhundale, A., Inouye, M. & Inouye, S. Branched RNA covalently linked to the 5′ end of a single-stranded DNA in *Stigmatella aurantiaca*: structure of msDNA. *Cell* **48**, 47–53 (1987).
13. Sun, J., Inouye, M. & Inouye, S. Association of a retroelement with a P4-like cryptic prophage (retronphage phi R73) integrated into the selenocystyl tRNA gene of *Escherichia coli*. *J. Bacteriol.* **173**, 4171–4181 (1991).
14. Stokar-Avihail, A. et al. Discovery of phage determinants that confer sensitivity to bacterial immune systems. *Cell* **186**, 1863–1876 (2023).
15. Azam, A. H. et al. Viruses encode tRNA and anti-retron to evade bacterial immunity. Preprint at *bioRxiv* https://doi.org/10.1101/2023.03.15.532788 (2023).
16. Mestre, M. R., González-Delgado, A., Gutiérrez-Rus, L. I., Martínez-Abarca, F. & Toro, N. Systematic prediction of genes functionally associated with bacterial retrons and classification of the encoded tripartite systems. *Nucleic Acids Res.* **48**, 12632–12647 (2020).
17. Carabias, A. et al. Retron-Eco1 assembles NAD-hydrolyzing filaments that provide immunity against bacteriophages. *Mol. Cell* **84**, 2185–2202 (2024).
18. Wang, Y. et al. Cryo-EM structures of *Escherichia coli* Ec86 retron complexes reveal architecture and defence mechanism. *Nat. Microbiol.* **7**, 1480–1489 (2022).
19. Wang, Y. et al. DNA methylation activates retron Ec86 filaments for antiphage defense. *Cell Rep.* **43**, 114857 (2024).
20. Li, X. et al. Architecture and mechanism of a dual-enzyme retron system in prokaryotic immunity. *Nat. Commun.* **17**, 487 (2026).
21. Wang, B., Hoffman, R., Hou, Y.-M. & Li, H. Structural basis for retron co-option of anti-phage ATPase-nuclease. *Nat. Struct. Mol. Biol.* **33**, 53–62 (2025).
22. George, J. T. et al. Structural basis of antiphage defence by an ATPase-associated reverse transcriptase. *Nat. Commun.* **16**, 8459 (2025).
23. Wang, C. et al. Disassembly activates Retron-Septu for antiphage defense. *Science* **389**, eadv3344 (2025).
24. Yuan, L. et al. Molecular mechanism of Eco8-mediated anti-phage defense. *Mol. Cell* https://doi.org/10.1016/j.molcel.2025.09.029 (2025).
25. Yu, C. et al. Phage SSB detection by retron Eco8 msDNA unleashes nuclease-mediated immunity. *Mol. Cell* https://doi.org/10.1016/j.molcel.2025.10.007 (2025).
26. Ji, C.-G. et al. Mechanistic insights into phage SSB-activated bacterial retron-Eco8 immunity. Preprint at *bioRxiv* https://doi.org/10.1101/2025.11.11.687766 (2025).
27. Sarafianos, S. G. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J. Mol. Biol.* **385**, 693–713 (2009).
28. Coté, M. L. & Roth, M. J. Murine leukemia virus reverse transcriptase: structural comparison with HIV-1 reverse transcriptase. *Virus Res.* **134**, 186–202 (2008).
29. Lampson, B. C. et al. Reverse transcriptase in a clinical strain of *Escherichia coli*: production of branched RNA-linked msDNA. *Science* **243**, 1033–1038 (1989).
30. Lampson, B. C., Viswanathan, M., Inouye, M. & Inouye, S. Reverse transcriptase from *Escherichia coli* exists as a complex with msDNA and is able to synthesize double-stranded DNA. *J. Biol. Chem.* **265**, 8490–8496 (1990).
31. Wang, Y. et al. Structural basis of the RNA-mediated Retron-Eco2 oligomerization. *Cell Discov.* **11**, 73 (2025).
32. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).

33. Lopatina, A., Tal, N. & Sorek, R. Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).
34. Azam, A. H. et al. Evasion of antiviral bacterial immunity by phage tRNAs. *Nat. Commun.* **15**, 9586 (2024).
35. Maffei, E. et al. Systematic exploration of *Escherichia coli* phage-host interactions with the BASEL phage collection. *PLoS Biol.* **19**, e3001424 (2021).
36. Ohshima, H., Hirano, N. & Takahashi, H. A hexanucleotide sequence (dC1-dC6 tract) restricts the dC-specific cleavage of single-stranded DNA by endonuclease IV of bacteriophage T4. *Nucleic Acids Res.* **35**, 6681–6689 (2007).
37. Hirano, N., Ohshima, H. & Takahashi, H. Biochemical analysis of the substrate specificity and sequence preference of endonuclease IV from bacteriophage T4, a dC-specific endonuclease implicated in restriction of dC-substituted T4 DNA synthesis. *Nucleic Acids Res.* **34**, 4743–4751 (2006).
38. Davison, J. Pre-early functions of bacteriophage T5 and its relatives. *Bacteriophage* **5**, e1086500 (2015).
39. McCorquodale, D. J., Chen, C. W., Joseph, M. K. & Woychik, R. Modification of RNA polymerase from *Escherichia coli* by pre-early gene products of bacteriophage T5. *J. Virol.* **40**, 958–962 (1981).
40. Beckman, L. D., Hoffman, M. S. & McCorquodale, D. J. Pre-early proteins of bacteriophage T5: structure and function. *J. Mol. Biol.* **62**, 551–564 (1971).
41. McCorquodale, D. J. & Lanni, Y. T. Patterns of protein synthesis in *Escherichia coli* infected by amber mutants in the first-step-transfer DNA of T5. *J. Mol. Biol.* **48**, 133–143 (1970).
42. Webster, M. W. et al. Molecular basis of mRNA delivery to the bacterial ribosome. *Science* **386**, eado8476 (2024).
43. Qu, X., Lancaster, L., Noller, H. F., Bustamante, C. & Tinoco, I. Jr. Ribosomal protein S1 unwinds double-stranded RNA in multiple steps. *Proc. Natl Acad. Sci. USA* **109**, 14458–14463 (2012).
44. Duval, M. et al. *Escherichia coli* ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol.* **11**, e1001731 (2013).
45. Sørensen, M. A., Fricke, J. & Pedersen, S. Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli in vivo*. *J. Mol. Biol.* **280**, 561–569 (1998).
46. Aravind, L., Leipe, D. D. & Koonin, E. V. Toprim–a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res.* **26**, 4205–4213 (1998).
47. Nowotny, M. H. M. F. RNases H: structure and mechanism. *DNA Repair* **84**, 102672 (2019).
48. Yang, W., Lee, J. Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg^{2+}-ion catalysis and substrate specificity. *Mol. Cell* **22**, 5–13 (2006).
49. Jain, I. et al. tRNA anticodon cleavage by target-activated CRISPR-Cas13a effector. *Sci. Adv.* https://doi.org/10.1126/sciadv.adl0164 (2024).
50. Burman, N. et al. A virally encoded tRNA neutralizes the PARIS antiviral defence system. *Nature* **634**, 424–431 (2024).
51. Songailiene, I. et al. HEPN-MNT toxin-antitoxin system: the HEPN ribonuclease is neutralized by oligoAMPylation. *Mol. Cell* **80**, 955–970 (2020).
52. LeRoux, M. & Laub, M. T. Toxin-antitoxin systems as phage defense elements. *Annu. Rev. Microbiol.* **76**, 21–43 (2022).
53. Levitz, R. et al. The optional E. coli prr locus encodes a latent form of phage T4-induced anticodon nuclease. *EMBO J.* **9**, 1383–1389 (1990).
54. Winther, K., Tree, J. J., Tollervey, D. & Gerdes, K. VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translation. *Nucleic Acids Res.* **44**, 9860–9871 (2016).
55. Cruz, J. W. et al. Growth-regulating *Mycobacterium tuberculosis* VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat. Commun.* **6**, 7480 (2015).
56. Klaiman, D., Steinfels-Kohn, E., Krutkina, E., Davidov, E. & Kaufmann, G. The wobble nucleotide-excising anticodon nuclease RloC is governed by the zinc-hook and DNA-dependent ATPase of its Rad50-like region. *Nucleic Acids Res.* **40**, 8568–8578 (2012).
57. Calcuttawala, F. et al. Apoptosis like symptoms associated with abortive infection of *Mycobacterium smegmatis* by mycobacteriophage D29. *PLoS ONE* **17**, e0259480 (2022).
58. Mayo-Muñoz, D., Pinilla-Redondo, R., Camara-Wilpert, S., Birkholz, N. & Fineran, P. C. Inhibitors of bacterial immune systems: discovery, mechanisms and applications. *Nat. Rev. Genet.* **25**, 237–254 (2024).
59. van den Berg, D. F., van der Steen, B. A., Costa, A. R. & Brouns, S. J. J. Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations. *eLife* https://doi.org/10.7554/eLife.85183 (2023).
60. Palka, C., Fishman, C. B., Bhattarai-Kline, S., Myers, S. A. & Shipman, S. L. Retron reverse transcriptase termination and phage defense are dependent on host RNase H1. *Nucleic Acids Res.* **50**, 3490–3504 (2022).
61. Rychlik, M. P. et al. Crystal structures of RNase H2 in complex with nucleic acid reveal the mechanism of RNA-DNA junction recognition and cleavage. *Mol. Cell* **40**, 658–670 (2010).
62. Khan, A. G. et al. An experimental census of retrons for DNA production and genome editing. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-024-02384-z (2024).
63. Fishman, C. B. et al. Continuous multiplexed phage genome editing using recombitrons. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-024-02370-5 (2024).
64. González-Delgado, A., Lopez, S. C., Rojas-Montero, M., Fishman, C. B. & Shipman, S. L. Simultaneous multi-site editing of individual genomes using retron arrays. *Nat. Chem. Biol.* **20**, 1482–1492 (2024).
65. Doman, J. L. et al. Phage-assisted evolution and protein engineering yield compact, efficient prime editors. *Cell* **186**, 3983–4002 (2023).
66. Lopez, S. C., Crawford, K. D., Lear, S. K., Bhattarai-Kline, S. & Shipman, S. L. Precise genome editing across kingdoms of life using retron-derived DNA. *Nat. Chem. Biol.* **18**, 199–206 (2022).
67. Buffington, J. D. et al. Discovery and engineering of retrons for precise genome editing. Preprint at *bioRxiv* https://doi.org/10.1101/2024.07.21.604473 (2024).
68. Cattle, M. A. et al. An enhanced Eco1 retron editor enables precision genome engineering in human cells without double-strand breaks. *Nucleic Acids Res.* **53**, gkaf716 (2025).

Methods

Phylogenetic analysis of Eco2

We retrieved the RT–TOPRIM sequences listed in Table S1 of ref. 16, as subtypes I-C1, I-C2, I-C3, XI and XII, directly from NCBI, UniParc and PATRIC. We increased sequence diversity by retrieving protein sequences predicted by Millman et al.[8] as 'RT/Predicted OLD family endonuclease (TOPRIM) fusion' from the Integrated Microbial Genomes system[69]. In addition, we ran a BLASTp search querying the RT–TOPRIM sequence obtained in that study, excluding *E. coli* from the subject database, and requesting 1,000 maximum target sequences. We retrieved a total of 1,003 sequences from the databases, accessed between November 2024 and January 2025.

To reduce redundancy of sequence diversity, we clustered all sequences into 412 clusters using CD-HIT v.4.8.1[70] and a 70% sequence identity threshold. We then aligned the representative sequences from each cluster using the MAFFT v.7.525 G-INS-I algorithm[71]. We evaluated alignments by eye using JalView v.2.11.4[72] and Geneious v.R11 (https://www.geneious.com). We used IQtree v.2.3.6[73] to infer the maximum likelihood phylogenies of concatenated RT–TOPRIM and separated RT from TOPRIM alignments (for the tanglegram), using Blosum62 for the amino acid substitution model. We visualized the phylogenies using Python library baltic v.0.3.0 (https://github.com/evogytis/baltic).

We used HMMER profiles to predict and confirm the presence of RT-like and TOPRIM-like domains in sequences of the I-C, XI and XII types. We downloaded the alignments for the RT–TOPRIM fused domain and RT-like, TOPRIM-like, TIR-like and protease-like domains from the National Institutes of Health Conserved Domains Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml; last used on January 2025). We used the commands hmmbuild and hmmsearch from HMMER v.3.4 (http://hmmer.org/) to build profiles and query the 412 sequences used in this study, respectively.

BASEL phage plaque assay screen

E. coli MG1655ΔRM[35] were transformed either with empty control vector pACYC184 (NEB) or Eco2-carrying plasmid pLG006[9] (Addgene 157884) and plated on LB-agar-chloramphenicol (Cm, 30 μg ml^{-1}), followed by incubation overnight at 37 °C. Subsequently, 4-ml LB-Cm cultures were inoculated and incubated for 16 h at 37 °C to prepare overnight cultures for double-layer agar plaque assays. Double-layer agar plates were prepared using 1.5% agar LB-Cm (30 μg ml^{-1} Cm) for the bottom layer and LB-Cm (30 μg ml^{-1} Cm) 0.75% agar mixed with 200 μl of overnight *E. coli* culture for the top layer. For effective infection, 20 mM $MgCl_2$ and 5 mM $CaCl_2$ were supplemented into the LB-agar. Ten-fold serial phage dilutions were spot-plated in 3.5-μl volumes onto the double-layer plates, before incubation at 37 °C for 16 h.

Isolation of T2 phage mutants

To generate Eco2-escaper phages, *E. coli* DH5α (Thermo Scientific) were transformed with plasmid pLG006[9] (Addgene 157884), plated on LB-agar (5 g l^{-1} yeast extract, 10 g l^{-1} tryptone, agar 15 g l^{-1}, 10 g l^{-1} NaCl) plates containing Cm (30 μg ml^{-1}) and incubated for 16 h at 37 °C. Subsequently, 4-ml LB-Cm cultures were inoculated and incubated for 16 h at 37 °C to prepare overnight cultures for double-layer agar plaque assays. Double-layer agar plates were prepared by layering 30 ml 1.5% LB-agar-Cm for the bottom layer and 5 ml 0.75% LB-agar-Cm mixed with 500 μl overnight culture for the top layer. Ten-fold dilutions series of T2 were then spot-plated in 3.5-μl volumes onto the top layer, before incubation for 16 h at 37 °C. Emerging escaper phage candidates that appeared to propagate well in the presence of Eco2, as indicated by large plaques at high dilution, were subsequently isolated by scraping spots for transfer into 1 ml phage buffer (50 mM Tris pH 7.4, 10 mM NaCl, 100 mM $MgCl_2$). Phages were incubated for 1 h at room temperature, followed by repeated vortexing for phage release. The agar was then removed by centrifugation at 3,200*g* for 10 min, and the phage supernatant was transferred to a new tube. Enriched escaper phages were then replated as serial dilutions on fresh Eco2-expressing *E. coli* DH5α double-layer plates (LB-agar-Cm) and incubated for 16 h at 37 °C. Enriched escaper phage spots were subsequently used for iterative phage isolation until the escaper phages had been purified. The final purification step was performed by scraping individual plaques for transfer in 4 ml LB-Cm containing *E. coli* DH5α at an optical density at 600 nm (OD_{600}) ≈ 0.5. Phages were subsequently incubated at 37 °C with shaking at 250 rpm for ~20 h. To separate phages from bacteria, cells were sedimented by centrifugation at 3,220*g* for 5 min. The phage supernatant was then filtered using a 0.22-μm filter and stored at 4 °C.

