





Structural asymmetry governs the assembly and GTPase activity of McrBC restriction complexes

Yiming Niu ^{1,2,3,6}, Hiroshi Suzuki ^{2,4,6}, Christopher J. Hosford^{1,5}, Thomas Walz ^{2✉} & Joshua S. Chappie ^{1✉}

McrBC complexes are motor-driven nucleases functioning in bacterial self-defense by cleaving foreign DNA. The GTP-specific AAA + protein McrB powers translocation along DNA and its hydrolysis activity is stimulated by its partner nuclease McrC. Here, we report cryo-EM structures of *Thermococcus gammatolerans* McrB and McrBC, and *E. coli* McrBC. The McrB hexamers, containing the necessary catalytic machinery for basal GTP hydrolysis, are intrinsically asymmetric. This asymmetry directs McrC binding so that it engages a single active site, where it then uses an arginine/lysine-mediated hydrogen-bonding network to reposition the asparagine in the McrB signature motif for optimal catalytic function. While the two McrBC complexes use different DNA-binding domains, these contribute to the same general GTP-recognition mechanism employed by all G proteins. Asymmetry also induces distinct inter-subunit interactions around the ring, suggesting a coordinated and directional GTP-hydrolysis cycle. Our data provide insights into the conserved molecular mechanisms governing McrB family AAA + motors.

¹Department of Molecular Medicine, Cornell University, Ithaca, NY, USA. ²Laboratory of Molecular Electron Microscopy, The Rockefeller University, New York, NY, USA. ³Present address: Laboratory Molecular Neurobiology and Biophysics, The Rockefeller University, New York, NY, USA. ⁴Present address: Advanced Research Institute, Tokyo Medical and Dental University, Tokyo, Japan. ⁵Present address: New England Biolabs, Inc., Ipswich, MA, USA. ⁶These authors contributed equally: Yiming Niu, Hiroshi Suzuki. ✉email: twalz@rockefeller.edu; chappie@cornell.edu

Infections by antibiotic-resistant bacteria pose a serious threat to human health^{1,2}. The slow progress in developing new drugs to combat these emerging “superbugs” and the rapid exchange of resistance genes among microbial populations has intensified the need for alternative therapeutic strategies³. One such strategy employs bacteriophages (phages)—viruses that infect a bacterial host, replicate, and then induce cell lysis to release the mature phage progeny, killing the host in the process⁴. The pharmaceutical application of phages dates back to the early 1920s⁵ and has resurged in recent years, bolstered by success in a number of clinical settings^{6,7}. Despite these promising results, phage therapy faces numerous challenges. One significant hurdle is that bacteria have evolved an array of defense mechanisms, including restriction modification systems, modification-dependent restriction systems (MDRSs), phage-exclusion systems, and CRISPR-Cas adaptive immune systems, that can hinder phage infection and diminish their subsequent killing potential^{8,9}. These machineries lack eukaryotic homologs and are conserved across antibiotic-resistant bacteria like methicillin-resistant *Staphylococcus aureus*, *Clostridium difficile*, and *Klebsiella pneumoniae*, making their components promising candidates for targeted inhibition. Some phages indeed already encode inhibitor proteins that can neutralize restriction and/or CRISPR systems^{10,11}, allowing them to survive and kill under conditions in which they would normally be suppressed. Elucidating the structure and function of bacterial defense systems will therefore extend these principles and aid in the development of new drugs that increase phage efficacy.

McrBC is a two-component MDRS that in *Escherichia coli* (Ec) restricts phage DNA and foreign DNA containing methylated cytosines^{12,13}. EcMcrB consists of an N-terminal DNA-binding domain that targets fully or hemi-methylated R^{MC} sites (where R is a purine base and ^{MC} is a 4-methyl-, 5-methyl-, or 5-hydroxymethyl-cytosine)^{14–19}, and a C-terminal AAA+ (extended ATPases Associated with various cellular Activities) domain that hydrolyzes GTP and oligomerizes into hexamers^{20,21}. EcMcrB's basal GTPase activity (~ 0.5 – 1 min^{-1}) is stimulated ~ 30 – 40 -fold in vitro via interaction with its partner EcMcrC¹⁶, a PD-(D/E)xK family endonuclease that cannot stably bind DNA on its own, and thus associates with the hexameric McrB AAA+ ring²¹. Biochemical data suggest that stimulated GTP hydrolysis powers DNA translocation^{18,22}, allowing EcMcrBC complexes bound to distant R^{MC} sites to interact and induce cleavage on both strands^{23,24}. While these activities have yet to be demonstrated in vitro for homologs beyond *E. coli*, other family members have also been shown to function in bacterial defense in vivo^{25–27}. These machines, however, exhibit different specificities for DNA modifications and/or sequences^{25,27–30}, suggesting that the core machinery for GTP hydrolysis and DNA cleavage is conserved and has been adapted to different targets throughout evolution in response to various selective pressures from invading phages. This flexibility holds a tremendous potential for engineering new endonucleases for biotechnology and biomedical applications, providing further motivation to study the structural organization and functional regulation of McrBC complexes.

AAA+ proteins are large, multimeric machines that use the energy of ATP hydrolysis to power a wide array of cellular processes³¹. These enzymes are built around a common structural core³² and contain numerous conserved sequence elements important for nucleotide binding and hydrolysis³³. AAA+ protein active sites are formed at the interface between two monomers, thus requiring higher-order assembly—predominantly as hexamers—for function³⁴. As a consequence, some catalytic residues like charge-compensating arginine fingers are provided in *trans* by the neighboring subunit. Despite sharing a common

architecture, McrB is the only AAA+ protein that preferentially binds and hydrolyzes GTP^{35,36}. All McrB homologs contain a conserved consensus sequence of MNxxDRS that replaces the AAA+ sensor I motif and is predicted to function as a G4 element, which confers guanine nucleotide specificity in GTPases³⁷. Mutation of this segment, however, does not significantly alter the nucleotide-binding profile of *E. coli* McrB^{16,36}, indicating that other regions of the protein dictate GTP selectivity. Stimulation of hydrolysis by a binding partner is rare among AAA+ proteins^{38,39}, but reminiscent of the activation of small GTPases by their corresponding GTPase-activating proteins (GAPs)⁴⁰. A key difference from other GTPases in this instance, however, is that the second component McrC only exerts its effects on the assembled McrB oligomer. Elucidating the structural basis for GTP recognition and stimulated hydrolysis is important for defining McrBC's divergence from other members of both the AAA+ and GTPase superfamilies.

A recent cryo-EM reconstruction of the hexameric EcMcrB AAA+ domain bound to EcMcrC at 3.6-Å resolution⁴¹ provided the first glimpse of this machine, showing the overall architecture of the complex and proposing a general mechanism for catalytic turnover. However, this study did not resolve the molecular details and chemistry underlying the GTP hydrolysis reaction and its stimulation by McrC and may not have identified the correct DNA-binding mode. Here, we present cryo-EM structures of an McrB hexamer and McrBC complexes from the evolutionarily distant archaeal species *Thermococcus gammatolerans* (Tg) and the well-characterized *E. coli* system. Our models confirm that McrBC complexes share the same general architecture, but lead to a different view of the GTP hydrolysis cycle wherein structural asymmetry drives the underlying physical interactions and conformational motions. Moreover, our structures provide a detailed molecular mechanism for how McrC-binding stimulates McrB GTP hydrolysis, which we show is conserved across the McrBC family. Our structures also establish that McrB homologs use the same general chemistry employed by all GTPases to recognize GTP, albeit through different structural elements upstream of the AAA+ domain. This observation establishes how distant McrB homologs have adapted and maintained guanine nucleotide specificity despite the individual constraints imposed by their structurally unrelated N-terminal domains. Together these data provide mechanistic insights into the structure, function, and regulation of motor-driven McrBC nucleases.

Results

TgMcrB^{AAA} forms an asymmetric hexamer. Given the widespread distribution of *mcrBC* genes among diverse bacteria and archaea, we sought to examine the structural and biochemical properties of different McrB homologs to understand how these AAA+ enzymes have evolved to preferentially bind and hydrolyze GTP. Our previous work identified the archaeal McrB homolog from *T. gammatolerans* (TgMcrB) as an ideal candidate for structural studies given its compact size and increased thermostability²⁹. The purified AAA+ domain from TgMcrB (TgMcrB^{AAA}) forms stable oligomers even in the absence of nucleotides (Supplementary Fig. 1a). Single-particle cryo-EM analysis of purified TgMcrB^{AAA} incubated with the non-hydrolyzable GTP analog GTP γ S yielded a density map at an overall resolution of 3.1 Å with no symmetry imposed (Fig. 1 and Supplementary Fig. 1). The cryo-EM map reveals that TgMcrB^{AAA} forms a ring-shaped, homo-hexameric assembly with six nucleotides bound at the subunit interfaces, similar to the closed-ring assembly seen in type I AAA ATPases^{42,43}. Each subunit displays a canonical AAA+ fold with the additional features of a β -hairpin inserted in helix 2 of the large subdomain

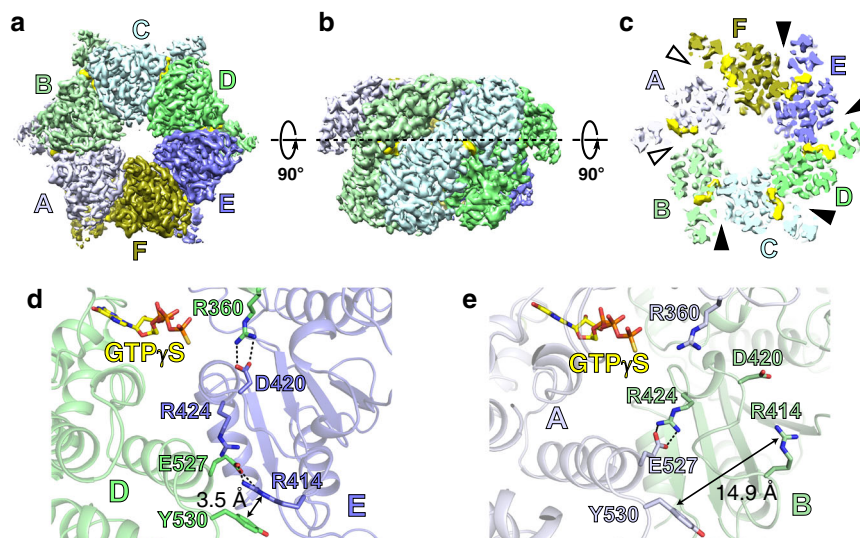


Fig. 1 TgMcrB^{AAA} forms an asymmetric hexamer. **a, b** Bottom and side views of the cryo-EM density map of the TgMcrB^{AAA} hexamer. Subunits are colored in shades of blue and green, and nucleotides are shown in yellow. **c** Slice section through the TgMcrB^{AAA} hexamer map at the level of the bound nucleotides, indicated by the dashed line in **b**. Solid and empty arrowheads indicate tight and loose interfaces, respectively. **d, e** Close-up views of interacting residues at the tight D/E interface (**d**) and the loose A/B interface (**e**). Dashed lines indicate hydrogen bonds.

as previously predicted⁴⁴ and “wing”-like helices in the small subdomain (Supplementary Figs. 2 and 3). The TgMcrB^{AAA} hexamer is asymmetric with four tight interfaces (between monomers B/C, C/D, D/E, and E/F) that bury a surface area ranging from 2393 to 2554 Å², and two loose interfaces (between monomers A/B and F/A) that bury surface areas of 1519 and 1772 Å² (Fig. 1c). Tight interfaces feature a hydrogen bond between Asp420 in one monomer and Arg360 in the adjacent monomer (Fig. 1d). Arg414 from the first monomer also extends into the neighboring monomer, where it forms hydrogen bonds with Glu527 and π -stacking interactions with Tyr530 (Fig. 1d). These interactions are absent at the loose interfaces, where Glu527 instead interacts in *trans* with Arg424 (Fig. 1e). All of these residues are highly conserved amongst McrB family proteins (Supplementary Fig. 2).

To determine if these interface residues affect McrB’s catalytic turnover, we mutated each side chain individually to alanine in the context of TgMcrB^{AAA} and measured basal GTPase activity using a colorimetric assay. All mutants show an approximate twofold increase in hydrolysis activity compared to that of the wild-type protein (Supplementary Fig. 1i). Alanine substitution of Arg337 in EcMcrB (corresponding to Arg414 in TgMcrB, Supplementary Fig. 3) was previously shown to increase the basal GTPase rate threefold¹⁶, consistent with our results.

In parallel, we also determined the structure of TgMcrB^{AAA} in the presence of GTP γ S by X-ray crystallography at 2.95-Å resolution (Supplementary Table 3). Symmetry-related hexamers abut against each other in the crystal lattice, deforming the planar arrangement of the six subunits in each molecule (Supplementary Fig. 1j, k). This produces an “open-ring” conformation in which the subunits at the loose A/B interface are significantly splayed apart, and the small subdomain of the F subunit becomes highly disordered (Supplementary Fig. 1k, l). The individual TgMcrB^{AAA} monomers, however, adopt the same overall conformation and organization of nucleotide binding, as is observed in the cryo-EM reconstruction (Supplementary Fig. 1m–o). The distorted appearance of the crystallographic hexamer suggests a greater flexibility at the loose interfaces, which could more readily be influenced by crystal packing forces.

TgMcrB^{AAA} contains the complete machinery for nucleotide hydrolysis. Nucleotide hydrolases harness the energy of ATP or GTP hydrolysis to catalyze energetically unfavorable biological reactions, coordinate signal transduction events, and power protein conformational changes that orchestrate a multitude of cellular processes⁴⁵. Efficient hydrolysis requires (i) the binding and recognition of the appropriate nucleotide substrate, (ii) the correct positioning of a water molecule for an in-line S_N2 attack on the γ -phosphate to initiate cleavage of the phosphoanhydride bond, and (iii) neutralization of a negative charge that develops between the β - and γ -phosphates in the transition state⁴⁶.

While a conserved sequence motif of GxxGxGK[T/S] (P-loop/Walker A motif) coordinates the α - and β -phosphates in both ATPases and GTPases⁴⁷, the remaining catalytic machinery, specificity determinants, and charge-compensating elements vary from enzyme to enzyme. AAA+ proteins contain four additional sequence motifs—Walker B, Sensor I, Sensor II, and second region of homology (SRH)—that contribute to ATP binding and hydrolysis along with the conserved P-loop/Walker A motif^{33,48}. The Walker B motif (D[D/E]xx) stabilizes an essential magnesium cofactor and acts in concert with a polar residue in the Sensor I motif to orient the catalytic water for nucleophilic attack on the γ -phosphate. The Sensor II motif localizes to helix 7 and contains a conserved arginine that interacts with the γ -phosphate. By convention, the subunit contributing these structural motifs to the nucleotide-binding pocket is referred to as the *cis* subunit. The neighboring, *trans* subunit inserts the arginine finger at the end of helix 4 of the SRH into the nucleotide-binding pocket, where it stabilizes the γ -phosphate and contributes to the charge compensation in the transition state.

Each composite active site of TgMcrB^{AAA} contains one GTP γ S molecule and a bound magnesium ion (Fig. 2a–c). In the *cis* subunit, the main-chain atoms of the Walker A motif interact with the α - and β -phosphates, and Lys221 contacts the γ -phosphate of GTP γ S (Fig. 2a, b). Thr222 (Walker A) and Asp356 (Walker B) coordinate the magnesium cofactor along with two ordered water molecules (Fig. 2a). Mutation of these conserved side chains to alanine impairs the basal GTPase activity of TgMcrB^{AAA} (Fig. 2c, d). Glu537 (Walker B) lies in close

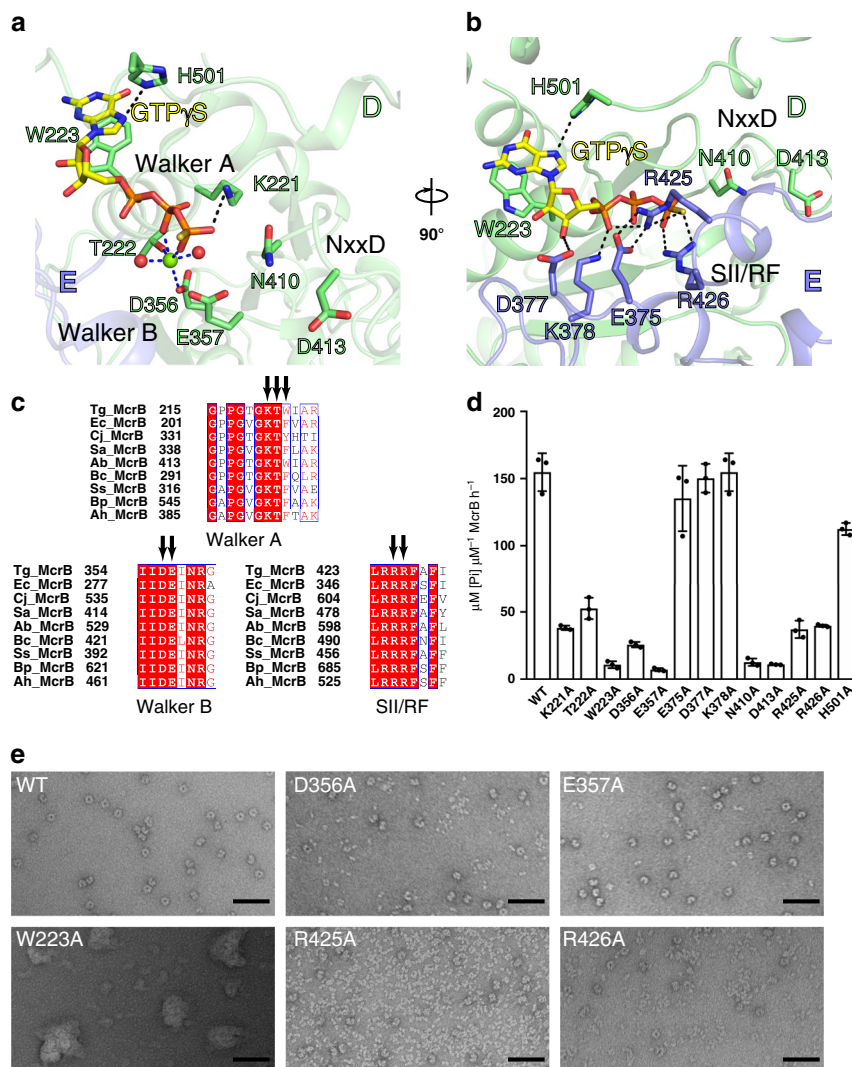


Fig. 2 Catalytic residues involved in the basal GTPase activity of TgMcrB. **a, b** Close-up views of the GTP-binding site at the tight D/E interface, highlighting residues involved in *cis* interactions, in particular those of the Walker A and B motifs, and the NxxD motif (**a**), and residues involved in *trans* interactions, in particular those of the Sensor II/arginine finger (SII/RF) motif (**b**). Spheres indicate waters (red) and a magnesium ion (green). Dashed lines indicate hydrogen bonds (black) and metal coordination (blue). **c** Sequence alignment of McrB homologs for the classic Walker A and B motifs, and the SII/RF motif. Arrows indicate the catalytic residues. Sequence alignment abbreviations are as follows: Tg *Thermococcus gammatolerans*, Ec *Escherichia coli*, Cj *Campylobacter jejuni*, Sa *Staphylococcus aureus*, Ab *Aciduliprofundum boonei*, Bc *Bacillus cereus*, Ss *Streptococcus suis*, Bp *Butyrivibrio proteoclasticus*, Ah *Anaerobutyricum hallii*. **d** Basal GTPase activity of wild-type TgMcrB and alanine mutants at the residues shown in **a** and **b** ($n = 3$, mean \pm standard deviation). **e** Selected micrograph areas of negatively stained wild-type and mutant TgMcrB^{AAA} incubated with GTPyS. Scale bars are 50 nm.

proximity to the γ -phosphate, primed to help stabilize a catalytic water (Fig. 2a). An alanine substitution at this position completely abolishes hydrolysis activity (Fig. 2d). Negative stain EM indicates that the Asp356Ala mutation has a higher propensity to disrupt the TgMcrB^{AAA} hexamer than the Glu357Ala mutation (Fig. 2e), consistent with their distinct functions in nucleotide binding/stabilization versus catalysis. This result mirrors the oligomerization defects observed in EcMcrB when the corresponding residues (Asp279 and Glu280) were mutated²⁰, suggesting that the aspartate residue functions in magnesium binding, while the glutamate residue is critical for coordinating the catalytic water as in other AAA+ proteins³³. Notably, the conserved McrB consensus loop (⁴⁰⁹MNxxDR⁴¹⁴) replaces Sensor I and is located close to the γ -phosphate (Fig. 2 and Supplementary Fig. 2). Asn410Ala and Asp413Ala mutants significantly impair basal GTPase activity (Fig. 2d), suggesting they are critical for catalytic turnover, rather than for nucleotide binding, as was previously predicted³⁵.