Phage DNA extraction

Prepared phage escaper solutions were propagated on a 10-ml *E. coli* DH5α culture in LB-Cm (Cm, 30 μg ml^{-1}) at 37 °C overnight. The next day, cultures were centrifuged for 2 min at 6,000*g*, and the supernatant was filtered through a 0.22-μm pore filter to remove cell debris and intact cells. Filtered supernatants were centrifuged at 12,000*g* for 1 h, the supernatant was discarded, and phages were resuspended in 400 μl of 30 mM Tris-HCl pH 7.5, 50 mM $MgCl_2$ and 0.5 mM $CaCl_2$ before incubation overnight at 4 °C. The next day, 1 μl of RNase A (10 mg ml^{-1}) and 5 μl of DNase I (1 U μl^{-1}) were added, and samples were incubated at 37 °C for 2 h. Subsequently, 15 μl of proteinase K (0.65 mg ml^{-1}) and EDTA (final concentration of 62.5 mM) were added to the mix, followed by incubation at 56 °C for 2 h. Phage genomic DNA was then purified using a GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific), according to the manufacturer's protocol.

Phage genome sequencing and escaper analysis

DNA libraries were prepared using a Colibri ES PCR Library Prep Kit (Invitrogen) according to the manufacturer's protocol and sequenced in 150 paired-end mode on a MiniSeq (Illumina). Adapters were trimmed with Bbduk from the BBmap package (v.38.18; https://sourceforge.net/projects/bbmap; ktrim = r, *k* = 15, mink = 11, hdist = 1, tpe, tbo). Reads were subsampled to 40,000 per dataset using seqtk –s 123, as breseq requires coverage of ~100. The initial WT T2 analysis was performed using breseq v.0.37.1 with default parameters on our laboratory strain T2 sequence data and reference genome (GenBank accession: LC348380.1). The reference genome of bacteriophage T2 was refined using the gdtools APPLY command to reflect the genetic composition of our laboratory T2 strain. This revised genome sequence served as the reference for subsequent breseq analyses using default parameters.

To identify DenB (T2, YP_010073922.1) and A1 (T5, AAU05157.1) homologs across BASEL phage genomes, 3 iterations of psi-blast against the NCBI clusteredNR database (2025-07-21 version) were performed, selecting sequences with 30–90% homology and >70% query coverage (expected threshold: 0.05, BLOSSUM62 substitution matrix). The sequences were clustered to 80% sequence identity for DenB and 70% sequence identity for A1, with 80% alignment coverage using the mmseqs2 algorithm (MPI Bioinformatics Toolkit v.c552cce6c3194c06bc0bba84f04c4ef13d62f0a5; https://toolkit.tuebingen.mpg.de/tools/mmseqs2). Sequences were aligned using UGENE v.51.0 with the default MAFFT algorithm (v.7.520) (gap opening penalty: 1.53). The alignment was visually inspected to remove sequences with long insertion and deletions, and realigned to remove unaligned N- and C-terminal stretches. The HMM profiles were built using the hmmbuild command from HMMER v.3.4 (http://hmmer.org/). The resulting HMM profiles were used to search the proteomes with the hmmsearch command (http://hmmer.org/E-value: 0.001).

E. coli growth assay

E. coli MG1655 were transformed with either empty control vector pACYC184 (NEB) or Eco2 plasmid pMJ001 and plated on LB-agar-Cm (25 μg ml^{-1} Cm), before incubation overnight at 37 °C. They were then inoculated with 4 ml LB-Cm (25 μg ml^{-1} Cm) medium, followed by incubation at 37 °C, 200 rpm, until the OD_{600} reached ~0.7. Subsequently,

cultures were diluted in 90 µl LB-Cm (25 µg ml^{-1}) and transferred into clear 96-well PS F-bottomed plates (Greiner), before addition of phages or medium to a total volume of 100 µl and final OD_{600} of 0.1. Phage titers were estimated by spot-plating before the experiment. Growth was monitored every 5 min for 16 h by OD measurement on a CLARIOstar Plus plate reader (BMG Labtech, software v.5.70 R2) at 37 °C, 200 rpm. Data were analyzed and plotted in Prism (v.10.5.0 (673), GraphPad).

Plasmid construction

Constructs for expression and affinity purification were cloned by fusing a sequence encoding a Strep-tag II downstream of the *rt–toprim* gene encoded on pLG006 (Addgene plasmid number: 157884), using Golden Gate assembly. Mutations encoding amino acid substitutions were subsequently introduced by Golden Gate mutagenesis. For expression of the N-terminally Strep-tag II-tagged apo *rt–toprim*, the gene was cloned into pRSFDuet-1 for Bacterial Adenylate Cyclase Two-Hybrid (BACTH) system experiments, and genes were cloned by Golden Gate assembly upstream of *t18* into pUT18. Plasmids listed in Supplementary Table 1 are available for reuse from the corresponding author (P.P., Vilnius University) upon request.

Protein production and purification from *E. coli* cultures

All proteins were purified at a constant temperature of 4 °C and kept on ice.

For small-scale purification of Eco2 and its variants, Eco2 (ncRNA-encoding and RT–TOPRIM-Strep-tag-encoding) plasmids were transformed in *E. coli* DH5α (Thermo Scientific), plated on LB-agar plates containing Cm (25 µg ml^{-1}) and incubated at 37 °C overnight. For expression of Eco2 or variants thereof, 30 ml TB-Cm (12 g l^{-1} tryptone, 24 g l^{-1} yeast extract, 4 ml l^{-1} glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4, 25 µg ml^{-1} Cm) was directly inoculated with cells from transformation plates with multiple colonies, followed by incubation at 37 °C with shaking at 200 rpm for 16 h. Cells were subsequently harvested and resuspended in 30 ml lysis buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA), followed by cell disruption at 10,000 PSI using an LM10 Microfluidizer (Microfluidics) equipped with an H10Z interaction chamber. Lysates were clarified by centrifugation at 20,000*g* for 20 min at 4 °C. Supernatants were loaded onto 100 µl Strep-Tactin XT resin (Cytiva) and incubated for 10 min on ice with shaking. Subsequently, the resin was centrifuged at 4,000*g* for 5 min at 4 °C, and the supernatant was discarded. The resin was then resuspended in 700 µl lysis buffer and transferred to a Zymo-Spin P1 column (Zymo Research). To remove nonspecific contaminants, the resin was washed 3 times with 700 µl lysis buffer before incubation with 100 µl elution buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 50 mM biotin) for 5 min on ice. Proteins were eluted by centrifugation 4,000*g* for 1 min at 4 °C. Concentrations were subsequently estimated based on the absorbance at 280 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific). Protein samples were analyzed by SDS–PAGE and SYPRO Ruby (Thermo Scientific) staining.

For preparative purifications, Eco2-encoding (ncRNA-encoding and RT–TOPRIM-Strep-tag-encoding) plasmids were transformed in *E. coli* DH5α (Thermo Scientific) plated on LB-agar plates containing Cm (25 µg ml^{-1}) and incubated at 37 °C overnight. For expression of Eco2 or variants thereof, 2 l TB-Cm (25 µg ml^{-1} Cm) was directly inoculated with cells from transformation plates with multiple colonies, followed by incubation at 37 °C with shaking at 200 rpm for 16 h. Cells were subsequently harvested and resuspended in 40 ml lysis buffer (c) before cell disruption using an LM10 Microfluidizer (Microfluidics). Lysates were clarified by centrifugation at 20,000*g* for 40 min at 4 °C. The supernatants were subsequently loaded onto 1-ml StrepTrap XP columns (Cytiva) pre-equilibrated in lysis buffer and washed with 20 column volumes (CV) of 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA buffer before elution with 6 CV elution buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 50 mM biotin). Samples were further purified by size-exclusion chromatography (SEC) using Superose 6 10/300 (Cytiva) columns for Eco2 and variants thereof, and a Superdex 200 10/300 (Cytiva) column for the Eco2 dRT variant, pre-equilibrated in SEC buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl). Peak fractions were concentrated to approximately 50 µl, and concentrations were estimated based on the absorbance at 280 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific). Samples were split in aliquots, snap-frozen in liquid nitrogen and stored at −70 °C before use in assays.

For production of Strep-tagged apo RT–TOPRIM, pRSFDuet-derived plasmids encoding the *strep–rt–toprim* gene were transformed in *E. coli* BL21 Star (DE3) (Invitrogen) and plated on LB-agar plates containing kanamycin (Kan, 50 µg ml^{-1}). To express the *strep–rt–toprim* gene, 10-ml LB-Kan overnight cultures (50 µg ml^{-1} Kan), grown with shaking at 37 °C, were used to inoculate 2 l TB-Kan (50 µg ml^{-1} Kan) and incubated with shaking at 200 rpm and 37 °C until the OD_{600} reached ~0.6. Gene expression was induced with 0.5 mM IPTG, followed by incubation for 16 h at 16 °C. Cells were subsequently collected by 3,600*g* centrifugation and resuspended in lysis buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA) before cell disruption using a microfluidizer. The lysate was clarified by centrifugation at 20,000*g* for 40 min at 4 °C. The supernatant was then loaded on a 5-ml StrepTrap XP column (Cytiva) pre-equilibrated in lysis buffer, washed with 5 CV lysis buffer and eluted with 2 CV elution buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA and 50 mM biotin). The elution was then applied to a 5-ml HiTrap Q HP column (Cytiva), pre-equilibrated in a buffer containing 20 mM NaCl and 100 mM Tris-HCl (pH 8.0). After a 4-CV column wash with 20 mM NaCl and 100 mM Tris-HCl (pH 8.0) buffer, a 15 CV gradient from 20 mM to 2 M NaCl in 100 mM Tris-HCl (pH 8.0) was applied for elution. The peak fraction was concentrated and further purified by SEC using a Superose 6 10/300 (Cytiva) column pre-equilibrated in SEC buffer. The protein was then concentrated to approximately 50 µl, and the concentration was estimated based on the absorbance at 280 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific). The sample was split in aliquots, snap-frozen in liquid nitrogen and stored at −70 °C.

For production of His-tagged S1 protein, pRSFDuet-derived plasmids encoding the *rpsA-his$_6$* gene were transformed in *E. coli* BL21 (DE3) (Invitrogen) and plated on LB-agar plates containing Kan (50 µg ml^{-1}). For expression of the *rpsA-his$_6$* gene, 10-ml LB-Kan overnight cultures (50 µg ml^{-1} Kan), grown with shaking (200 rpm at 37 °C), were used to inoculate 2 l TB-Kan (50 µg ml^{-1} Kan), followed by incubation with shaking at 200 rpm and 37 °C until the OD_{600} reached ~0.6. Gene expression was induced with 0.5 mM IPTG, before incubation for 16 h at 16 °C. Cells were subsequently collected by centrifugation (3,600*g*) and resuspended in lysis buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$, 40 mM imidazole) before cell disruption by sonication using a Vibra-Cell ultrasonic processor at 40% amplitude for 5 min with pulses of 3 s at 6-s intervals. The lysate was clarified by centrifugation at 20,000*g* for 60 min at 4 °C. The supernatant was then loaded on a 5-ml HisTrap column (Cytiva) pre-equilibrated in lysis buffer, washed with 3 CV lysis buffer and eluted with 15 CV elution buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$, 500 mM imidazole). Peak elution fractions were collected and applied to a 5-ml HiTrap Q HP column (Cytiva), pre-equilibrated in a buffer containing 20 mM NaCl, 20 mM Tris-HCl (pH 8.0) and 5 mM $MgCl_2$. After a 4-CV column wash with 20 mM NaCl, 20 mM Tris-HCl (pH 8.0) and 5 mM $MgCl_2$ buffer, a 15 CV gradient from 20 mM to 2 M NaCl in 100 mM Tris-HCl (pH 8.0) and 5 mM $MgCl_2$ was applied for elution. The peak fraction was concentrated and further purified by SEC using a Superdex 200 10/300 (Cytiva) column pre-equilibrated in SEC buffer (150 mM NaCl, 100 mM Tris-HCl (pH 8.0) and 5 mM $MgCl_2$). The protein was then concentrated to approximately 200 µl, and the concentration was estimated based on the absorbance at 280 nm using a NanoDrop Eight spectrophotometer

(Thermo Scientific). The sample was split in aliquots, snap-frozen in liquid nitrogen and stored at −70 °C.

Complex reconstitution

S1 and apo RT–TOPRIM were produced as described above. The ncRNA was produced using a MEGAscript T7 Transcription Kit (Thermo Fisher Scientific) according to the manufacturer's manual and purified with TRIzol (Thermo Fisher Scientific). Concentrations were estimated based on the absorbance at 260 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific). The ncRNA was split in aliquots, snap-frozen in liquid nitrogen and stored at −70 °C.

Before complex assembly, the ncRNA was heated to 65 °C for 1 min and cooled to 20 °C to facilitate secondary structure formation. S1, apo RT–TOPRIM and ncRNA were combined in a 1:1:1.2 ratio (2.75 µM S1, 2.75 µM apo RT–TOPRIM, 3.29 µM ncRNA) and incubated at 20 °C for 10 min in reconstitution buffer (100 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10% glycerol). S1 and apo RT–TOPRIM were combined in a 1:1 ratio (2.75 µM S1, 2.75 µM apo RT–TOPRIM) and incubated at 20 °C for 10 min in reconstitution buffer. Subsequently, assembly reactions or separate components (S1, apo RT–TOPRIM or ncRNA) were injected into a Superdex 200 3.2/300 column (Cytiva) pre-equilibrated in reconstitution buffer at 4 °C.

Urea–PAGE analysis of nucleic acids

Protein–nucleic acid complex samples (10 µl, 0.5–7.0 µM) were combined with 1 µl proteinase K (~20 mg ml^{-1}, Thermo Fisher Scientific) and optionally 1 µl DNase I (1 U $µl^{-1}$, Thermo Fisher Scientific) and/or 1 µl RNase A (10 mg ml^{-1}, Thermo Fisher Scientific) for 30 min at 37 °C in DNAse I reaction buffer (Thermo Fisher Scientific). Subsequently, samples were mixed with one volume of 2X RNA Loading Dye (Thermo Fisher Scientific), incubated at 70 °C for 10 min, and cooled on ice before separation by 12.5% Urea–PAGE. Gels were stained with SYBR Gold (Thermo Fisher Scientific) and visualized using a GelDoc Go imaging system (Bio-Rad, software v.3.0.0.07).