His501 and Trp223 in the *cis* subunit sandwich the guanine base of GTPyS (Fig. 2a). His501 is situated above and forms a hydrogen bond with the 7' nitrogen. Trp223, which lies adjacent to the Walker A motif, forms an unusual parallel π -stacking interaction from below that is absent in the majority of both GTPases and AAA+ proteins^{44,49}, but has recently been observed in the YCJK stress protein⁵⁰. Mutation of Trp223 to Ala completely abolishes the basal GTPase activity (Fig. 2d) and causes the protein to aggregate, as seen by negative stain EM imaging (Fig. 2e). These observations indicate that π -stacking is critical for both McrB GTP binding and the stability of the oligomeric assembly. Although the aromatic residue is not strictly conserved across the McrB family, every homolog contains a residue at this position that is capable of π -stacking (Trp, Phe, Tyr, or Arg; Fig. 2c and Supplementary Fig. 2), including Phe209 in EcMcrB⁴¹.

The *trans* subunit also contributes numerous conserved side chains that stabilize different portions of the bound nucleotide.

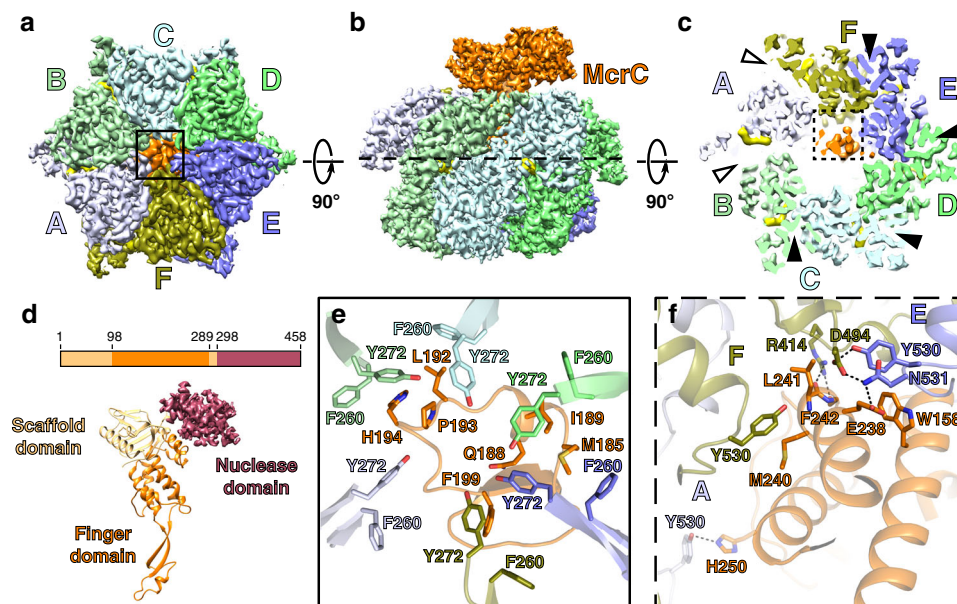


Fig. 3 Asymmetric assembly of the TgMcrBC complex. **a, b** Bottom and side views of the cryo-EM density map of the TgMcrBC half-complex. TgMcrB subunits are colored as in Fig. 1, TgMcrC is shown in orange, and nucleotides in yellow. **c** Slice section through the map of the TgMcrBC half-complex at the level of the bound nucleotides, indicated by the dashed line in **b**. Solid and empty arrowheads indicate tight and loose interfaces, respectively. **d** Domain architecture of TgMcrC. **e** Close-up view of the interactions of TgMcrC with TgMcrB at the bottom of the hexamer, indicated by the black square in **a**. **f** Close-up view of the interactions of TgMcrC with the TgMcrB hexamer at the E/F and F/A interfaces, indicated by the dashed black square in **c**. Dashed lines indicate hydrogen bonds.

Asp377 interacts with the 3' ribose hydroxyl group, while Glu375 and Lys378 coordinate the α -phosphate (Fig. 2b and Supplementary Fig. 2). Mutations of these side chains had negligible effects on basal GTPase activity (Fig. 2d). Arg426 in helix α 11 acts as the charge-compensating arginine finger, here forming hydrogen bonds with the γ -phosphate in the ground state (Fig. 2b). A second neighboring arginine located on the same helix, Arg425, assumes the role of the missing Sensor II motif (Fig. 2b, c). Arg425Ala and Arg426Ala mutations impair basal GTPase activity, and disrupt hexamer formation (Fig. 2d, e). All the *trans* interactions with GTP γ S are prominent at the tight interfaces but are lost at the loose interfaces. Since the cryo-EM density for the Arg side chains in the Sensor II/arginine finger motif are also weaker at the loose interfaces (Supplementary Fig. 1o), the GTP-binding sites at these locations are likely in a non-catalytic state. Taken together, these results indicate that TgMcrB^{AAA} possesses all the critical residues needed to bind and hydrolyze GTP.

TgMcrB and TgMcrC form an asymmetric complex. We next sought to elucidate structural and biochemical consequences of TgMcrC binding to TgMcrB. Purified TgMcrC was very sensitive to buffer conditions and could only be concentrated in the presence of TgMcrB. Together full-length TgMcrB and TgMcrC formed stable, dumbbell-shaped complexes in the presence of GTP γ S that were suitable for structure determination by cryo-EM (Supplementary Fig. 4a). Initial image processing showed that the complex consists of two TgMcrB hexamers connected through TgMcrC dimerization (Supplementary Fig. 4b). Because of structural variability, however, we were only able to refine a “half”-complex (Supplementary Fig. 4c), which yielded a map at an overall resolution of 2.4 Å (Fig. 3 and Supplementary Fig. 4d–g). In this reconstruction, a single TgMcrC binds the TgMcrB hexamer by inserting itself through the central pore of the AAA+ ring in an asymmetric fashion (Fig. 3a–c).

The resolution of our reconstruction allowed us to build the TgMcrC structure *de novo*. Each monomer contains a scaffold

domain, a “finger domain” and a C-terminal endonuclease domain (Fig. 3d and Supplementary Fig. 5). The scaffold domain (residues 1–98 and 289–298) consists of a barrel-like structure that centrally positions the two flanking domains, forming a rigid connection between the finger and endonuclease domains. The finger domain (residues 99–288) adopts an extended, segmented structure with two antiparallel helices that contact the nuclease domain above, a helical bundle, and a long β -sheet “stalk” that protrudes downward, terminating in a loop-helix-loop region at the tip (Fig. 3d and Supplementary Fig. 6a). The C-terminal endonuclease domain (residues 299–458) rests atop the structure and though poorly resolved in our map exhibits a fold characteristic of PD-(D/E)xK family enzymes.

The finger domain spans the entire length of the hexamer and its binding interface changes along the axis of the central pore (Supplementary Fig. 6a–e). At the top of the ring, the helical bundle associates with the F and E subunits, and then tilts to contact the E and D subunits near the middle of the assembly (Fig. 3c and Supplementary Fig. 6a–c). We also observe interactions between the β -sheet stalk and the E subunit at this midpoint (Supplementary Fig. 6a, d). The loop-helix-loop at the distal tip of the finger domain plugs a narrow opening at the very bottom of the McrB hexamer (Fig. 3a and Supplementary Figs. 5 and 6a, e). Conserved aromatic residues Phe260 and Tyr272 from the helix 2 inserts of each McrB subunit surround and stabilize the tip (Fig. 3e and Supplementary Fig. 2). While the finger domain interacts with all six subunits of TgMcrB at the bottom of the hexamer, TgMcrC binds the hexamer in a highly asymmetric fashion.

TgMcrC binding breaks the parallel π -stacking interaction between Arg414^F and Tyr530^E at the E/F interface (Fig. 3c, f), which has the smallest interaction area among the four tight interfaces (\sim 2400 Å² versus $>$ 2500 Å² for all the others). This perturbation changes the conformation of the 414–420 loop in subunit F as Arg414^F rotates to hydrogen bond with the main-chain atoms of Leu241^{McrC} and Phe242^{McrC} (Fig. 3f and

Supplementary Fig. 6h, i). Concomitantly, Tyr530 and Asn531 in subunit E hydrogen bond to Asp494 in subunit F. Glu238 in the finger domain further stabilizes this conformation through an additional hydrogen bond with Asp494^F. TgMcrC binding also generates some additional interactions in the F/A interface, where His250^{McrC} hydrogen bonds with Tyr530^{McrB} from the A subunit, and Met240^{McrC} and Leu241^{McrC} form van der Waals interactions with Tyr530^{McrB} in the F subunit (Fig. 3f and Supplementary Fig. 5). These interactions, which bury a combined surface area of 1298 Å², serve to anchor McrC at the top of the ring, restricting its motion and orientation. Despite the localized differences at the E/F interface, the conformation of the TgMcrB hexamer remains largely unchanged in the TgMcrBC complex (overall RMSD of 0.75 Å compared to TgMcrB^{AAA} alone, based on corresponding C α atoms), with its intrinsic asymmetry, and the remaining tight and loose interface interactions preserved (Fig. 3c and Supplementary Fig. 6f, g). These findings indicate that TgMcrC does not induce substantial remodeling of the TgMcrB hexamer, but instead adapts and exploits its intrinsic asymmetry when binding.

TgMcrC binding optimally positions existing catalytic machinery to stimulate GTP hydrolysis. A distinguishing feature of the *E. coli* McrBC system is the ability of McrC to stimulate McrB's GTP hydrolysis in vitro^{35,51}. Purified TgMcrC similarly stimulates TgMcrB's basal GTPase activity, demonstrating that this is also a conserved property of other homologs (Fig. 4d). Our high-resolution TgMcrBC structure reveals the underlying molecular mechanism governing this stimulation. As a consequence of the structural asymmetry imposed by the TgMcrB hexamer, TgMcrC's finger domain engages only a single active site at a time (Fig. 4a). Here, the helical bundle wedges against the D/E interface and inserts a highly conserved arginine (Arg263^{McrC}) at the edge of the pocket (Fig. 4b, c and Supplementary Fig. 6a, c). Acting through a hydrogen-bonding network that includes Asn359^{McrB}, Asn410^{McrB}, Asp413^{McrB}, and a bridging water (H₂O^{Bridge}), Arg263^{McrC} ultimately alters the conformation of the McrB consensus loop (Supplementary Fig. 2). This reorganization allows Asn410 and E357 of the Walker B motif to position a second water (H₂O^{Cat}) that is poised for nucleophilic attack on the γ -phosphate (Fig. 4b). Glu357 also acts in concert with the Asp356 of the Walker B motif to stabilize a third water molecule that completes the octahedral coordination of the magnesium cofactor (Fig. 4b).

Alanine substitutions at Asn410 and Asp413 in full-length TgMcrB abolish both basal and McrC-stimulated GTPase activity (Fig. 4d), underscoring their crucial catalytic function. Mutation of Arg263^{McrC} to alanine selectively abrogates the stimulatory effect of McrC binding without impairing basal turnover (Fig. 4d). The apparent GAP function thus arises from an indirect reconfiguration of the side chains that orient the catalytic water rather than promoting charge compensation in the transitions state.

Sequential rearrangements of the consensus loop control the cycle of McrB GTP hydrolysis. The consensus loop and charge-compensating arginine finger (Arg426^{trans}) adopt different conformations at each of the six interfaces within the McrC-bound TgMcrB hexamer (Fig. 4e–j). As described above, the tight D/E interface shows an McrC-activated conformation with Arg426^{trans} stabilizing the γ -phosphate and Asn410 properly arranged to orient the catalytic water (Fig. 4b, e). In the adjacent tight E/F interface, Asn410 and Arg426^{trans} appear in close contact with the γ -phosphate of GTP γ S in a manner that excludes a potential catalytic water (Fig. 4f). The loose F/A interface uniquely contains

GDP with the side-chain oxygen of Asn410 forming a hydrogen bond with the β -phosphate. This partially occludes the space normally occupied by the γ -phosphate and forces Arg426^{trans} into a conformation in which it is angled away from the nucleotide (Fig. 4g). At the loose A/B interface, Arg426^{trans} and Asn410 both point away from GTP γ S, likely a consequence of the weakened inter-subunit interactions (Fig. 4h). In both the tight B/C and C/D interfaces, Arg426^{trans} interacts with the γ -phosphate, but Asn410 faces away from the nucleotide (Fig. 4i, j). These pockets appear primed for hydrolysis but unable to proceed efficiently, as Asp413 adopts random orientations in the absence of McrC, and thus cannot help stably redirect Asn410 to position the catalytic water (Fig. 4i, j).

These conformational differences likely reflect different states in the hydrolysis cycle, with the B/C and C/D active sites occupying a GTP-bound, pre-hydrolysis state, D/E most likely the activated transition state, E/F assuming a post-hydrolysis state, and the loose GDP-bound F/A and GTP γ S-bound A/B sites, depicting the phosphate release and subsequent nucleotide exchange steps, respectively. Together these data imply that TgMcrB GTP hydrolysis proceeds through a coordinated, sequential mechanism.

McrBC homologs share a conserved architecture and catalytic mechanism. To establish whether different homologs use a conserved mechanism for stimulated hydrolysis, we determined the single-particle cryo-EM structure of the complex formed by the full-length *E. coli* proteins (EcMcrBC) in the presence of GTP γ S. EcMcrBC also formed dumbbell-shaped particles and we refined a “half”-map reconstruction of these assemblies to an overall resolution of 3.3 Å (Fig. 5 and Supplementary Fig. 7). The half-complex structure shares the same overall asymmetric architecture as the previously reported structure of the truncated *E. coli* restriction complex that lacks the N-terminal domain of McrB (EcMcrB Δ NC)⁴¹ (Supplementary Fig. 8). Despite being stabilized by different guanine nucleotide analogs (5'-guanylyl imidodiphosphate (GMPPNP) versus GTP γ S), the two models superimpose with an overall RMSD of 2.97 Å, even across the asymmetrically interacting McrC subunit (Supplementary Fig. 8a). The orientations of interacting subunits are also spatially conserved (Supplementary Fig. 8c–e), suggesting that the assembly and asymmetric architecture of the restriction complex are fundamentally maintained, regardless of the used nucleotide analog.

As with TgMcrBC, a single EcMcrC monomer inserts into the central pore of the EcMcrB hexamer (Fig. 5a, b). A cross-section slice through the map at the height of the bound nucleotides reveals that the same intrinsic asymmetry is present, with loose F/A and A/B interfaces, and tight B/C, C/D, D/E and E/F interfaces (Fig. 5c). The unique interactions stabilizing each tight interface are also conserved in EcMcrBC and absent in the loose interfaces: Arg337 (Arg414 in TgMcrB) and Asp343 (Asp420 in TgMcrB) interact in *trans* with Phe428 (Tyr530 in TgMcrB) and Arg283 (Arg360 in TgMcrB), respectively (Supplementary Fig. 9a, b).

EcMcrC shares the same general architecture as TgMcrC, featuring an extended finger domain and a C-terminal nuclease domain (Supplementary Fig. 9c). EcMcrC, however, lacks the N-terminal portion of the scaffold domain, retaining only a small β -hairpin insertion between the finger and nuclease domains (Fig. 3d and Supplementary Fig. 9c). Sequence alignment of McrC family proteins suggests that these insertion strands serve as a conserved linker between the finger and nuclease domains (Supplementary Fig. 5). The finger domains superimpose with an RMSD of 2.4 Å (sequence identity: 20%), confirming the

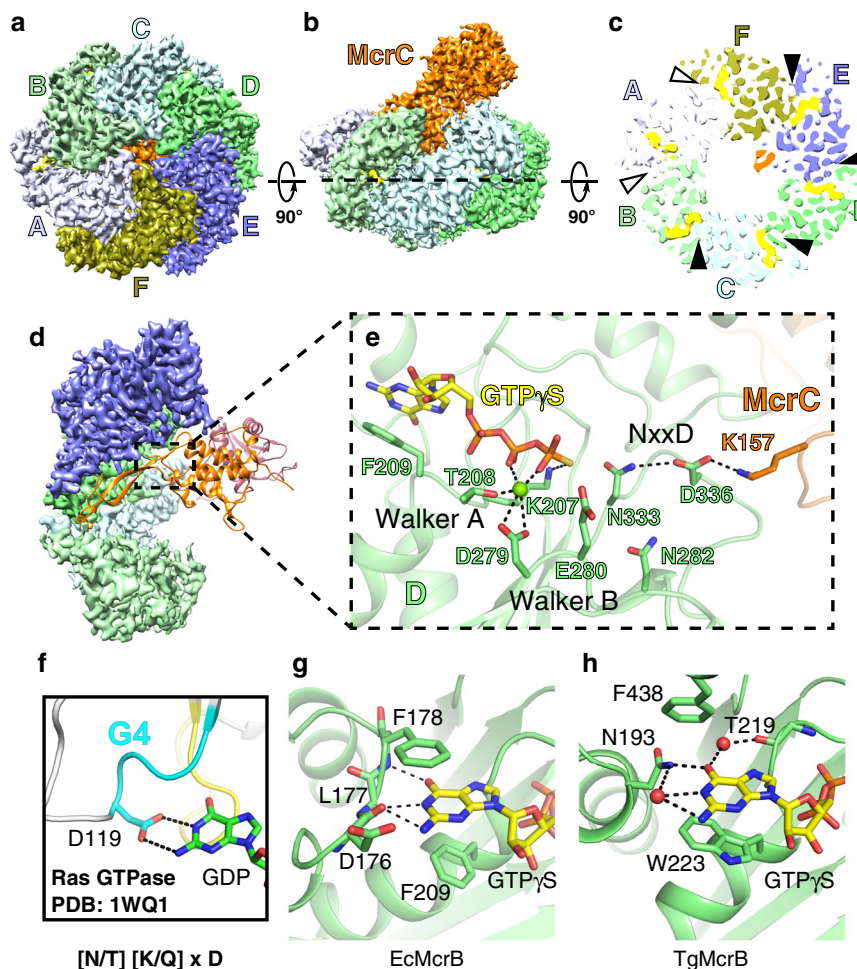


Fig. 5 The McrBC complexes of *E. coli* and *T. gammatolerans* show a conserved architecture. **a, b** Bottom and side views of the cryo-EM density map of the EcMcrBC half-complex. **c** Slice section through the map of the EcMcrBC half-complex at the level of the bound nucleotides, indicated by the dashed line in **b**. Solid and empty arrowheads indicate tight and loose interfaces, respectively. **d** Close-up view of the interaction of EcMcrC with the EcMcrB hexamer at the D/E interface. EcMcrB and EcMcrC are shown in surface and ribbon representation, respectively. For clarity, subunits A and F are not shown. **e** Hydrogen-bonding network formed by EcMcrC with residues of the NxxD motif at the D/E interface of the EcMcrB hexamer. Spheres indicate waters (red) and a magnesium ion (green). Dashed lines indicate hydrogen bonds and metal coordination. **f–h** Structural basis for guanine recognition in the Ras GTPase (PDB: 1WQ1)⁶⁸ and in the McrB homologs (EcMcrB and TgMcrB).

activity due to an alanine mutation at Asn282 would arise from a similar structural perturbation¹⁶. Interestingly, substituting a lysine for TgMcrC's catalytic Arg263 partially restores the stimulatory effect that is lost when this side chain is replaced with alanine (Fig. 4d). Together these data demonstrate that stimulated GTP hydrolysis in different McrBC homologs occurs via a conserved molecular mechanism.

Divergent McrB homologs employ the same generalized principles for nucleotide specificity. In every GTPase, the conserved sequence [N/T] [K/Q]xD (termed the “G4 element”) confers nucleotide specificity^{37,40,49}. The absolutely conserved aspartate side chain in this motif forms specific hydrogen bonds with the 1' amine and 2' amino group of the guanine base, thereby distinguishing it from ATP (Fig. 5f). Nothing in the TgMcrB AAA+ domain makes contact with this portion of the nucleotide (Fig. 2a, b), suggesting other structural features fulfill this role. Our reconstructions of the full-length EcMcrBC and TgMcrBC complexes reveal how each individually achieves this end (Fig. 5g, h). In EcMcrBC, a loop that lies directly upstream of the AAA+ domain coordinates the guanine base through main-chain interactions (Fig. 5g). The backbone carbonyl of Asp176 hydrogen

bonds with both the 1' amine and 2' amino group of the guanine base, while the main-chain nitrogen of Phe178 reads out the 6' carbonyl group. The same hydrogen bonds were observed in the truncated, GMPPNP-stabilized EcMcrBΔNC complex containing residues 162–465⁴¹. Collectively these interactions would discriminate against the substitution of an amino group at the 6' position (as in ATP and XTP) and the absence of an amino group at the 2' position (as in ATP and ITP), consistent with EcMcrB's nucleotide selectivity preferences of GTP > ITP > XTP >> ATP⁵¹. TgMcrB, in contrast, specifically coordinates the guanine base through two water-mediated interactions (Fig. 5h). Asn193 at the very beginning of the AAA+ domain directly hydrogen bonds to guanine's 6' carbonyl and orients a water molecule to interact with the 1' amine and 2' amino group. The backbone carbonyl of Thr219 also interacts with the 6' carbonyl group of the base via a second bridging water. Importantly, the fundamental chemistry underlying guanine nucleotide recognition is conserved between both homologs despite each utilizing different structural elements.