For Urea–PAGE analysis of DenB cleavage products, the nucleic acids underwent a two-step treatment process. First, nucleic acids were treated with alkaline phosphatase (FastAP, Thermo Scientific) in a 10-µl reaction containing 0.5× PNK A buffer (50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mM spermidine), 500 nM DNA and 1.5 µl FastAP enzyme (1 U $µl^{-1}$). To stop this reaction, 5 µl of 15 mM EDTA was added (final concentration: 5 mM), followed by heating at 65 °C for 15 min. The samples were then supplemented with 2 µl of 10× PNK A buffer, 2 µl of 3.3 µM γ-[^{32}P]-ATP (6,000 Ci $mmol^{-1}$) and 1 µl of T4-PNK enzyme (10 U $µl^{-1}$, Thermo Scientific). This mixture was incubated at 37 °C for 30 min. The reaction was stopped by addition of 5 µl of 50 mM EDTA and heating at 75 °C for 10 min (25 µl final volume). Parallel reactions were prepared following the same protocol, with the addition of RNase A (Thermo Scientific) to a final concentration of 0.4 mg ml^{-1} during the phosphorylation step. Urea–PAGE was used to analyze 5 µl of the reaction mixtures. Samples were separated on a 15% (29:1 acrylamide/bis-acrylamide) sequencing gel with 7 M of urea in 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA). Gels were run for 2 h at 60 °C and visualized by phosphor imaging on a Amersham Typhoon scanner (GE Healthcare, software v.2.0.0.6).

Cryo-EM grid preparation and data acquisition

Eco2 complexes were purified as described above either with or without 10 mM $MgCl_2$ supplementation in the SEC buffer. The activated Eco2 sample was purified as described in the 'Protein purification from CFEs' section. $MgCl_2$ was added to the thawed sample to a final concentration of 10 mM. Cryo-EM grids were prepared by applying 3 µl of Eco2 sample (7 µM) to Quantifoil R 1.2/1.3 300 copper mesh grids with a 2-nm continuous carbon support layer; these were glow-discharged at 20 mA for 7 s (GLOQUBE PLUS, Quorum) or at 12 W for 5 s (Zepto-W6, Diener) before sample application. The grids were flash-frozen in liquid ethane using Vitrobot Mark IV (Thermo Scientific) with 100% relative humidity at 4 °C, with no preincubation and a blotting time of 5 s.

Datasets were collected on a Glacios Cryo-TEM (Thermo Scientific) operated at 200 kV and equipped with a Falcon III directed electron detector (Thermo Scientific) at ×92,000 magnification, with a pixel size of 1.1 Å, 30 frames with a total dose of 30 $e^{-}/Å^2$, and defocus range of −2.2 µm to −1.0 µm. Data collection was performed using EPU v.3.2 and v3.10 software (Thermo Fisher Scientific).

Cryo-EM data processing

Data for Eco2 without $MgCl_2$ were processed using CryoSPARC (v.4.3)[74]. A dataset of 3,826 micrographs was collected, and patch motion correction and contrast transfer function (CTF) estimation were done using on-the-fly processing in CryoSPARC Live. Blob autopicking (box size: 360 pixels) resulted in 2,914,940 particles, which were sorted by two-dimensional (2D) classification. An initial model was generated from the best 2D classes with 583,457 particles. The particles were further classified using heterogeneous refinement. The best class, with 508,057 particles, was refined using homogeneous refinement, and the data were further polished by global and local CTF refinement and reference-based motion correction. The final reconstruction on polished particles contained 504,525 particles and had a global resolution of 2.76 Å.

For the Eco2 sample in the presence of $MgCl_2$, a dataset of 2,309 movies was collected. Data processing was performed using CryoSPARC (v.4.6)[74]. Blob autopicking resulted in 1,037,763 particles extracted with a box size of 360 pixels, and these were sorted by 2D classification. The best class, containing 339,226 particles, was selected following several rounds of 2D classification and heterogenous refinement. The dataset was further polished using per-particle CTF and reference-based motion correction. A symmetry expansion job was run for C3 symmetry, and local refinement was performed using a mask covering a single Eco2 monomer. After refinement, a final reconstruction with global resolution of 3.04 Å was obtained containing 1,017,678 C3 symmetry expanded particles. The final map was sharpened using tools in CryoSPARC (v.4.6).

A dataset of 2,968 micrographs was collected for DenB-activated Eco2. Patch motion correction and CTF estimation were done using on-the-fly processing in CryoSPARC Live, and data processing was performed using CryoSPARC (v.4.7)[74]. Template picking was performed using a template generated from the map of Eco2 in the absence of $MgCl_2$, resulting in 1,430,679 particles, which were sorted by 2D classification. An initial model was generated from the best 2D classes and further classified using heterogeneous refinement. Global and local CTF refinement, reference-based motion correction and rebalance orientations jobs were performed, followed by nonuniform refinement. A symmetry expansion job was run for C3 symmetry, and local refinement was performed using a mask covering a single Eco2 monomer. A final reconstruction with global resolution of 2.90 Å was obtained containing 617,364 C3 symmetry expanded particles. The final map was sharpened using tools in CryoSPARC (v.4.6).

Model building, refinement and figure preparation

The initial RT–TOPRIM model was generated using AlphaFold2[75] under the ColabFold[76] framework with default parameters and fitted into the cryo-EM map in UCSF ChimeraX (v.1.6.1)[77]. The model was manually modified using Coot (v.0.9.8.7)[78] to fit the map. RNA and DNA nucleotides observed in the map were built manually in Coot. Further model refinement and evaluation were performed using Phenix real space refine (v.1.20.1–4487)[79] against sharpened maps. Secondary structure restraints were generated for nucleic acid base pairs. Global minimization and local grid search strategies were used. The resulting model was then used as a starting model for Eco2 in the presence of $MgCl_2$ and activated Eco2 and was fitted into the experimental maps in

UCSF ChimeraX (v.1.8). The model was manually modified using Coot (v.0.9.8.95) and refined using Phenix (v.1.21.2-5419) as described above. Figures were prepared in UCSF ChimeraX.

AlphaFold structure prediction

DenB and RT–TOPRIM structures in the presence of ligands were predicted using AlphaFold3[80] (default settings) via the AlphaFold Server (https://deepmind.google/technologies/alphafold/alphafold-server/). Predicted local distance difference test values were mapped on structures using UCSF ChimeraX.

Sample preparation for proteomic analysis

Protein bands of interest were excised from SDS–PAGE gels, stained with colloidal Coomassie (0.08% (w/v) Coomassie Brilliant Blue G250, 10% (w/v) citric acid, 8% (w/v) ammonium sulfate, 20% (v/v) methanol) overnight. For all following steps, buffers were exchanged by two consecutive 15-min incubation steps of the gel pieces with 200 µl of acetonitrile (ACN), with ACN removed after each step. Proteins were reduced by addition of 200 µl of a 10 mM DTT solution in 100 mM ammonium bicarbonate (AmBiC, Sigma Aldrich, A6141), samples were incubated at 56 °C for 20 min, 180 µl ACN was added, and samples were incubated for a further 15 min at room temperature. Proteins were alkylated for 20 min by addition of 200 µl of a 55 mM chloroacetamide solution in 100 mM AmBiC. Gel pieces were incubated twice with 200 µl ACN for 15 min at room temperature.

In-gel enzymatic digest

A 0.2 µg µl^{-1} stock solution of trypsin (Promega, V511A) in resuspension buffer (Promega, V542A) was diluted with ice-cold freshly prepared 50 mM AmBiC buffer to achieve a final concentration of 2 ng µl^{-1}. Then, 50 µl of this solution was added to gel pieces, followed by incubation for 30 min on ice and thereafter overnight at 37 °C. Gel pieces were sonicated for 15 min and spun down, and the supernatant was transferred into a glass vial (VDS optilab, 93908556). The remaining gel pieces were washed with 50 µl of an aqueous solution of 50% (v/v) ACN and 1% (v/v) formic acid and sonicated for 15 min. The combined supernatants were dried in a SpeedVac and reconstituted in 10 µl of an aqueous solution of 0.1% (v/v) formic acid.

LC–MS/MS analysis of peptides

Peptides were analyzed by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS) on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) as previously described[81]. To this end, peptides were separated using an Ultimate 3000 RSLCnano system (Dionex) equipped with a trapping cartridge (precolumn C18 PepMap100, 5 mm, 300 µm i.d., 5 µm, 100 Å) and an analytical column (Acclaim PepMap 100. 75 × 50 cm C18, 3 mm, 100 Å) connected to a Nanospray Flex ion source. The peptides were loaded onto the trap column at 30 µl per min using solvent A (0.1% (v/v) formic acid in water), and peptides were eluted using a gradient from 2% to 85% solvent B (0.1% (v/v) formic acid in ACN) over 30 min at 0.3 µl per min (all solvents were of LC–MS grade). The Orbitrap Fusion Lumos was operated in positive ion mode with a spray voltage of 2.2 kV and capillary temperature of 275 °C. Full-scan MS spectra with a mass range of 375–1200 m/z were acquired in profile mode using a resolution of 120,000 (maximum injection time of 50 ms); the AGC target was set to 400% and a maximum injection time of 86 ms. Precursors were isolated using the quadrupole with a window of 1.2 m/z, and fragmentation was triggered by high-energy collision dissociation in fixed collision energy mode with fixed collision energy of 34%. MS2 spectra were acquired with the Orbitrap with a resolution of 30,000 and a max injection time of 86 ms.

The Orbitrap Fusion Lumos was operated in positive ion mode with a spray voltage of 2.2 kV and capillary temperature of 275 °C. Full-scan MS spectra with a mass range of 350–1500 m/z were acquired in profile mode using a resolution of 120,000 (maximum injection time of 100 ms); the AGC target was set to standard, and a RF lens setting of 30% was used. Precursors were isolated using the quadrupole with a window of 1.2 m/z, and fragmentation was triggered by high-energy collision dissociation in fixed collision energy mode with a fixed collision energy of 30%. MS2 spectra were acquired in ion trap normal mode. The dynamic exclusion was set to 5 s.

Acquired data were analyzed using FragPipe[82] and an *E. coli* UniProt FASTA database (UP000000625, ID83333, 4.402 entries, date: 27.10.2022, downloaded: 11.01.2023) including common contaminants as well as the protein sequence of our protein of interest. The mass error tolerance was set to 10 ppm for full-scan MS spectra and 0.02 Da for MS/MS spectra. A maximum of two missed cleavages was allowed. A minimum of 2 unique peptides with a peptide length of at least 7 amino acids and a false discovery rate less than 0.01 were required on the peptide and protein level[83].

Bacterial two-hybrid assay

Protein–protein interactions were assayed using a BACTH system kit (Euromedex). In brief, the *rpsA* (*s1*) gene and its truncated variants were fused with the T25 adenylate cyclase fragment in plasmid vector pKNT25. WT Eco2, dRT Eco2 and apo RT–TOPRIM loci were fused downstream of the *rt–toprim* gene with the T18 adenylate cyclase fragment in pUT18. The BTH101 cells were cotransformed with plasmid pairs and grown overnight at 30 °C in LB medium containing 100 µg ml^{-1} ampicillin and 50 µg ml^{-1} Kan. Then, 2 ml LB-ampicillin-Kan was inoculated with 40 µl of overnight culture and grown to $OD_{600} \approx 0.8$. For spot tests, 7 µl of fresh cultures were spotted on M9 minimal medium (47.7 mM $Na_2HPO_4 \times 7H_2O$, 22 mM KH_2PO_4, 8.5 mM NaCl, 18.7 mM NH_4Cl, 0.2% (w/v) glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.5 mM IPTG, 60 µg ml^{-1} X-Gal, 100 µg ml^{-1} ampicillin and 50 µg ml^{-1} Kan) agar plates, before incubation at 30 °C for 20 h. Spots were then assessed for blue color, indicative of protein–protein interactions.

Hydrogen/deuterium exchange mass spectrometry

HDX-MS experiments to investigate the impact of msDNA on Eco2 conformation and topology were conducted essentially as described previously[84]. The msDNA-bound Eco2 and apo RT–TOPRIM were purified as described above. HDX reactions were automatically set up with a two-arm robotic autosampler (LEAP Technologies) by dispension of 6.5 µl of Eco2 sample (40 µM) in a 96-well plate, followed by addition of 58.5 µl HDX buffer (20 mM Tris-Cl (pH 8.0), 5 mM $MgCl_2$, 150 mM NaCl) prepared with 99.9% D_2O. After incubation at 25 °C for various durations (10 s, 100 s, 1,000 s or 10,000 s), 55 µl of the HDX reaction was withdrawn and added to 55 µl of predispensed quench buffer (400 mM KH_2PO_4/H_3PO_4 (pH 2.2), 2 M guanidine-HCl) temperated at 1 °C, and 95 µl of the resulting mixture was injected into an ACQUITY UPLC M-Class System with HDX Technology (Waters)[85]. Nondeuterated samples were prepared similarly (incubation for approximately 10 s at 25 °C) by 10-fold dilution of samples with HDX buffer prepared with H_2O. The injected samples were delivered with H_2O + 0.1% (v/v) formic acid (100 µl per min) to a column (2 mm × 2 cm, 12 °C) containing porcine pepsin immobilized to beads, and the resulting peptic peptides were trapped on an ACQUITY UPLC BEH C18 1.7-µm 2.1 ×5 mm VanGuard Pre-column (Waters) kept at 0.5 °C. After 3 min of digestion and trapping, the trap column was placed in line with an ACQUITY UPLC BEH C18 1.7-µm 1.0 ×100 mm column (Waters), and the peptides were eluted at 0.5 °C with a gradient of H_2O + 0.1% (v/v) formic acid (eluent A) and ACN + 0.1% (v/v) formic acid (eluent B) at 30 µl per min as follows: 0–7 min: 95–65% A; 7–8 min: 65–15% A; 8–10 min: 15% A; 10–11 min: 5% A; 11–16 min: 95% A. The peptides were guided to a G2-Si high-definition MS (HDMS) mass spectrometer with ion mobility separation (Waters) and ionized with an electrospray ionization source (250 °C capillary temperature, 3.0 kV spray voltage), and mass spectra were acquired in positive ion mode over a range of 50–2,000 m/z in enhanced HDMS

(HDMS^E) or HDMS mode for nondeuterated and deuterated samples, respectively[86]. Lock-mass correction was implemented with a [Glu1] Fibrinopeptide B-Standard (Waters). During separation of the peptide mixtures on the C18 column, the protease column was washed three times with 80 µl of wash solution (0.5 M guanidine hydrochloride in 4% (v/v) ACN), and blank injections were performed between each sample to reduce peptide carry-over. Measurements were conducted in technical triplicates (individual HDX reactions).