McrBC forms a tetradecameric assembly through the dimerization of McrC. Previous studies reported that EcMcrBC

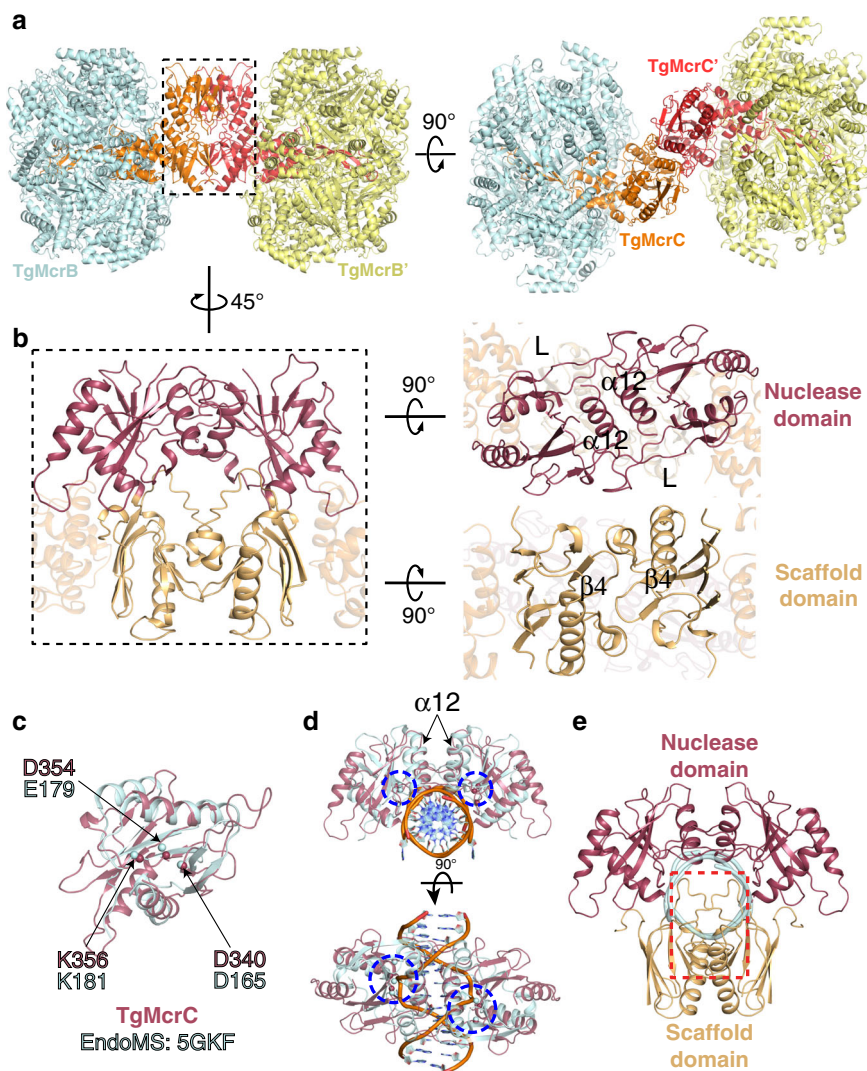


Fig. 6 The tetradecameric assembly of the TgMcrB^{AAA}C complex shows a cleavage-incompetent conformation. **a** Side views of the full TgMcrB^{AAA}C complex are shown in ribbon representation. Two TgMcrC (orange and red) form a dimer that bridges two TgMcrB hexamers (cyan and yellow). **b** Close-up views of the TgMcrC dimer interfaces formed by the two nuclease domains (upper right panel) and the two N-terminal domains (lower right panel). The $\alpha 12$ helix and a loop between the $\beta 10$ and $\beta 11$ strands are labeled as “ $\alpha 12$ ” and “L”, respectively. **c** Superposition of the monomeric structures of TgMcrC and EndoMS (sequence identity: 12%; PDB: 5GKF)⁵³. The conserved residues involved in the cleavage activity are labeled and shown as spheres. **d** Structural comparison between the TgMcrC dimer in the TgMcrB^{AAA}C complex and the EndoMS dimer in a DNA-bound state. The blue dashed circles indicate the active sites for DNA cleavage. **e** Illustration of the cleavage-incompetent conformation of TgMcrC. For clarity, the structure of the EndoMS protein is not shown. The backbone of the DNA substrate bound to EndoMS is colored in cyan. The red square indicates the regions of potential steric clashes.

complexes form tetradecameric assemblies *in vitro*^{20,21,41}. In our hands, dimeric McrBC complexes generated using the full-length Tg and Ec proteins exhibit a high degree of conformational variability, which hampered efforts to calculate complete, interpretable maps for these larger oligomeric states. To overcome this limitation, we produced complexes containing full-length TgMcrC bound to the AAA+ domain of TgMcrB (TgMcrB^{AAA}C) in the presence of GTP γ S. This assembly was structurally more homogeneous and allowed us to calculate maps of the “half”-complex at 3.7-Å resolution, as well as a C2-symmetrized map of the entire TgMcrB^{AAA}C tetradecameric complex at 4.2-Å resolution (Supplementary Fig. 10). A TgMcrC dimer bridges two TgMcrB^{AAA} hexamers in this structure (Fig. 6a), with the scaffold and nuclease domains forming the dimer interface (Fig. 6b). The nuclease domains associate through their $\alpha 12$ helices and a loop between the $\beta 10$ and $\beta 11$ strands (“L”), whereas the neighboring scaffold domains interact with each other through their $\beta 4$ strands that form main-chain

hydrogen bonds with each other. EcMcrBANC shows a similar overall arrangement, although numerous single-particle classes with different angles between the two half-complexes were reported for this assembly⁴¹. Interestingly, the half-complex reconstruction of the full-length EcMcrBC tetradecamer shows density for an additional ordered EcMcrC nuclease domain (Supplementary Fig. 11a). The organization of the EcMcrC nuclease domains at this dimer interface is identical to that seen in other McrBC complexes⁴¹ (Fig. 6a, b), with the $\alpha 10$ helix and an analogous extended loop serving as the primary points of contact (Supplementary Fig. 11b). This observation implies the same tetradecameric assembly is formed by the full-length construct.

The McrC dimer adopts a cleavage-incompetent conformation in the absence of a DNA substrate. The DNA-bound structures of other PD-(D/E)xK nucleases provide a template for modeling

McrC's cleavage activity. Of the many structural homologs identified by the DALI server⁵², the coordinates of the *Thermococcus kodakarensis* EndoMS endonuclease⁵³ (sequence identity 12%; PDB: 5GKF; Z-score 7.9) provided the best framework for these purposes. EndoMS binds DNA as a dimer, with each active site attacking a single strand of the DNA duplex to induce a double-strand break. As with other PD-(D/E)xK enzymes, Asp165^{EndoMS}, Glu179^{EndoMS}, and Lys181^{EndoMS} coordinate a divalent metal cofactor that is required for catalytic function^{53,54}. Structural superposition confirms TgMcrC's C terminus shares the same fold and identifies Asp340^{TgMcrC}, Asp354^{TgMcrC}, and Lys356^{TgMcrC} as putative catalytic side chains based on their spatial alignment with the EndoMS metal-binding residues (Fig. 6c). EcMcrC also shares this structural homology (Supplementary Fig. 11c). Importantly, our modeling is consistent with previous biochemical data showing that mutation of the predicted catalytic residues in EcMcrC (Asp224^{EcMcrC}, Asp257^{EcMcrC}, and Lys259^{EcMcrC}) impairs cleavage of modified DNA *in vitro*⁵⁵. Further comparison shows that the organization and location of the active sites in the TgMcrC and EcMcrC dimers is conserved between the two species (Supplementary Fig. 11d).

To gain insight into McrC cleavage, we overlaid two copies of the TgMcrC and EcMcrC endonuclease domains independently onto the dimeric, DNA-bound EndoMS complex (Fig. 6d and Supplementary Fig. 11e). The nuclease domains align in an orientation that resembles the dimer configuration captured in our cryo-EM structures; however, we observe numerous steric clashes in both models. TgMcrC's scaffold domain and the α 12 nuclease helices collide with the DNA substrate (Fig. 6e). EcMcrC lacks an N-terminal scaffold domain yet still clashes with the DNA backbone, owing to the first helix of its nuclease domain being significantly longer (Supplementary Fig. 11f). Attempts to model similar interactions with other structurally related homologs like EcoRV⁵⁶ (PDB: 1AZ0; sequence identity: 8% with TgMcrC and 13% with EndoMS) and the *Sulfolobus solfataricus* Holliday junction endonuclease⁵⁷ (PDB: 1OB8; sequence identity: 17% with TgMcrC and 9% with EndoMS) resulted in substantial clashes between the two McrB hexamers. We therefore speculate that our dimeric McrBC structures depict a conformation that is incompatible with DNA cleavage, and that a major conformational change would be required for nuclease activity to proceed unencumbered.

Discussion

Our structural analysis reveals that TgMcrB^{AAA} forms an asymmetric hexamer, similar to the architecture adopted by many other AAA+ family proteins^{58–65}. In the hexameric arrangement, four of the subunits, B, C, D, and E, occupy a radial sector (measured as the radius between the C α positions of Lys221 residues in neighboring subunits) of 59°, with the other two subunits, A and F, occupying radial sectors of 60° and 64°, respectively. This distortion of the hexameric assembly, which results in four tight and two loose interfaces (Fig. 1c–e), appears to be maintained by the conformation of key interface residues—Arg360, Glu527, and Tyr530 in one monomer and Arg414, Asp420, and Arg424 in its neighbor—acting in *trans*. Alanine substitutions of these residues increase basal GTPase activity by ~two-fold (Supplementary Fig. 1i). We speculate that interface mutations alter the programmed asymmetry, causing the unrestrained individual subunits to wobble randomly and leading to uncoordinated, stochastic GTP hydrolysis throughout the hexamer. The asymmetry in the ring also explains how crystal packing forces could induce and/or sustain the open conformation observed in our TgMcrB^{AAA} X-ray structure (Supplementary Fig. 1j–l), as the loose interfaces likely have a greater propensity

for flexibility resulting from the fewer stabilizing interactions. These observations argue that asymmetry is an intrinsic characteristic of the McrB^{AAA} hexamer rather than being induced upon McrC binding, as has recently been proposed⁴¹.