ProteinLynx Global SERVER (v.3.0.1, Waters) and DynamX (v.3.0, Waters) facilitated peptide identification and analysis of deuterium incorporation essentially as described previously[84]. In brief, peptides were identified with ProteinLynx Global SERVER from the nondeuterated samples acquired with HDMS^E by employing low energy, elevated energy and intensity thresholds of 300, 100 and 1,000 counts, respectively. Identified ions were matched to peptides with a database containing the amino acid sequences of Eco2 and porcine pepsin and their reversed sequences with the following search parameters: peptide tolerance = automatic; fragment tolerance = automatic; min fragment ion matches per peptide = 1; min fragment ion matches per protein = 7; min peptide matches per protein = 3; maximum hits to return = 20; maximum protein mass = 250,000; primary digest reagent = non-specific; missed cleavages = 0; false discovery rate = 100. Only peptides identified in at least three nondeuterated samples with a minimum intensity of 10,000 counts, a peptide length of 5–40 residues, a minimum number of 2 products, a maximum mass error of 25 ppm, and a retention time tolerance of 0.5 min were considered for further analysis. All spectra were manually inspected with DynamX 3.0, and, if necessary, peptides were omitted (for example, in cases of low signal-to-noise ratio or presence of overlapping peptides). Residue-specific HDX was calculated from overlapping peptides with DynamX 3.0 using the shortest peptide covering a residue. Where multiple peptides were of the shortest length, the peptide with the residue closest to the C terminus was used. Numerical peptide and residue-specific HDX data are contained in the source data.

RNase activity assay

Reactions were assembled by combining 1 µM protein and 200 nM RNase Alert (IDT) FQ probe in reaction buffer (10 mM HEPES (pH 7.5), 250 mM KCl, 5 mM $MgCl_2$, 5% (v/v) glycerol) in a total volume of 30 µl in 384-well flat-bottomed black polystyrene assay plates (Thermo Fisher Scientific). Reactions to assess the effects of DNA degradation also contained 16.7 mU µl^{-1} DNase I (Thermo Fisher Scientific). The reactions were monitored by fluorescence measurements (λ_{ex}: 489 nm; λ_{em}: 522 nm) in a CLARIOstar Plus plate reader (BMG Labtech, software v.5.70 R2) at 37 °C every 30 s. Data were analyzed and plotted in Prism (v.10.4.0 (527), GraphPad).

CFE preparation

CFEs were prepared following a protocol adapted from ref. 87. Briefly, a saturated overnight culture of *E. coli* BL21 Star (DE3) (Invitrogen) was inoculated into sterile 2-l baffled flasks, containing 400 ml of 2× YTPG media (16 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} NaCl, 22 mM KH_2PO_4, 40 mM K_2HPO_4, 100 mM D-(+)-glucose, pH 7.2). The initial inoculation density was adjusted to an OD_{600} of 0.05, and cultures were incubated at 37 °C, 200 rpm. To induce T7 RNA polymerase expression, IPTG was added to a final concentration of 1 mM once the cultures reached an OD_{600} of ~0.5. Cells were harvested during the mid-exponential phase (OD_{600}: 2.5–3.0) via centrifugation and washed three times with ice-cold filtered S30A buffer (50 mM Tris-acetate (pH 7.8), 14 mM magnesium L-glutamate, 60 mM potassium L-glutamate, 2 mM DTT). The resulting biomass pellets were collected by centrifugation, decanted, weighed, flash-frozen in liquid nitrogen and stored at −80 °C. Next, the biomass pellet was thawed on ice and resuspended in ice-cold filtered S30A buffer (1 ml of buffer per 1 g of biomass). Cells were then lysed using a Sonics Vibra-Cell ultrasonic liquid processor (Sonics & Materials) equipped with a 13-mm probe. Lysis was performed in an ice bath at 50% amplitude, with a total lysis time of 30 s per 1 g of biomass, using a pulsing cycle of 3 s on and 6 s off. The resulting lysate was cleared by centrifugation at 48,384*g* for 30 min at 4 °C, followed by incubation for 1 h at 37 °C and shaking at 200 rpm for 1 h. To remove any remaining debris, the subsequent centrifugation step was performed for 10 min at 3,113 rcf, 4 °C. The total protein concentration of the lysate was determined using a Pierce Bradford Plus Protein Assay Kit (Thermo Fisher Scientific). The lysate was aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C.

Preparation of templates for expression in CFEs

Linear DNA templates for *denB* WT, the *denB*-G71R escaper mutant and mCherry expression in CFEs were generated by two-step fusion PCR[88]. In the first step, the target genes were amplified by PCR, attaching a GC-rich fusion sequence at the 5′ ends. In the second step, the amplified genes were combined with a short DNA cassette, similar to those described by Garamella et al.[89], containing the native *E. coli* RNA polymerase P70a or T7 RNA polymerase PT714 promoters, a ribosome-binding site, a Strep II-tag, and the GC-rich fusion sequence at the 3′ end. These components were fused via PCR using flanking ORG13 and ORG113 oligonucleotides. PCR products were analyzed by agarose gel electrophoresis, and the resulting fusion products were purified using a GeneJET PCR Purification Kit (Thermo Fisher Scientific).

The pCI-T7Max-UTR1-deGFP-8xHis-T500 reporter plasmid was a gift from K. Adamala (Addgene plasmid 178422)[90]. The plasmid was propagated in *E. coli* DH5α (Thermo Scientific) and purified using a ZymoPURE II Plasmid Maxiprep Kit (Zymo Research).

Transcription–translation interference assay in CFEs

Transcription–translation interference experiments were performed following guidelines from Garenne et al. and Marshall et al.[91,92]. Transcription–translation interference reactions were assembled on ice and aliquoted in 9-µl volumes into a Nunc 384 Well ShallowWell Standard Height Black microplate (Thermo Fisher, 264705). Each reaction contained 10 mg ml^{-1} *E. coli* BL21 Star (DE3) CFE, 57 mM HEPES-K (pH 8.0), 0.2 mg ml^{-1} tRNA mix from *E. coli* MRE600 (Sigma Aldrich), 1.5 mM ATP, 1.5 mM GTP, 0.9 mM CTP, 0.9 mM UTP, 0.75 mM cAMP, 30 mM 3-PGA, 0.33 mM NAD, 0.26 mM coenzyme A, 68 µM folinic acid, 1 mM spermidine, 30 mM D-ribose, 60 mM maltodextrin, 2 mM of each of the 20 L-amino acids, 7 mM magnesium L-glutamate, 50 mM potassium L-glutamate, 1 mM DTT, 20 g l^{-1} PEG 8000, 2.5 µM GamS inhibitor, 5 nM pCI-T7Max-UTR1-deGFP-8xHis-T500 reporter plasmid, 20 nM linear P70a DNA expression templates encoding either mCherry, DenB or dDenB (G71R), and 1 µM purified Eco2 or dTOPRIM. The plate was centrifuged for 1 min at 200*g* and 4 °C, equilibrated at room temperature for 10 min, sealed with MicroAmp Optical Adhesive Film (Thermo Fisher Scientific) and transferred to a CLARIOstar Plus microplate reader (BMG Labtech, software v.5.70 R2) for measurements. The fluorescence intensities of GFP (λ_{ex}: 470 nM ± 15 nM, λ_{em}: 515 nM ± 20 nM) and mCherry (λ_{ex}: 570 nM ± 15 nM, λ_{em}: 620 nM ± 20 nM bandpass window) were measured every 5 min for 16 h at 30 °C with double orbital shaking at 200 rpm for 30 s before each measurement.

As fluorescence signals plateaued after 4 h, data from this time point were used for further analysis. The GFP fluorescence in each reaction was first normalized to background noise by subtraction of the fluorescence of corresponding negative control reactions lacking the reporter plasmid. To obtain relative values, each normalized GFP value was divided by the fluorescence of a corresponding positive control reaction containing no linear DNA templates. Data were analyzed and plotted in Prism (v.10.4.0 (527), GraphPad).

RNA isolation from CFEs and Urea–PAGE analysis

CFE reactions were performed as described above. Reactions were mixed with 50 µl of TRIzol reagent (Thermo Fisher Scientific), and

total RNA was extracted according to the manufacturer's recommendations. The single-stranded RNA concentration in each sample was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Then, 5 µl 100 ng µl^{-1} single-stranded RNA from each sample was mixed with 5 µl of 2× RNA loading dye (NEB), denatured for 10 min at 75 °C and separated on a denaturing urea polyacrylamide gel (8 M urea, 12% polyacrylamide (29:1 acrylamide/bis-acrylamide ratio) in 1× TBE (89 mM Tris-borate (pH 8.4), 2 mM EDTA)), prerun for 30 min at 35 mA. The electrophoresis was performed for ~2.5 h at 35 mA. The gel was subsequently stained in 1× TBE (89 mM Tris-borate (pH 8.4), 2 mM EDTA) supplemented with 1× SYBR Gold dye (Thermo Fisher Scientific) on an orbital shaker for 10 min at room temperature and imaged with a Amersham Typhoon scanner (GE Healthcare, control software v.2.0.0.6) using the SYBR Gold channel, with analysis using with Fiji (v.2.14.0/1.54f).

Protein purification from CFEs

DenB was synthesized using NEBExpress (NEB) cell-free expression systems following the manufacturer's recommendations. A total of 250 µl of reaction mixture, containing 50 nM of linear PT714-Strep-DenB expression template and 2.5 µM homemade GamS was prepared. The RecBCD inhibitor GamS was purified from *E. coli* for CFE supplementation as described before[93]. The reaction mix was divided into 50-µl reactions in 1.5-ml tubes and incubated in a thermomixer (Eppendorf) for 18 h at 37 °C, 800 rpm. Lysates were cleared by centrifugation for 5 min at 17,000*g*, 4 °C. The reactions were pooled, combined with 250 µl of 2× binding buffer (200 mM Tris-HCl (pH 8.0), 300 mM NaCl, 2 mM EDTA, 10 mM 2-mercaptoethanol) and loaded onto Zymo-spin P1 columns (Zymo Research, P2003-1) preloaded with dry Strep-Tactin XT Sepharose chromatography resin (Cytiva, 29401324). Before use, 100 µl of the resin solution was loaded into the column, washed 3 times with 1× binding buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 5 mM 2-mercaptoethanol) and decanted by centrifugation (1–2 min at 1,000*g*, 4 °C). The bottom of the column was covered with parafilm, and it was incubated for 30 min at room temperature, with vigorous shaking on a 3D shaker (Biosan). Subsequently, the flow-through was collected by centrifugation, and the beads were washed 3 times with wash buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM 2-mercaptoethanol) and decanted by centrifugation. Subsequently 100 µl of elution buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM 2-mercaptoethanol, 50 mM D-biotin) was added to dry aspirated beads, the bottom of the column was covered with parafilm, and the mix was incubated for 15 min at room temperature, with vigorous shaking on a Multi Bio 3D shaker (Biosan). The elution was collected by centrifugation. DenB synthesis and purification was assessed by SDS–PAGE and western blotting. The concentration of the purified WT DenB protein was determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific). The protein was aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C.

For the Eco2 activation and tRNA cutting assay, WT DenB was synthesized using NEBExpress (NEB) cell-free expression systems following the manufacturer's recommendations. A reaction mixture containing 100 nM linear PT714-His$_6$-DenB-WT expression template, 2.5 µM homemade GamS and the NEBExpress cell-free expression system mix was prepared to a total volume of 500 µl. The reaction was divided into 50-µl aliquots in 1.5-ml tubes and incubated at 37 °C for 18 h, with shaking at 800 rpm, in a thermomixer (Eppendorf). Lysates were clarified by centrifugation at 17,000*g* for 10 min at 4 °C. Pooled lysates were split into two aliquots. Each aliquot was mixed with 25 µl of 15.76 µM purified Eco2 or dTOPRIM, respectively, and incubated at 37 °C for 1 h in a thermomixer (Eppendorf) at 800 rpm. Subsequently, 238 µl of each reaction was mixed with an equal volume of 2× binding buffer (200 mM Tris-HCl (pH 8.0), 300 mM NaCl, 2 mM EDTA, 10 mM 2-mercaptoethanol) and loaded onto Zymo-spin P1 columns (Zymo Research, P2003-1) preloaded with dry Strep-Tactin XT Sepharose chromatography resin (Cytiva, 29401324). Before loading, 100 µl of the resin solution was loaded into the column, washed three times with 1× binding buffer (100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 5 mM 2-mercaptoethanol), followed by centrifugation (1–2 min at 500–1,000*g*, 4 °C). After addition of the lysate–resin mixture, the bottom of each microcolumn was sealed with parafilm, and the mixture was incubated for 1 h at room temperature with vigorous shaking on a 3D orbital shaker (Biosan). The flowthrough was collected by centrifugation. The beads were washed three times with wash buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM 2-mercaptoethanol), with removal of the supernatant after each wash via centrifugation. For elution, 50 µl of elution buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM 2-mercaptoethanol, 50 mM D-biotin) was added to the aspirated beads. The microcolumns were sealed with parafilm and incubated for 30 min at room temperature with vigorous shaking on a 3D orbital shaker (Biosan). The eluates were collected by centrifugation and analyzed by SDS–PAGE and western blotting. Protein concentrations of the DenB-activated and purified Eco2 variants were determined using a Qubit Protein Assay (Thermo Fisher). Proteins were aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C. The thawed samples were analyzed by SEC on a Superdex 200 3.2/300 column pre-equilibrated with a buffer containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl and 10 mM $MgCl_2$.

tRNA cleavage assay

The tRNA mix from *E. coli* MRE 600 (Sigma Aldrich) was dephosphorylated in a 10-µl reaction containing 0.5× PNK Buffer A (Thermo Fisher), 250 nM tRNA mix and 0.15 U µl^{-1} FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher). The reaction was incubated at 37 °C for 10 min, and subsequent enzyme inactivation was achieved by addition of 5 µl of 15 mM EDTA, followed by heating of the reaction mixture at 75 °C for 10 min. Dephosphorylated tRNA was subsequently 5′-radiolabeled in a 20-µl reaction. The radiolabeling reaction mixture contained 1.25× PNK Buffer A (Thermo Fisher), 125 nM tRNA mix, 5 U µl^{-1} T4 Polynucleotide Kinase (Thermo Fisher), 125 nM γ-[^{32}P]-ATP (6000 Ci mmol^{-1}), 3.75 mM EDTA and 0.075 U µl^{-1} heat-inactivated FastAP Thermosensitive Alkaline Phosphatase. The reaction was incubated at 37 °C for 30 min, after which 5 µl of 50 mM EDTA was added, and the reaction mixture was heated at 75 °C for 10 min to stop the reaction. Following heating, the thermoblock containing the reaction mixture was allowed to cool to room temperature to facilitate proper folding of tRNA molecules. The labeled tRNA was aliquoted and stored at −20 °C.

DenB-activated Eco2 and dTOPRIM were produced as described above. Activated Eco2 was combined with radiolabeled tRNA mix in a 30-µl reaction (final composition: 25 mM Tris-HCl (pH 8.0), 250 mM NaCl, 5 mM $MgCl_2$, 5% (v/v) glycerol, 0.05 nM ^{32}P-tRNA mix, 877 nM DenB-activated and purified Eco2/dTOPRIM) and subsequently incubated at 37 °C. At 1 h, 2 h and 17 h, 10 µl of each reaction was extracted and mixed with 10 µl of phenol/chloroform/isoamyl alcohol solution (ROTH). The mixtures were centrifuged for 5 min at 17,000*g*, and the water fraction was mixed with 10 µl of 2× RNA loading dye (Thermo Fisher Scientific), followed by heating for 10 min at 75 °C. The denatured samples were analyzed by 12 % Urea–PAGE. The gel was dried, exposed to a phosphor screen overnight and read using an Amersham Typhoon scanner (GE Healthcare, software v.2.0.0.6) via the phosphor channel and analyzed with Fiji (v.2.14.0/1.54f).

Protein analysis by SDS–PAGE and western blotting

Samples (12 µl) were mixed with 4 µl of 4× SDS dye (10% glycerol (v/v), 2% SDS (w/v), 63 mM Tris-HCl (pH 6.8), 0.1% 2-mercaptoethanol (v/v), 0.01% bromophenol blue (w/v)) and denatured for 5 min at 95 °C. Samples were then separated on homemade 12% or 15% Tris-glycine SDS polyacrylamide gels and stained with Coomassie solution (0.1% Coomassie R-250, 40% (v/v) EtOH, 10% (v/v) acetic acid) or used for western blotting.

For western blot analysis, samples were separated on 12% or 15% Tris-glycine SDS polyacrylamide gels and transferred onto 0.22-µm or 0.45-µm polyvinylidene fluoride membranes (Thermo Fisher Scientific) in a semi-dry transfer with a Pierce Power Station blotter (Thermo Fisher Scientific). The membranes were blocked (5% (w/w) bovine serum albumin, 1× phosphate-buffered saline, 0.2% (v/v) Tween-20) for 30 min to 1 h with shaking at room temperature. The membranes were incubated with a 1:4,000 dilution of Strep·Tag II Antibody HRP Conjugate (Merck, 71591) in blocking buffer overnight with shaking at 4 °C, and subsequently washed 3–5 times (1× phosphate-buffered saline, 0.2% (v/v) Tween-20) and developed using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) for visualization on a Chemidoc system (Bio-Rad). The images were analyzed with Fiji (v.2.14.0/1.54f).

RNA sequencing and analysis of CFEs

Ten-microliter CFEs were transferred into TRIzol for RNA extraction according to the manufacturer's protocol (Thermo Fisher Scientific). Complementary DNA libraries were prepared using a Small RNA-Seq Library Prep Kit (Lexogen) with 500 ng RNA input material, according to the manufacturer's protocol. Libraries were subsequently pooled and prepared for nanopore sequencing using an Oxford Nanopore Ligation Sequencing Kit V14 (SQK-LSK114). The resulting libraries were sequenced on the Oxford Nanopore platform (R10.4.1 flow cell), and base calling was performed using dorado v.0.8.0, sup v.5.0.0 (Oxford Nanopore Technologies; https://github.com/nanoporetech/dorad). Datasets were demultiplexed and adapters trimmed using dorado v.0.8.0 with Illumina barcodes as custom input. For identification of RNA cleavage positions, reads were mapped to the *E. coli* MG1655 genome (GenBank: U00096.3) using minimap2 v.2.17[94] (-ax map-ont), and samtools v.1.13[95] was used to remove ambiguously mapped reads (-F 256 -F 2048). Data were analyzed using custom Python scripts. Read end positions were identified using python pysam package v.0.22.1 (https://github.com/pysam-developers/pysam), and counts were normalized according to the sample with fewer reads. Positions without coverage were assigned a pseudocount of 1. The read end ratio was calculated at each position to identify read ends enriched in the WT Eco2 sample (cutoff: 20). Hits present in genomic regions without gene annotations were not further analyzed.

Sequencing and analysis of protein-bound ncRNA

Protein–nucleic acid complexes were purified in the preparative scale as described above. Samples (100 µl, 14–17 µM) were transferred into TRIzol for RNA extraction according to the manufacturer's protocol (Thermo Fisher Scientific). Complementary DNA libraries were prepared using a Small RNA-Seq Library Prep Kit (Lexogen) with 300 ng RNA input material, according to the manufacturer's protocol. Libraries were subsequently barcoded using an Oxford Nanopore Native Barcoding Kit V14 (SQK-NBD114.96) and sequenced on the Oxford Nanopore platform (R10.4.1 flow cell), and base calling was performed using dorado v.0.7.3, sup v.4.3.0 (Oxford Nanopore Technologies; https://github.com/nanoporetech/dorado). Reads were aligned to the *E. coli* DH5α genome (GenBank: CP026085.1) and to the retron operon with minimap2 v.2.24[94] using default parameters. The dataset was filtered to exclude ambiguously mapped reads using samtools v.1.13[95] (-F 256 -F 2048). Reads were visualized using the Integrative Genomics Viewer[96] and custom Python scripts. The bed file with genome features was prepared from the gff3 file using awk, and *E. coli* gene coverage was evaluated using bedtools v.2.29.2[97] with default parameters.

Deep RNA sequencing and analysis of infected *E. coli*

E. coli MG1655 were transformed with pMJ001 or pMJ003 plasmids and plated on LB-agar-Cm (25 µg ml^{-1} Cm), before incubation overnight at 37 °C. 4 ml LB-Cm (25 µg ml^{-1} Cm) medium was subsequently inoculated with 6–7 colonies, followed by incubation at 37 °C, 200 rpm, until the OD_{600} reached ~0.7. Bacterial cultures were transferred into sterile 1.5-ml Eppendorf tubes, normalized by dilution in LB-Cm to an OD_{600} of 0.7, and infected with T2 or T5 at an estimated MOI of ~10 in a final volume of 660 µl. Cultures were incubated at 37 °C, 200 rpm. Samples (150 µl) were taken 10 min, 30 min, 1 h and 3 h after infection and immediately centrifuged for 30 s at 20,000*g*. The supernatant was rapidly decanted, and cell pellets were flash-frozen in liquid nitrogen. For RNA extraction, the pellets were resuspended in 400 µl TRIzol, and RNA was extracted according to the manufacturer's manual (Thermo Fisher Scientific). Concentrations were estimated based on the absorbance at 260 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific).

Small RNA fractions were enriched from purified RNA (0.5–1.0 µg) using a SPLIT RNA extraction kit (Lexogen) according to the manufacturer's manual. Concentrations were estimated based on the absorbance at 260 nm using a NanoDrop Eight spectrophotometer (Thermo Scientific).

To identify cleaved RNA positions, an adapter ligation-based RNA library preparation method was used. Sequencing libraries were prepared using a Small RNA-Seq Library Prep Kit for Illumina (Lexogen) according to the manufacturer's manual, using 200–500 ng of RNA as input. Library quality was evaluated with Bioanalyzer using a DNA high sensitivity kit (Agilent) and Qubit 4 fluorimeter (Thermo Fisher Scientific). The library was sequenced using an Illumina NextSeq P3 flow cell in 100 PE mode (EMBL Genomics Core facility, Heidelberg, Germany). Data quality was evaluated using fastqc v.0.12.1 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed using Bbduk v.38.18 (https://sourceforge.net/projects/bbmap/; ktrim = r, *k* = 20, mink = 10, hdist = 1, tpe, tbo, minlen = 17, minavgquality = 30). The *E. coli* MG1655 genome (GenBank ID: NZ_LR881938.1) was joined with phage T5 (GenBank ID: NC_005859.1) and T2 (GenBank ID: NC_054931.1) genomes using the Bash cat command. Annotation files.gff and. gbff were created in the same way. The corresponding reads were aligned to the joined genome files using bwa-mem2 v.2.2.1[98] with default parameters and filtered using samtools v.1.13 (-F 260 -f 2). For indexing and sorting, samtools was used. The data were further analyzed using custom Python scripts. In brief, read end positions were identified with pysam v.0.23.3 (read.reference_end and read.reference_start methods), genomic positions without read ends were assigned a pseudo count of 1. Read end ratios (WT Eco2/dTOPRIM) at each genomic position were calculated, and enriched positions were selected by setting the filtering cutoff to 100. Finally, only enriched positions present in both biological replicates were further analyzed. Corresponding transcripts and sequences were extracted from the annotation files, and the most prominent putative cleavage positions were verified manually using the Integrative Genomics Viewer[96].

Reporting summary

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

Data availability

Cryo-EM maps and model coordinates were deposited to the Electron Microscopy Data Bank EMD-(52583, EMD-52584, EMD-54448) and Protein Data Bank (PDB) (9I2F, 9I2G, 9S1F). Cryo-EM maps, half-maps, models and PDB validation reports are available via figshare at https://doi.org/10.6084/m9.figshare.29930282 (ref. 99). MS data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with dataset identifiers PXD060313 and PXD067212. Data related to sequencing experiments and RT–TOPRIM phylogeny are available via Mendeley Data, V1, at https://data.mendeley.com/datasets/v9b289v6ct/1. Raw sequencing reads generated from in vivo experiments were deposited in the NCBI Sequence Read Archive under BioProject accession PRJNA1358792. Source data are provided with this paper.

Code availability

Code related to sequencing experiments and RT–TOPRIM phylogeny is available via Mendeley Data, V1, at https://data.mendeley.com/datasets/v9b289v6ct/1.

References

69. Chen, I.-M. A. et al. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* **47**, D666–D677 (2019).
70. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
71. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
72. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
73. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
74. 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).
75. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
76. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
77. Meng, E. C. et al. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
78. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
79. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
80. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
81. Sridharan, S. et al. Proteome-wide solubility and thermal stability profiling reveals distinct regulatory roles for ATP. *Nat. Commun.* **10**, 1155 (2019).
82. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520 (2017).
83. Savitski, M. M., Wilhelm, M., Hahne, H., Kuster, B. & Bantscheff, M. A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteomics* **14**, 2394–2404 (2015).
84. Düring, J. et al. Structural mechanisms of autoinhibition and substrate recognition by the ubiquitin ligase HACE1. *Nat. Struct. Mol. Biol.* **31**, 364–377 (2024).
85. Wales, T. E., Fadgen, K. E., Gerhardt, G. C. & Engen, J. R. High-speed and high-resolution UPLC separation at zero degrees Celsius. *Anal. Chem.* **80**, 6815–6820 (2008).
86. Li, G.-Z. et al. Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics* **9**, 1696–1719 (2009).
87. Dopp, J. L., Jo, Y. R. & Reuel, N. F. Methods to reduce variability in *E. coli*-based cell-free protein expression experiments. *Synth. Syst. Biotechnol.* **4**, 204–211 (2019).
88. Cha-Aim, K., Hoshida, H., Fukunaga, T. & Akada, R. Fusion PCR via novel overlap sequences. *Methods Mol. Biol.* **852**, 97–110 (2012).
89. Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The All E. coli TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol.* **5**, 344–355 (2016).
90. Deich, C. et al. T7Max transcription system. *J. Biol. Eng.* **17**, 4 (2023).
91. Garenne, D., Thompson, S., Brisson, A., Khakimzhan, A. & Noireaux, V. The all-E. coliTXTL toolbox 3.0: new capabilities of a cell-free synthetic biology platform. *Synth. Biol.* **6**, ysab017 (2021).
92. Marshall, R., Beisel, C. L. & Noireaux, V. Rapid testing of CRISPR nucleases and guide RNAs in an *E. coli* cell-free transcription-translation system. *STAR Protoc.* **1**, 100003 (2020).
93. Batista, A. C. et al. Differentially optimized cell-free buffer enables robust expression from unprotected linear DNA in exonuclease-deficient extracts. *ACS Synth. Biol.* **11**, 732–746 (2022).
94. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572–4574 (2021).
95. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
96. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
97. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
98. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 314–324 (IEEE, 2019).
99. Jasnauskaitė, M., Skorupskaite, A., Miksys, A., Malinauskaite, L. & Pausch, P. Structure and mechanism of anti-phage retron Eco2. *figshare* https://doi.org/10.6084/m9.figshare.29930282 (2025).

Acknowledgements

We thank the department of protein–DNA interactions at the Life Sciences Center at Vilnius University (Virginijus Šikšnys) for access to equipment, and the Vilnius University cryo-EM facility (Giedrius Sasnauskas) for support. We also thank A. Harms (ETH Zürich, Switzerland) for providing the BASEL phage collection; L. Aitmanaitė and G. Russo for performing preliminary nanopore RNA sequencing runs; D. Smalakytė for protein purification advice; L. Gao and F. Zhang (MIT, USA) for providing pLG006 (Addgene plasmid 157884) and K. Adamala (University of Minnesota, USA) for Addgene plasmid 178422; M. Rodríguez Mestre (University of Copenhagen, Denmark) for enabling access to sequences and for discussion; and B. Wiedenheft, N. Burman and J. George (Montana State University, USA) for sharing the structure of Eco FORC 82 before publication. We further thank the EMBL Proteomics Core Facility for MS analysis of protein samples, and the EMBL Genomics Core Facility for deep RNA sequencing. P.P. receives funding from the Research Council of Lithuania, Lithuania under the EMBO Installation Grant programme under grant No 5342 and is supported by the Horizon Europe European Research Council (ERC) grant JUPITER number 101163749. R.G. and A.S. acknowledge support by the Research Council of Lithuania (LMTLT) under grant agreements S-PD-24-151 (R.G.) and S-PD-24-74 (A.S.). This project has received funding from the Research Council of Lithuania, Lithuania under the EMBO Installation Grant programme under grant No 5305 (G.D.). M.F.T.J. acknowledges the support of the ERA fellowship project ENTWINE No. 101130794 (2024-2026). I.S. acknowledges support by the LMTLT under grant agreement S-MIP-24-51 (I.S.). I.S. and J.J. were supported by Vilnius University Science Promotion fund projects MSFJM-30/2022 (I.S.) and MSFJM-12/2021 (J.J.). We thank the DFG for support through the Marburg core facility for Interaction, Dynamics and Assembly of Biomolecular Structures, project numbers 324652314 (G.B.) and 260989694 (G.B.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

P.P. conceived the study. M.F.T.J. and G.D. performed the phylogenetic analysis. M.J. performed growth and infection assays. T.L. tested protection against BASEL collection and isolated phage escapers with guidance by I.S. M.J. and K.K. cloned constructs. M.J. and R.G. produced and purified proteins. M.J. and R.G. performed nucleic acid cleavage assays. M.J., A.S. and A.M. prepared cryo-EM samples. A.S. and A.M. collected and processed cryo-EM data with input from P.P. and L.M. M.J., A.S. and P.P. built and refined structure models. J.J. sequenced phage escaper genomes and RNA and analyzed msDNA samples on the sequencing gel. M.J. and T.L. performed EOP experiments testing Eco2 variants for antiphage defense. W.S. performed the HDX-MS analysis. L.T. performed the BACTH experiments. S.K. produced CFEs with guidance by R.G., and R.G. performed CFE assays. All authors designed experiments and analyzed data. M.J. and P.P. wrote the paper and prepared figures with input from J.J., T.L., R.G., A.S., W.S., A.M., M.F.T.J, G.D., G.B., L.M. and I.S. The paper was reviewed and approved by all authors.