While all McrBC structures presented here display the same arrangement of four tight interfaces and two loose interfaces, the previous EcMcrB Δ NC structure showed three GMPPNP-bound interfaces and three GDP-bound interfaces in the McrB hexamer, which were likely to be tight and loose interfaces, respectively⁴¹. This discrepancy might be due to subtly different binding affinities for nucleotide analogs and the sensitivity of the EcMcrB assembly to nucleotide depletion. TgMcrB, in contrast, exists as stable hexamers even in the absence of any nucleotide (Supplementary Fig. 1a), suggesting that the balance between nucleotide affinity, occupancy, and structural integrity could affect the dynamics of McrB AAA rings.

The TgMcrB^{AAA}+ domain possesses all the catalytic machinery needed for nucleotide hydrolysis. We find that the canonical *cis*-acting Sensor II arginine is replaced with a *trans*-acting arginine (Arg425) that is positioned adjacent to the charge-compensating arginine finger (Arg426) in helix α 11 (Fig. 2b). Our cryo-EM and X-ray structures of TgMcrB^{AAA} reveal that Arg425 is not only important for stabilizing Glu375 in *cis* as predicted from the previous structures of *E. coli* complexes⁴¹, but also interacts with the phosphates of GTP in *trans* (Fig. 2b). Asn410 (consensus loop) and Glu357 (Walker B motif) together position the catalytic water. We also note that Trp223 forms a crucial π -stacking interaction with the guanine base that is present in the EcMcrBC reconstructions⁴¹ (Fig. 5e) and functionally conserved at the sequence level in other homologs. Perturbing any of these side chains reduces basal GTP hydrolysis of TgMcrB^{AAA}. Similar phenotypes were observed with the corresponding mutations in the *E. coli* protein¹⁶, indicating that the basic catalytic machinery is hardwired into the McrB AAA+ fold across evolution.

We demonstrate that McrC-stimulated GTP hydrolysis is a broadly conserved property of the McrBC family and not simply a unique feature of the *E. coli* homolog (Fig. 4d)^{16,35}. While this type of stimulation is uncommon among AAA+ proteins, it resembles the activation of small G proteins by their cognate GAPs. GAPs enhance catalytic turnover either by contributing essential catalytic residues in *trans* or by conformationally stabilizing and/or reorienting active site elements into an optimal configuration⁶⁶. In nearly every case, these interactions affect the charge-compensating element⁶⁷. RasGAP, for example, provides the arginine finger needed for Ras turnover, while RGS4 binding to G _{α} 1 reorients an existing arginine in the switch I motif^{68,69}. A notable exception is RapGAP, which provides in *trans* an asparagine that positions the nucleophilic water⁷⁰. Our structures show that TgMcrC and EcMcrC stimulate hydrolysis indirectly by altering the conformation of the McrB consensus loop. Both proteins insert a conserved basic residue (Arg263^{TgMcrC} and Lys157^{EcMcrC}) at the edge of the McrB active site and, via a hydrogen-bonding network, reposition a conserved asparagine (Asn410^{TgMcrC} and Asn333^{EcMcrC}) that in turn correctly orients the catalytic water for nucleophilic attack on the γ -phosphate (Figs. 4b and 5e, and Supplementary Fig. 9e, f). This conserved molecular mechanism thus represents a unique variation on a common theme. We note that the helical bundle of the McrC finger domain wedges into the E/F interface at the top of the McrB hexamer in both structures (Fig. 3f). This interaction not only anchors McrC, but also directs its catalytic machinery to the adjacent active site at the D/E interface (Fig. 4b and Supplementary Fig. 6c). These constraints dictate that McrC stimulation can only occur at a single active site at any given time.

In our structures, the consensus loop and the in *trans* arginine finger adopt different conformations in each active site around

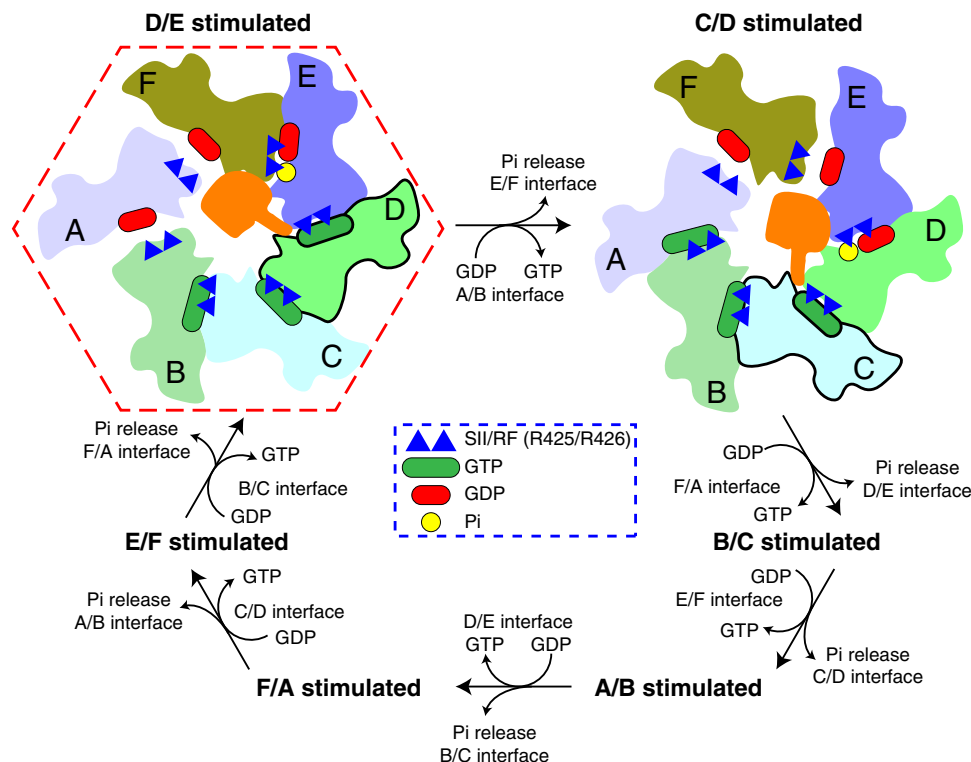


Fig. 7 Rotation model for the catalytic cycle of McrB. Schematic representation of the putative GTP hydrolysis cycle that proceeds sequentially in a clockwise manner around the hexameric McrB ring relative to McrC in the central pore. The “finger” extending from McrC represents the arginine/lysine residue that interacts with the NxxD motif. Ovals at the interfaces of the hexamer represent GTP (green) and GDP (red). The subunits indicated by the thick outlines are those in which the NxxD has been reorganized by McrC inserting its arginine/lysine residue. The red hexagon indicates the state observed in our cryo-EM structures.

the McrB hexamer (Fig. 4e–j). We interpret each configuration as representing a different state in the hydrolysis cycle (Fig. 7, dashed red outline). The McrC-engaged D/E active site assumes a transition state-like conformation with the catalytic machinery optimally positioned for stimulated turnover. In the tight C/D and B/C active sites, GTP is bound but the catalytic components are in a suboptimal conformation. This configuration suggests a pre-hydrolysis state that is primed for interaction with McrC. The loose A/B and F/A sites represent low-affinity, post-hydrolysis states that allow for free exchange of GTP and GDP, consistent with McrB not requiring a guanine nucleotide exchange factor. In support of this notion, we find GDP at the A/B site in our EcMcrBC structure (Supplementary Fig. 12a–f, s), the F/A site in the structure of the TgMcrBC complex (Supplementary Fig. 12g–l, t), and at both sites in the structure of the TgMcrB^{AAA}C complex (Supplementary Fig. 12m–r, u). The final tight E/F site likely adopts a post-hydrolysis state that is partially destabilized, but still remains intact due to the presence of the γ -phosphate in the bound GTP γ S. These data suggest that McrC-stimulated GTP hydrolysis proceeds via a coordinated mechanism that cycles around the McrB hexamer, engaging each composite active site sequentially (Fig. 7). In this scheme, the release of the γ -phosphate and the intrinsic asymmetry of the complex serve as the driving forces for a rotational movement. Release of the phosphate would destabilize the E/F interface, converting it from a tight to a loose configuration. This could promote a transition of the A/B interface from loose to tight, where exchange of GTP for GDP has presumably occurred. Weakening the E/F interface would destabilize the interactions with the helical bundle that anchor the finger domain (Fig. 3f), thereby releasing McrC and allowing it to rotate. The asymmetry of the structure would bias the movement in a clockwise direction, as

the helical bundle of the finger domain would not be able to associate with the loose F/A interface, and thus would have to intercalate into the D/E interface. This engagement would orient McrC to insert its catalytic arginine/lysine into the C/D active site, where it could trigger the next hydrolysis event to power the motor (Fig. 7). The extensive contacts formed between the helix-loop-helix tip of the finger domain and all six subunits of the McrB hexamer (Fig. 3e) would ensure that McrC does not dissociate from the complex following stimulated turnover. The stepwise transition from one binding interface to the next (Fig. 7) is reminiscent of F/V-type ATPases^{71–74}.

A similar sequential mechanism for GTP hydrolysis and clockwise movement of McrC were previously proposed based on the EcMcrBANC structure⁴¹. In the half-complex of the EcMcrBANC tetradecameric structure, three GMPPNP and three GDP were assigned in the subunit interfaces of McrBAN (Supplementary Fig. 12v). One of the GMPPNP-bound interfaces (the “CD interface”, which corresponds to the E/F interface in this study) was assumed to be the McrC-stimulated active site. We interpret this interface as a post-hydrolysis site, and instead believe the stimulation and formation of the catalytic transition state occurs at the adjacent D/E interface. It was further speculated that the interaction of the “CD interface” with the β -sheet “stalk” of McrC-initiated GTP hydrolysis⁴¹. The resulting conformational changes in the McrB signature motif were not fully appreciated, however, due to the limited resolution of the EcMcrBANC structure. Our cryo-EM reconstruction of TgMcrBC unambiguously reveals the catalytic water molecules and illustrates how McrC’s insertion of a basic residue specifically repositions the signature motif to trigger hydrolysis, providing a chemically and energetically favorable description of stimulated turnover. Given the conserved structural features and asymmetry

present in both the Tg and Ec complexes, we anticipate that other McrBC homologs will follow this mechanochemical model.