Competing interests

The authors declare no competing interests.

Additional information

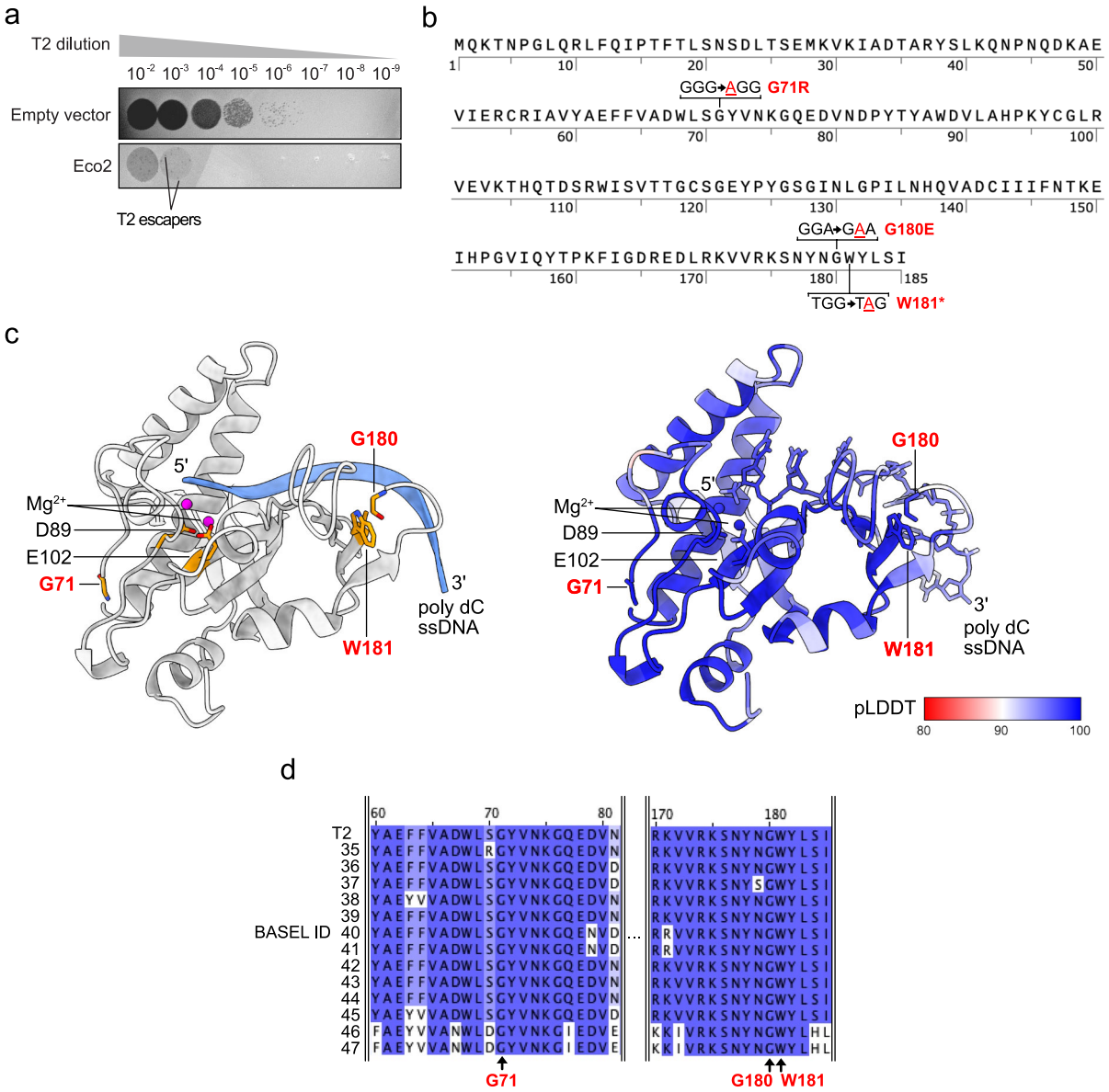
Extended data is available for this paper at https://doi.org/10.1038/s41594-026-01754-2.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-026-01754-2.

Correspondence and requests for materials should be addressed to P. Pausch.

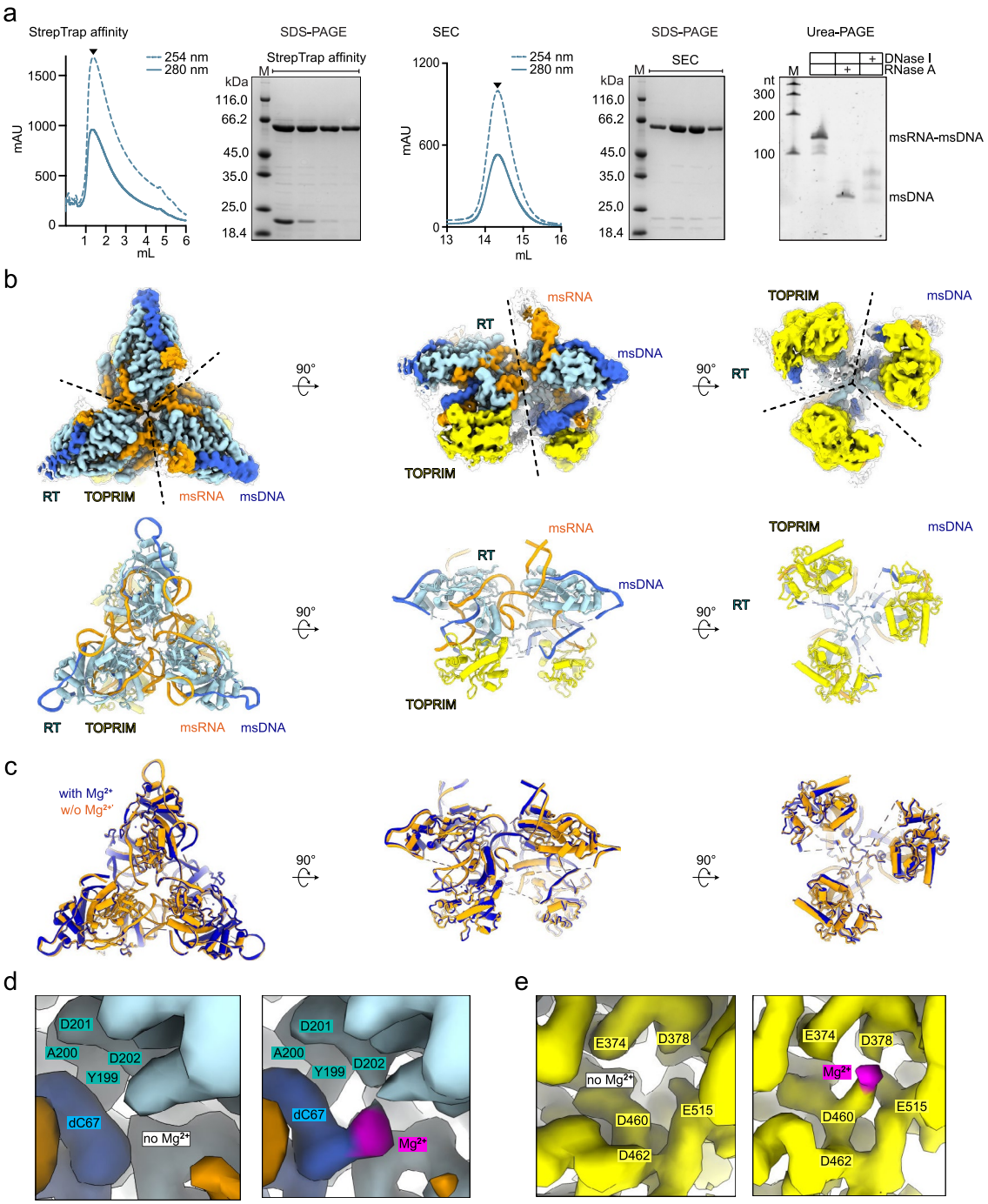
Peer review information *Nature Structural & Molecular Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available. Primary Handling Editor: Sara Osman, in collaboration with the *Nature Structural & Molecular Biology* team.

https://doi.org/10.1038/s41594-026-01754-2



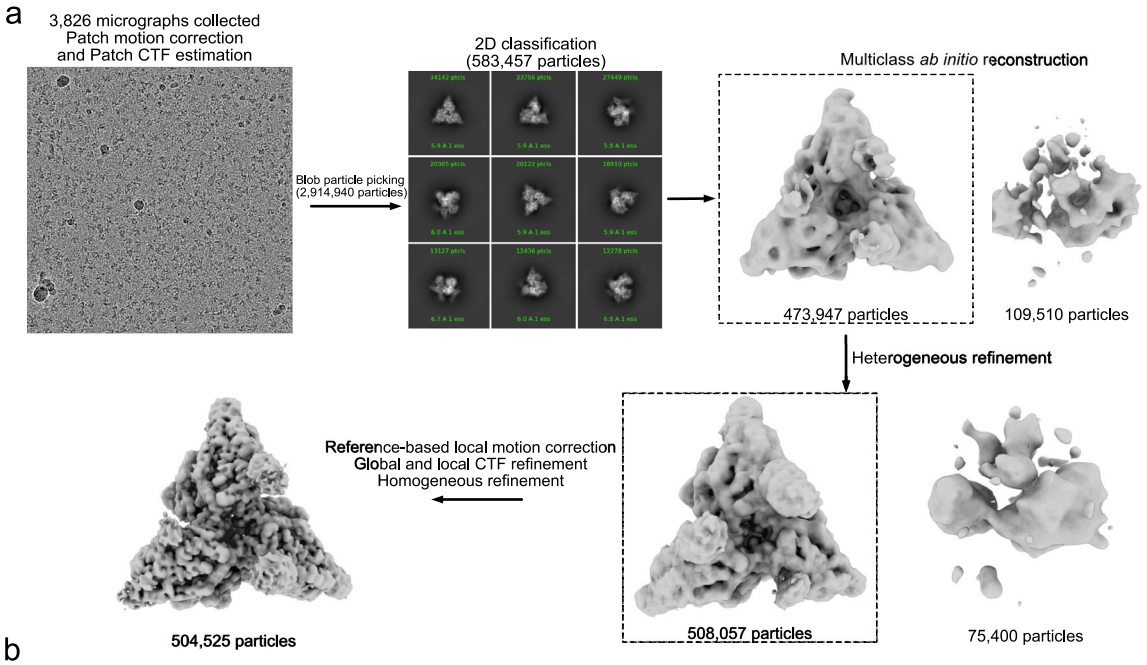
Extended Data Fig. 1 | T2 escapes by mutation of *denB*. **a**, Representative spot plating for T2 on *E. coli*, either in presence of an empty vector control (pACYC184) or a vector expressing Eco2 (pLG006). Escaper plaques are marked with lines. Representative data; $n \geq 2$ technical replicates. **b**, DenB protein sequence highlighting escaper mutations in bold red. Corresponding DNA mutations are shown. **c**, Left: AlphaFold3-predicted structure of T2-DenB (grey) with catalytic magnesium ions (purple spheres) and poly-C DNA (blue band). The amino acids mutated in the escaper variant are highlighted in orange and labeled in red. Right: pLDDT values mapped on the structure (color code), indicating prediction confidence (>90 highly confident). **d**, Multiple Sequence Alignment of T2 and BASEL phage collection-encoded DenB homologs, colored in blue according to identity. G71, G180 and W181 (red labels) are indicated by arrows.

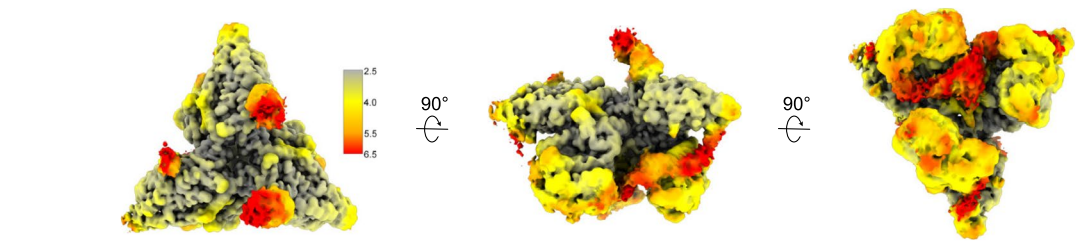
https://doi.org/10.1038/s41594-026-01754-2

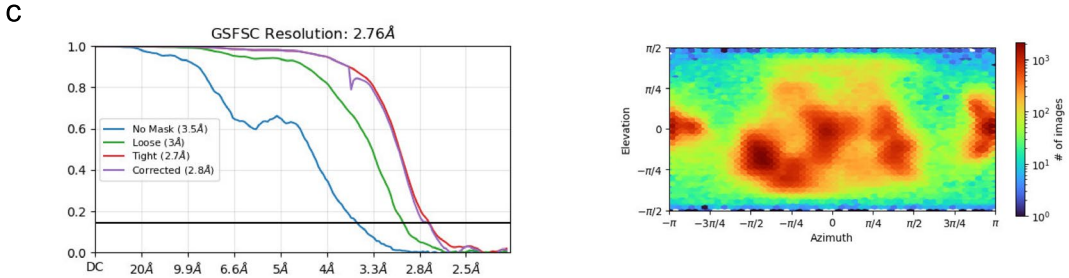


Extended Data Fig. 2 | Purification of Eco2 and cryo-EM structure in absence of magnesium. a, StrepTrap-affinity and SEC chromatograms (traces) with corresponding analytical SDS- and Urea-PAGEs. Representative data; $n \geq 2$ biological replicates. **b**, Cryo-EM maps (above) and corresponding model (below) of Eco2 trimers in absence of magnesium, shown in three 90°-rotations. Maps and models are color-coded according to Fig. 1, with unfiltered maps shown as translucent surfaces. **c**, Superimposition of Eco2 structures determined in absence (orange) and presence (blue) of magnesium, shown in three 90°-rotations. **d** and **e**, Close-up side-by-side comparison of the RT (**d**) and TOPRIM (**e**) active sites in absence (left) and presence (right) of magnesium (purple). The sharpened cryo-EM maps are shown.

https://doi.org/10.1038/s41594-026-01754-2

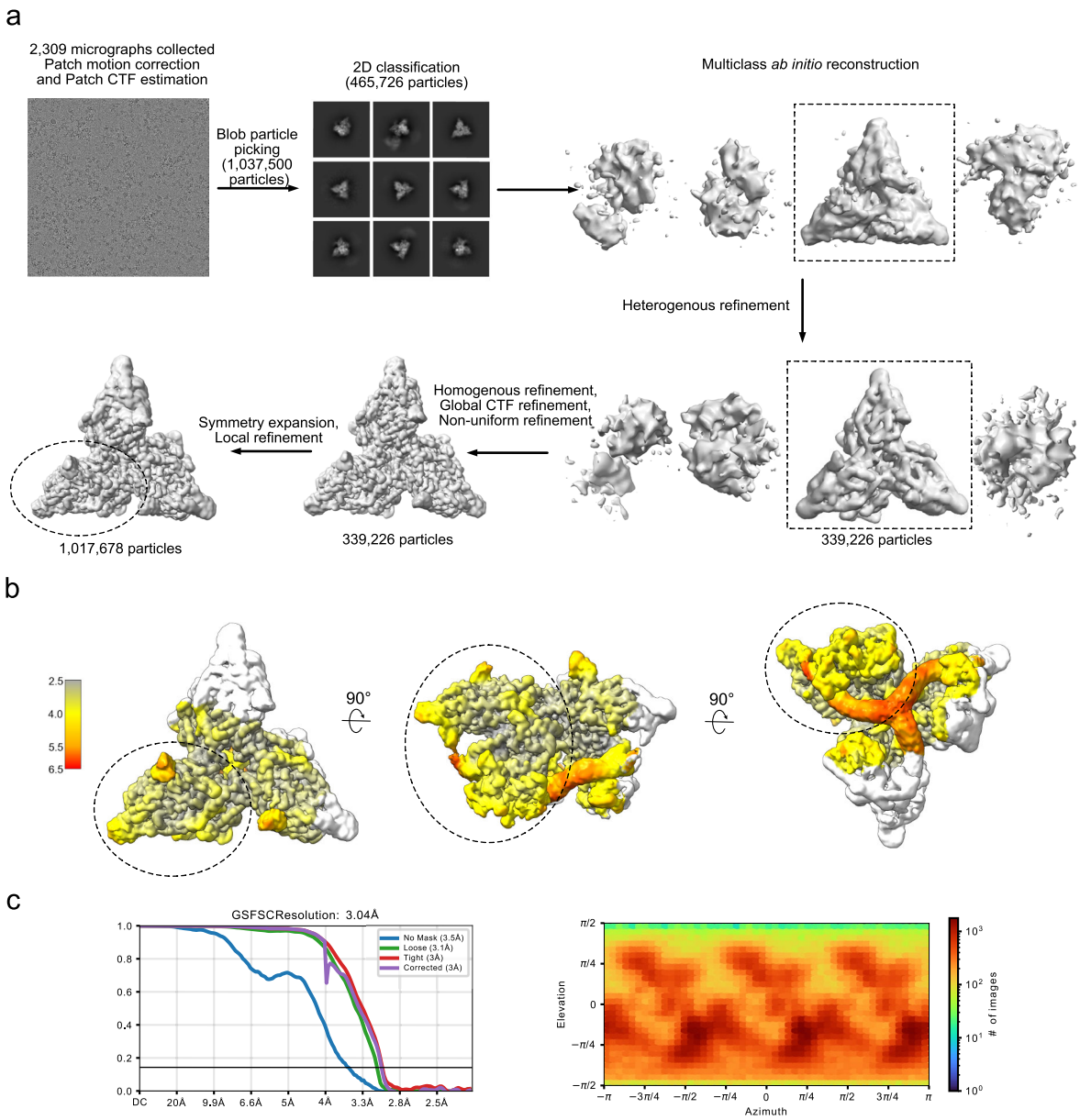






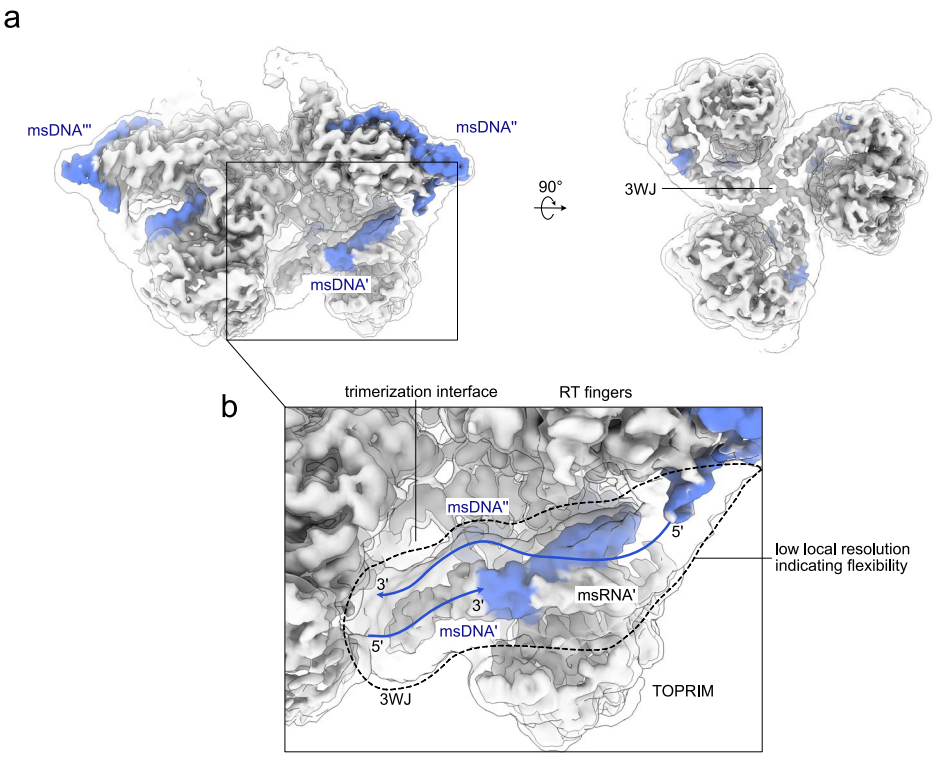
Extended Data Fig. 3 | Cryo-EM data processing for Eco2 in absence of magnesium. a, Processing workflow as performed in CryoSPARC. **b**, Local resolution map. **c**, FSC curves and angular distribution plots calculated in CryoSPARC.

https://doi.org/10.1038/s41594-026-01754-2



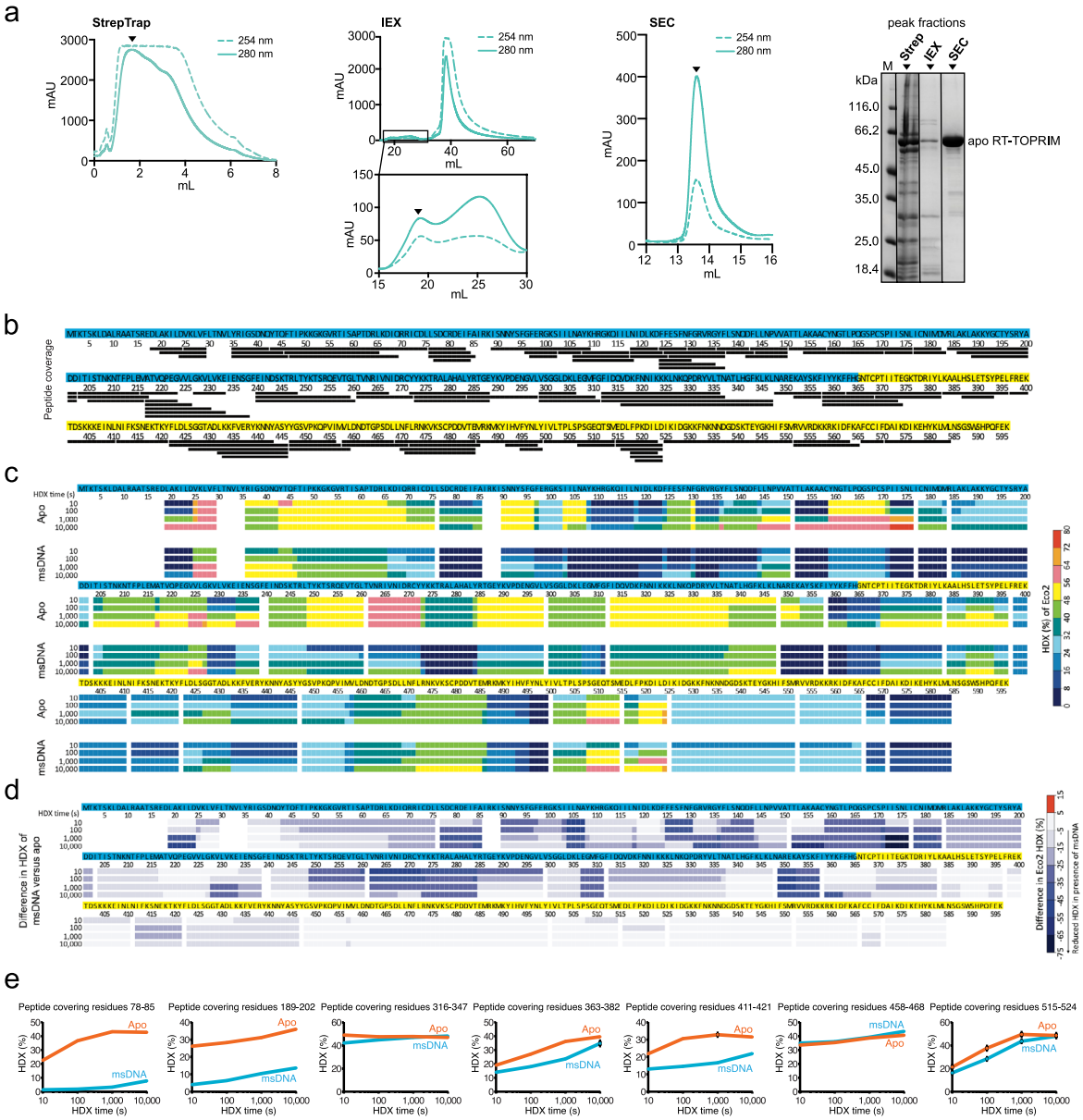
Extended Data Fig. 4 | Cryo-EM data processing for Eco2 in presence of magnesium. a, Processing workflow as performed in CryoSPARC. **b**, Local resolution map. Dashed outlines indicate the locally refined subunit. **c**, FSC curves and angular distribution plots calculated in CryoSPARC.

https://doi.org/10.1038/s41594-026-01754-2



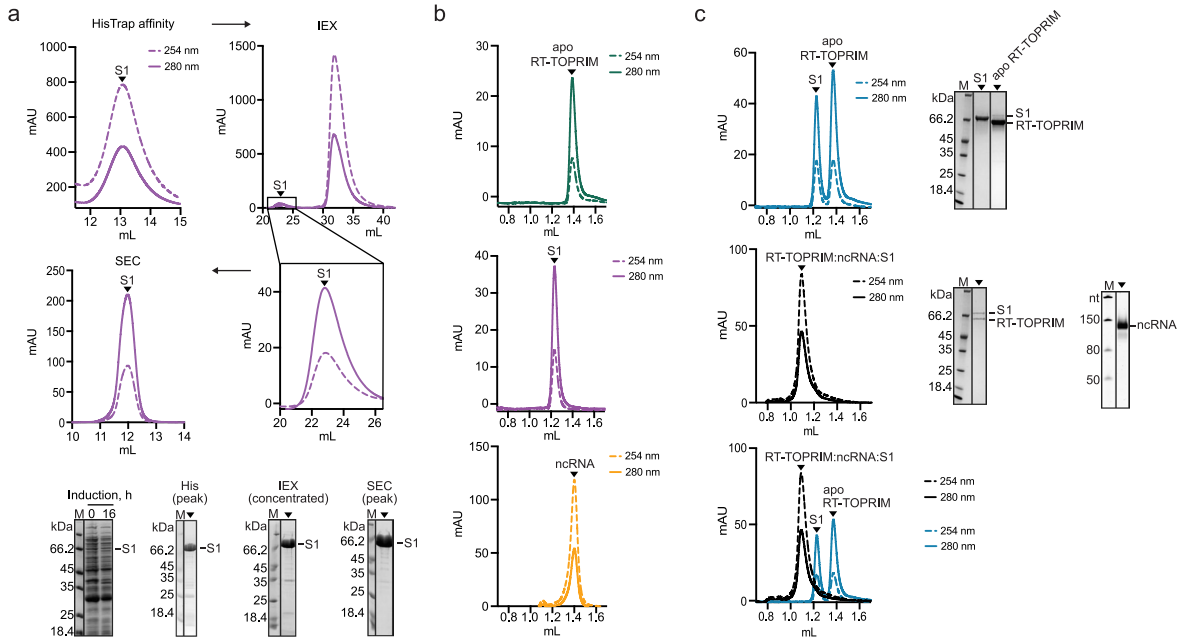
Extended Data Fig. 5 | Conformational flexibility of the msDNA. a, Cryo-EM maps of magnesium-bound Eco2 in two 90°-rotated orientations. The sharpened map is colored in gray and blue (msDNA). The unfiltered map is shown as a translucent surface. **b**, Close-up on the msDNA segments that lie adjacent to the trimerization interface. Blue arrows indicate msDNA-paths and polarity. The dotted outline highlights local low-resolution areas.