Efficient hydrolysis also depends on the ability of an enzyme to bind and differentiate its appropriate nucleotide substrate. GTPases use the conserved aspartate in the G4 element to coordinate substituents at the 1' and 2' positions of the guanine base, while AAA+ proteins recognize the amino group at the 6' position in adenine^{37,68}. By reading out the 1', 2', and 6' positions of the guanine base, McrB homologs appear to have combined both strategies to fine-tune their specificity for GTP in the context of a AAA+ fold. Ec and TgMcrB both use the same basic chemistry for this recognition, but each employs different structural components to mediate these contacts (Fig. 5g, h). Interestingly, these pieces lie outside the core AAA+ fold and localize to either the flexible linker that connects to EcMcrB's N-terminal DNA-binding domain or the very start of helix α 1 in TgMcrB (Supplementary Figs. 2 and 3a, colored in gold). Although the motor and cleavage machineries are conserved among McrB homologs (Figs. 3–5), the N-terminal domains and connecting linkers are highly divergent. Crystallographic studies have shown that EcMcrB uses the DUF3578 fold to bind methyl-cytosine modifications^{17,19}, whereas the N-terminal domain of TgMcrB consists of a YTH fold that specifically targets 6mA-modified DNA²⁹. The related LlaJI restriction system from *Helicobacter pylori* binds DNA site-specifically via an N-terminal B3 domain²⁸. The subtle distinctions we observe with regard to nucleotide recognition are therefore significant, and provide a blueprint for how divergent homologs can maintain the necessary pattern of hydrogen bonding even in radically different structural contexts. Future structural characterization will determine if these principles hold true for other McrBC family members.

Previous biochemical studies suggest that McrBC's stimulated GTP hydrolysis powers DNA translocation^{18,22}. While we do not directly address how this may occur in this study, our structures impose constraints with regard to the potential pathway of DNA and the organization of a cleavage-competent McrBC complex. DNA and RNA typically pass through the central pore of hexameric AAA+ helicases and translocases driven by ATP hydrolysis^{59,75}. Based on recent cryo-EM reconstructions, a similar mechanism has been proposed for EcMcrBC, in which the McrB N-terminal domains might interact with DNA on the bottom of the hexamer and thread it into the central channel⁴¹. Although we see weak density in our full-length EcMcrBC map that corresponds to the N-terminal domains near the top of the complex (Supplementary Fig. 11g), numerous structural observations oppose this potential trajectory. First, McrC specifically binds in the center of the McrB hexamer, blocking access to this pathway in both the Tg and Ec complexes. The asymmetric association of the finger domain's helical bundle with the D/E/F subunits shrinks the pore diameter at the top of the hexamer to ~10 Å (Supplementary Fig. 6b), while the loop-helix-loop region completely occludes the pore at the bottom of the hexamer (Fig. 3a, and Supplementary Figs. 6e and 13a, b), which narrows to a diameter of ~8 Å even without McrC. Passage through the ring in this state would require both distortion and/or melting of the DNA duplex to conform to the narrow dimensions of the structure, as well as either a complete displacement or gross conformational reorganization of McrC. Such changes would uncouple the sequential, coordinated stimulation of GTP hydrolysis suggested by our structures and yield a translocation mechanism that would use a completely stochastic catalytic process, and would depend on alternating cycles of binding and dissociation for both McrC and DNA. While we cannot rule out that additional conformational changes occur upon DNA binding, biochemical characterization of EcMcrBC has shown that DNA binding and GTP hydrolysis are separate and distinct

properties *in vitro*^{14,22,35,36}. It therefore seems unlikely that DNA binding would significantly alter the architectural and catalytic interactions that have been conserved across kingdoms. Second, we resolve clear density decorating the outside edges of the TgMcrB^{AAA} hexamer that we attribute to the TgMcrB N-terminal domains (Supplementary Fig. 11h). The localization of these domains nearly perpendicular to the pore axis would require DNA, if it were to pass through the center of the TgMcrBC complex, to bend dramatically, more than has been observed in any structure to date. Energetically, such a configuration would be extremely unfavorable⁷⁶. The short seven amino acid linker connecting the N-terminal domains to the Tg AAA+ domains combined with the structural requirements of nucleotide selectivity would likely prohibit a large-scale rearrangement of these domains within the restriction complex. Taken together, these findings argue against a mechanism in which DNA passes through the central channel in the McrB hexamer. We speculate that McrBC complexes use an alternative, yet to be elucidated means to translocate DNA.

Both EcMcrBC and TgMcrBC form tetradecameric complexes that are bridged by an McrC dimer (Fig. 6a and Supplementary Fig. 11a). Our structural modeling, however, suggests that the conformation of this McrC dimer is incompatible with DNA binding and cleavage. Superposition with EndoMS shows that the N-terminal scaffold domain of TgMcrC and the first helix in the nuclease domain of EcMcrC clash with the modeled DNA substrate (Fig. 6d, e and Supplementary Fig. 11e, f). Modeling these interactions in the context of the tetradecameric complexes indicates further steric hindrance: superimposed DNA strands would clash with the McrB subunits (Supplementary Fig. 13c, d; left panels) and have a trajectory that is directed away from the central pore in the hexamer, offset by nearly 30° and 10° for the TgMcrB^{AAA}C and EcmMcrB^{ANC} structures, respectively (Supplementary Fig. 13c, d; right panels). These observations support the idea that the current structures represent binding/cleavage-incompatible conformations. It remains to be seen whether DNA binding alone could induce a cleavage-competent conformation. Interestingly, GTP γ S does not support EcMcrBC DNA cleavage *in vitro*²², consistent with our structural findings here. Moreover, mutation of Pro203 to valine in EcMcrB, a residue in a loop close to the γ -phosphate and the hexamer interface, significantly reduces both EcMcrC-stimulated GTP hydrolysis and DNA cleavage of an “ideal” substrate with R^MC sites optimally spaced 63 base pairs apart so as not to require translocation²². This finding raises the possibility that GTP hydrolysis is also needed for the transient reorganization of the McrC monomers, and that blocking this activity would lead to a nonproductive arrangement. Further experiments will be needed to fully understand how the McrBC complex cleaves DNA.

Modification-dependent restriction systems function as a conserved barrier to lytic phage infections. In the ongoing arms race between virus and host, phages have evolved inhibitors against McrBC and GmrSD^{77,78}, which confer the ability to bypass these defense machineries and allow phages to survive under conditions, in which they would normally be restricted. Knowing how these defense systems work, and how they have been naturally subverted is clinically important and will aid in the long-term development of small-molecule inhibitors that can impair conserved defense systems and improve the efficacy of phage-based treatments.

Methods

Cloning, expression, and purification of TgMcrB^{AAA}. The gene for the *T. gammatolerans* EJ3 McrB protein (JGI IMG/M ID 644807740) was codon optimized for expression in *E. coli* and synthesized commercially by GENEART (Supplementary Table 1). The DNA for the AAA+ domain of TgMcrB (residues

186–613) was amplified by PCR and cloned via Gibson assembly (New England Biolabs) into the pET15bP vector, a modified version of the pET15b vector, in which the Factor Xa cleavage site after the N-terminal 6xHis tag was replaced with an HRV 3C cleavage site. Cleavage by HRV 3C protease leaves a glycine and a proline residue immediately upstream of TgMcrB^{AAA}'s N-terminal methionine. Primers used in this study are summarized in Supplementary Table 2.

Selenomethionine-labeled (SeMet) TgMcrB^{AAA} was expressed in minimal medium using methionine auxotrophs (T7 Express Crystal Competent *E. coli*, New England Biolabs) according to manufacturer's protocols. For the expression of native TgMcrB^{AAA}, the construct was transformed into *E. coli* BL21(DE3) cells, which were grown at 37 °C in Terrific Broth. When OD600 reached 1.0, protein expression was induced by addition of 0.3 mM isopropyl β-D-thiogalactoside (IPTG) and cells were grown overnight at 19 °C. Cells were harvested by centrifugation at 6000 × g for 15 min at 25 °C, and washed twice with nickel-loading buffer (NLB; 20 mM HEPES, pH 7.5, 500 mM NaCl, 30 mM imidazole, 5% glycerol (v/v), and 5 mM β-mercaptoethanol). Pellets were typically flash frozen in liquid nitrogen and stored at –80 °C.

Thawed pellets from 500-mL cultures were resuspended in 30 mL of NLB supplemented with 10 mM PMSF, 5 μg/mL DNase I (Roche), 5 mM MgCl₂, and a tablet of complete protease inhibitor cocktail (Roche). Lysozyme was added to a final concentration of 1 mg/mL and the mixture was incubated for 15 min at 4 °C with rocking. Cells were disrupted by sonication and the lysate was cleared of debris by centrifugation at 19,700 × g for 30 min at 4 °C. The supernatant was filtered using a 0.45-μm cutoff syringe filter, incubated at 65 °C for 20 min, centrifuged at 6000 × g for 15 min at 4 °C, and loaded onto a 5-mL HiTrap chelating column (GE Healthcare) charged with NiSO₄ and then washed with NLB. TgMcrB^{AAA} was eluted with an imidazole gradient from 30 mM to 1 M. Peak fractions were pooled, HRV 3C protease was added, and the sample was dialyzed overnight at 4 °C against cleaning buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 5% glycerol (v/v), and 5 mM β-mercaptoethanol (10 mM for SeMet-labeled protein)). Another 5-mL HiTrap chelating column charged with NiSO₄ was equilibrated with cleaning buffer and the sample was applied to this column, followed by elution with a NLB to remove the cleaved 6xHis tag. Pooled peak fractions were concentrated to 2 mL with a centrifugal concentrator (50 kDa cutoff, Millipore). The concentrated protein was further purified by size-exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare). During SEC, all proteins were exchanged into SEC₁₅₀ buffer (20 mM HEPES, pH 7.5, 150 mM KCl, 5 mM MgCl₂, and 1 mM DTT (5 mM for SeMet TgMcrB^{AAA})). For crystallographic analysis, the protein was concentrated to 40–80 mg/mL. Concentrations of purified proteins were determined by SDS-PAGE with BSA standards. All point mutations were introduced into TgMcrB^{AAA} in the pET15bP vector by quick-change PCR and the proteins were purified as described above.

Cloning, expression, and purification of TgMcrB. The gene for full-length TgMcrB (residues 1–613) was amplified by PCR and cloned into the pET15bP vector via Gibson assembly. Cleavage by HRV 3C protease leaves a glycine and a proline residue immediately upstream of TgMcrB's N-terminal methionine. The construct was transformed into *E. coli* BL21(DE3) cells, which were grown at 37 °C in Terrific Broth. When OD600 reached 1.5, protein expression was induced with 0.3 mM IPTG and cells were grown overnight at 19 °C. Cells were harvested and washed twice with NLB. Pellets were typically flash frozen in liquid nitrogen and stored at –80 °C. Thawed pellets from 2-L cultures were resuspended in 30 mL of NLB supplemented with 10 mM PMSF, 5 μg/mL DNase I, 5 mM MgCl₂, and a tablet of complete protease inhibitor cocktail. Cells were lysed and the full-length TgMcrB protein was purified as described above with the slight modification of using 250 mM KCl in the SEC buffer (20 mM HEPES, pH 7.5, 250 mM KCl, 5 mM MgCl₂, and 1 mM DTT). The protein was concentrated to 20–40 mg/mL.