https://doi.org/10.1038/s41594-026-01754-2



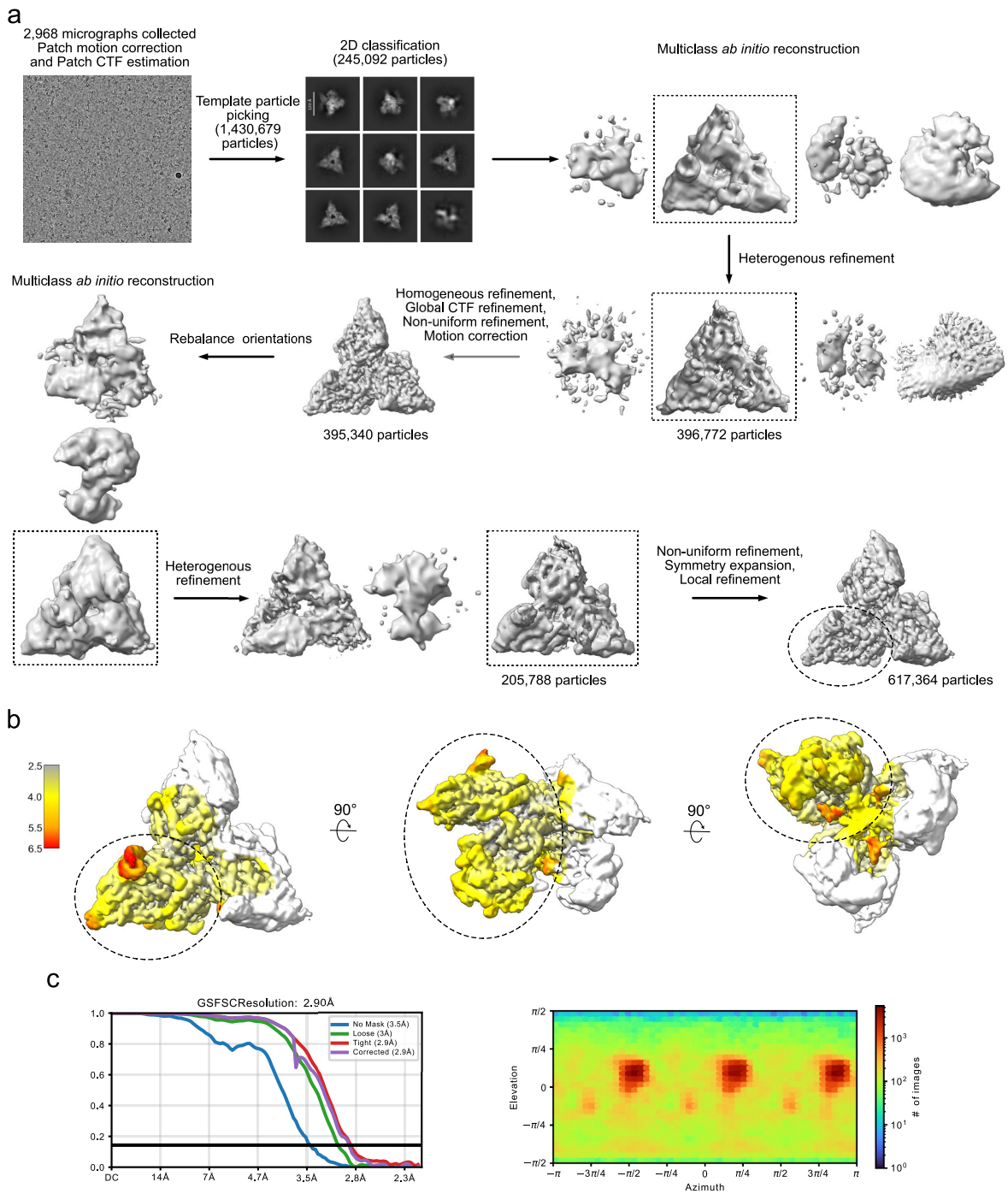
Extended Data Fig. 6 | HDX-MS of Eco2. a, StrepTrap affinity, ion-exchange (IEX) and SEC chromatograms (traces), as well as the corresponding analytical SDS-PAGEs of the apo RT-TOPRIM protein purification. Representative data; $n \geq 2$ biological replicates. **b**, HDX-MS peptide coverage of the RT-TOPRIM protein. Each black bar denotes a peptide identified in the HDX-MS experiments. Blue and yellow highlighted sequences correspond to the RT and TOPRIM domains, respectively. **c**, Residue-specific HDX of the apo state or when bound to msDNA, derived from HDX analysis of peptides from **b**. **d**, Difference in residue-specific HDX of the RT-TOPRIM between the msDNA-bound and apo states. Blue color denotes reduced HDX in presence of msDNA. **e**, Progression of HDX of selected representative peptides. Data represent the mean ± s.d. of $n = 3$ replicates.

https://doi.org/10.1038/s41594-026-01754-2



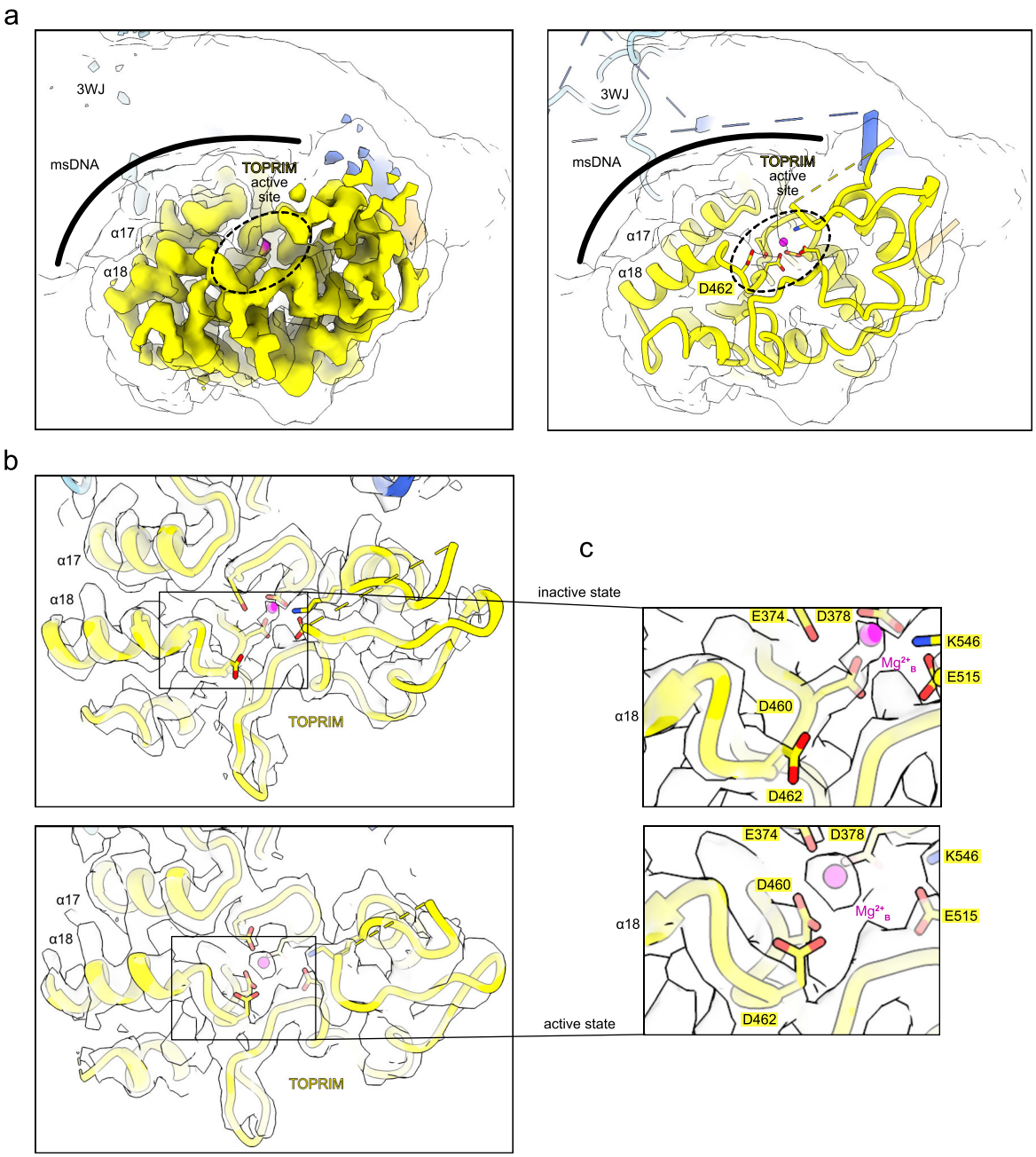
Extended Data Fig. 7 | *In vitro* reconstitution of the RT-TOPRIM:ncRNA:S1 complex. a, HisTrap affinity, ion-exchange (IEX), and size-exclusion chromatography (SEC) chromatograms, along with the corresponding analytical SDS-PAGEs, showing expression and purification of the apo S1 protein. Representative data; $n \geq 2$ biological replicates. **b**, Analytical SEC (S200, 3.2/300 column) traces of apo RT-TOPRIM (green trace), S1 (magenta trace), and the ncRNA (yellow trace; nts 1-118; a2-*msr-msd*). Representative data; $n \geq 2$ technical replicates. **c**, Analytical SEC of RT-TOPRIM and S1 proteins in absence (cyan trace) or presence (black trace) of the ncRNA. Right: analytical SDS- and urea-PAGEs; below: superimposition of both traces. Representative data; $n \geq 2$ technical replicates.

https://doi.org/10.1038/s41594-026-01754-2



Extended Data Fig. 8 | Cryo-EM data processing for activated Eco2. a, Processing workflow as performed in CryoSPARC. **b**, Local resolution map. Dashed outlines indicate the locally refined subunit. **c**, FSC curves and angular distribution plots calculated in CryoSPARC.

https://doi.org/10.1038/s41594-026-01754-2



Extended Data Fig. 9 | The msDNA 3WJ lies in proximity to TOPRIM α-helices 17 and 18. a, Left: Sharpened (colored) and unfiltered (translucent surface) cryo-EM maps of Mg^{2+}-bound Eco2 in the inactive, msDNA-bound state. The bold line highlights the msDNA:TOPRIM interface near TOPRIM α-helices 17 and 18. The dashed circle marks the TOPRIM active site. Right: Model (colored) and unfiltered (translucent surface) cryo-EM map of Mg^{2+}-bound Eco2 in the inactive, msDNA-bound, state. **b**, Sharpened cryo-EM maps (translucent surfaces) and TOPRIM models (cartoon) of the inactive- (above) and active- (below) state structures. **c**, Close-up views of the TOPRIM active site in its inactive (above) and active state (below).

nature portfolio

Corresponding author(s): Patrick Pausch

Last updated by author(s): Jan 17, 2026

Reporting Summary

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

Statistics

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

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Cryo-EM data was collected using EPU (versions 3.2 and 3.10) software (Thermo Fisher Scientific). Gel images were taken using the Image Lab Touch v3.0.0.07 (Biorad) software. UNICORN 7.8 (GE) was used for recording size exclusion traces. Radioactive gels were imaged using Amersham Typhoon v.2.0.0.6 (GE) software. Plate reader data was recorded using CLARIOstar plus plate reader software version 5.70 R2 (BMG Labtech). Long read RNA sequencing data was collected using Oxford Nanopore Technologies software Minknow v23.11. Short read RNA sequence data was obtained using NextSeq 2000, Control Software v1.7.1.46395. Phage genomes were sequenced using Illumina Miniseq Control Software v2.2.1.9. Protein sequences were collected by querying blastp PSI_BLAST on the NCBI clusteredNR database (2025-07-21), hosted at https://blast.ncbi.nlm.nih.gov/Blast.cgi. Protein sequences with accession numbers listed in Rodriguez-Mestre et al., (2020) or Millman et al., (2020) were downloaded by querying NCBI, EBI, and PATRIC databases (2024-11-28).

Data analysis

Cryosparc v.4.3-4.7 (Structura Biotech. Inc.), Phenix v.1.20.1-4487 and 1.21.2-5419, Coot v.0.9.8.7 and 0.9.8.95, and ChimeraX v.1.8 were used for cryo-EM data processing, model building and refinement, as well visualization/analysis.
Western blot images were analyzed with Fiji version 2.14.0/1.54f.
Data were plotted in Prism software version 10.5.0 (673)(Graphpad).
PLGS v3.0.1 (Waters) and DynamX v.3.0 (waters) were used for peptide identification and analysis of deuterium incorporation in HDX-MS experiments.
For ncRNA analysis, the following software was used: conda v23.7.4; pip v25.0; jupyterlab v3.6.6; dorado v1.0.0; samtools v1.21; minimap2 v2.24; python v3.10.16; pysam v0.22.1; matplotlib v3.8; numpy v1.26.4.
For the RT-TOPRIM phylogeny analysis, the following software was used: conda v25.1.1; pip v24.0; baltic v0.3.0; biopython v1.85; CD-HIT v4.8.1; hmmer v3.4; iqtree v2.3.6; jupyterlab v4.2.1; mafft v7.525; matplotlib v3.8.4; numpy v1.26.4; python v3.12.3.
For the deep sRNA-seq analysis, the following software was used: conda v4.12; pip v21.2.4; jupyterlab v4.3.4; python v3.13.5; bbmap v38.18; samtools v1.13; bwa-mem2 v2.2.1; pandas v2.3.1; pysam v0.23.3; matplotlib v3.10.3; biopython v1.85; numpy v2.3.1; matplotlib_venn v1.1.2.

For the escaper phage genome analysis, the following software was used: mamba v1.5.15; jupyterlab v4.4.5; python v3.12.3; bbmap v38.18; fastqc v0.12.2; multiqc v1.30; seqtk v1.4; breseq v0.37.1.
For the CFE sRNA-seq, the following software was used: conda v23.7.4; pip v25.0; minimap2 v2.17; samtools v1.13; python v3.10.16; pysam v0.22.1; pandas v2.2.3; numpy v1.26.4; biopython v1.81.
To identify DenB and A1 across the BASEL collection, the following software was used: conda v23.7.4; jupyterlab v3.6.6; hmmer v3.4.
Real data and code related to sequencing experiments and RT-TOPRIM phylogeny are available at: Juozapaitis, Jonas (2025), "Structure and mechanism of anti-phage retron Eco2, Jasnauskaite et al.", Mendeley Data, V1, https://data.mendeley.com/preview/v9b289v6ct.

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

Cryo-EM maps and model coordinates were deposited to the EMDB (52583, 52584, 54448) and PDB (9I2F, 9I2G, 9S1F). Cryo-EM maps, half-maps, models and PDB validation reports are available via Figshare "Retron Eco2 breaks tRNA for antiphage defense" by Jasnauskaitė et al., provided alongside the submission of this manuscript.
MS data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD060313 and PXD067212.
Data related to sequencing experiments and RT-TOPRIM phylogeny are available at: Juozapaitis, Jonas (2025), "Structure and mechanism of anti-phage retron Eco2, Jasnauskaite et al.", Mendeley Data, V1, https://data.mendeley.com/preview/v9b289v6ct. Raw sequencing reads generated from in vivo experiments were deposited in the NCBI Sequence Read Archive under BioProject accession number PRJNA1358792.
Protein sequences were accessed by querying blastp PSI_BLAST on the NCBI clusteredNR database (2025-07-21), hosted at https://blast.ncbi.nlm.nih.gov/Blast.cgi. Protein sequences with accession numbers listed in Rodriguez-Mestre et al., (2020) or Millman et al., (2020) were downloaded by querying NCBI, EBI, and PATRIC databases (2024-11-28).
Source data underlying figures is available alongside the manuscript.

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	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No sample size calculation was performed for all experiments. Experiments were performed in three independent replicates when data was quantified (generally accepted standard in the field).
Data exclusions	No data was excluded
Replication	Whenever data was quantified, experiments were repeated three times under identical conditions, either in parallel, or subsequently. All replication attempts were successful.
Randomization	Allocation was not random. Controls were included in all experiments and showed consistent results across different experiments, demonstrating that the results can be compared across sets of replicates.
Blinding	Investigators were not blinded. Controls were included in all experiments and showed consistent results across different experiments, demonstrating that the results can be compared across sets of replicates.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*