Cloning, expression, and purification of TgMcrC. The gene for the *T. gamma-tolerans* EJ3 McrC protein (JGI IMG/M ID 644807739) was codon optimized for expression in *E. coli* and synthesized commercially by Integrated DNA Technologies (Supplementary Table 1). The DNA encoding full-length TgMcrC (residues 1–458) was amplified by PCR and cloned via Gibson assembly into the pCAV6 vector, a modified version of the pMAL c5x T7 expression vector, in which a 6xHis tag was introduced upstream of the N-terminal MBP sequence and an HRV 3C cleavage site replaces that for Factor Xa in the multiple cloning site. Cleavage by HRV 3C protease leaves a glycine and a proline residue immediately upstream of TgMcrC's N-terminal methionine.

The TgMcrC construct was transformed into *E. coli* BL21(DE3) cells, which were grown at 37 °C in Terrific Broth. When OD600 reached 1.0, protein expression was induced with 0.3 mM IPTG and the cells were grown overnight at 19 °C. Cells were harvested and washed twice with NLB. Pellets were typically flash frozen in liquid nitrogen and stored at –80 °C. Thawed pellets from 500-mL cultures were resuspended in 30 mL of NLB supplemented with 10 mM PMSF, 5 μg/mL DNase I, 5 mM MgCl₂, and a tablet of complete protease inhibitor cocktail. Lysozyme was added to a final concentration of 1 mg/mL and the mixture was incubated for 15 min at 4 °C with rocking. Cells were disrupted by sonication and the lysate was cleared of debris by centrifugation at 19,700 × g for 30 min at 4 °C. The supernatant was filtered using a 0.45-μm cutoff syringe filter, loaded onto a 5-mL HiTrap chelating column charged with NiSO₄ and then washed with NLB.

TgMcrC was eluted with an imidazole gradient from 30 mM to 1 M. Peak fractions were pooled, HRV 3C protease was added, and the sample was dialyzed overnight at 4 °C against SP-loading buffer (SPLB; 20 mM HEPES, pH 7.5, 250 mM NaCl, 1 mM EDTA, 5% glycerol (v/v), and 1 mM DTT). The sample was applied to a 5-mL HiTrap SP HP column (GE Healthcare) equilibrated with SPLB and then washed with SPLB. TgMcrC was eluted with a NaCl gradient from 250 mM to 1 M. Because TgMcrC is prone to precipitate, no further purification steps were attempted and the pooled peak fractions yielded protein at a purity of ~70% and a concentration of ~0.8 mg/mL. All point mutations were introduced into TgMcrC in the pCAV6 vector by quick-change PCR and the proteins were purified as described above.

Cloning, expression, and purification of EcMcrB. The gene for full-length *E. coli* McrB (Uniprot P15005; JGI IMG ID 646316336) was codon optimized for expression in *E. coli* and synthesized commercially by GENEART (Supplementary Table 1). The DNA encoding full-length EcMcrB (residues 1–459) was cloned into the pMAL-c2XP vector, a modified version of the pMAL-c2X vector (New England Biolabs), in which the Factor Xa cleavage site after the N-terminal MBP tag was replaced with an HRV 3C cleavage site. Cleavage by HRV 3C protease leaves a glycine and a proline residue immediately upstream of EcMcrB's N-terminal methionine.

EcMcrB was transformed into *E. coli* BL21(DE3) cells, which were grown at 37 °C in Terrific Broth. When OD600 reached 1.0, protein expression was induced with 0.3 mM IPTG, and cells were grown overnight at 19 °C. Cells were harvested and washed once with TGED₅₀₀ buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol (v/v), 1 mM EDTA, 1 mM DTT, and 500 mM NaCl). Pellets were flash frozen in liquid nitrogen and stored at –80 °C. Thawed pellets from 500-mL cultures were resuspended in 30 mL of TGED₅₀₀ buffer supplemented with 10 mM PMSF, 5 μg/mL DNase I, 5 mM MgCl₂, and a tablet of complete protease inhibitor cocktail. Lysozyme was added to a final concentration of 1 mg/mL and the mixture was incubated for 15 min at 4 °C with rocking. Cells were disrupted by sonication and the lysate was cleared of debris by centrifugation at 19,700 × g for 30 min at 4 °C. The supernatant was filtered using a 0.45-μm cutoff syringe filter, loaded onto a 5-mL HiTrap MBP column (GE Healthcare), washed with TGED₅₀₀, and eluted with 10 mM D-maltose in TGED₅₀₀ buffer. Peak fractions were pooled, HRV 3C protease was added, and the sample was dialyzed overnight at 4 °C against TGED₅₀ buffer (TGED₅₀₀ buffer but with 50 mM NaCl instead of 500 mM). The sample was then applied to a 5-mL HiTrap Q HP ion-exchange column (GE Healthcare) in TGED₅₀ and eluted with a NaCl gradient from 50 to 500 mM. Peak fractions were pooled, concentrated, and further purified by SEC using a HiLoad 16/600 Superdex 200 pg column, during which the protein was exchanged into SEC₁₅₀ buffer. The protein was then concentrated to ~25 mg/mL.

Cloning, expression, and purification of EcMcrC. The gene encoding full-length *E. coli* McrC protein (Uniprot P15006; JGI IMG ID 637004274) was codon optimized for expression in *E. coli* and synthesized commercially by GENEART (Supplementary Table 1). The DNA encoding full-length EcMcrC (residues 1–348) was cloned into the pMAL-c2XP vector. Cleavage by HRV 3C protease leaves a glycine and a proline residue immediately upstream of EcMcrC's N-terminal methionine.

EcMcrC was transformed into *E. coli* BL21(DE3) cells and grown at 37 °C in Terrific Broth. When OD600 reached 1.0, protein expression was induced with 0.3 mM IPTG and cells were grown overnight at 19 °C. Cells were harvested and washed once with TGED₅₀₀ buffer. Pellets were flash frozen in liquid nitrogen and stored at –80 °C. Thawed pellets from 500-mL cultures were resuspended in 30 mL of TGED₅₀₀ buffer supplemented with 10 mM PMSF, 5 μg/mL DNase I, 5 mM MgCl₂, and a tablet of complete protease inhibitor cocktail. Lysozyme was added to a final concentration of 1 mg/mL and the mixture was incubated for 15 min at 4 °C with rocking. Cells were disrupted by sonication and the lysate was cleared of debris by centrifugation at 19,700 × g for 30 min at 4 °C. The supernatant was filtered using a 0.45-μm cutoff syringe filter, loaded onto a 5-mL HiTrap MBP column, washed with TGED₅₀₀ buffer, and eluted with 10 mM D-maltose in TGED₅₀₀ buffer. Peak fractions were pooled, HRV 3C protease was added, and the sample was dialyzed overnight at 4 °C against HGED₂₅₀ buffer (20 mM HEPES, pH 7.5, 5% glycerol (v/v), 1 mM EDTA, 1 mM DTT, and 250 mM NaCl). The sample was then applied to a 5-mL HiTrap SP HP ion-exchange column in TGED₂₅₀ buffer and eluted with a NaCl gradient from 250 mM to 1 M. Because EcMcrC is prone to precipitate, no further purification steps were attempted. The pooled peak fractions yielded protein at a purity of ~70% and a concentration of ~6 mg/mL.

GTPase activity assays. GTPase activity was measured by using a colorimetric malachite green assay that monitors the amount of free phosphate released over time⁷⁹. To measure the basal GTPase activity of TgMcrB^{AAA}, 0.4 μM TgMcrB^{AAA} was incubated with 1 mM GTP at 65 °C in reaction buffer (20 mM Tris-HCl, pH 8.0, 150 mM KCl, and 5 mM MgCl₂). To measure the GTPase activity of TgMcrB^{AAA} stimulated by TgMcrC, the same conditions were used but 0.1 μM TgMcrC was added. At time points of 0, 5, 10, 20, 30, 45, 60, 80, 100, and 120 min, 20-μL aliquots were taken and quenched with 5 μL of 0.5 M EDTA, pH 8.0. For colorimetric reactions, 150 μL of filtered malachite green solution were added to

each sample and incubated for 5 min. The absorbance at 650 nm of the samples was measured with a Multiskan GO Microplate Spectrophotometer (Thermo Scientific). The amount of phosphate released was determined using a standard curve. To account for the spontaneous hydrolysis of GTP at 65 °C, a protein-free sample containing GTP and magnesium was incubated in parallel, and the measured amount of phosphate released at each time point was subtracted from the corresponding measurements of protein-containing samples. The specific activity is reported for all wild-type and mutant proteins. Quantified data represent the average of three independent experiments using multiple independently purified batches of protein with error bars indicating the standard deviation from the mean ($n = 3$, mean \pm standard deviation). GraphPad Prism and Microsoft Excel were used for statistical analysis and to plot the data.

Negative-stain EM. Negatively stained samples were prepared as described⁸⁰. Freshly purified proteins were diluted to ~0.05 mg/mL with SEC₁₅₀ buffer supplemented with 2.5 mM GTP γ S before applying 5- μ L aliquots to glow-discharged grids. Grids were stained with 0.7% uranyl formate (Pfaltz & Bauer, U01000) and imaged with a Philips CM10 electron microscope equipped with a tungsten filament and operating at 100 kV. All images were recorded on an AMT XR16L-ActiveVu charge-coupled device camera (Woburn, MA, USA) using a defocus of approximately -1.5 μ m and a nominal magnification of 52,000 \times .

Crystallization, X-ray data collection, and structure determination of TgMcrB^{AAA}. SeMet TgMcrB^{AAA} was diluted to 16 mg/mL with SEC₁₅₀ buffer containing 2.5 mM GTP γ S and crystallized by sitting drop vapor diffusion in 0.1 M sodium acetate, pH 6.5, 17.5% 2-methyl-2,4-pentanediol (v/v) with a drop size of 2 μ L, and a reservoir volume of 650 μ L. Crystals appeared within 3–4 days at 20 °C and were cryo-protected with Parabar 10312 (Hampton Research) and frozen in liquid nitrogen. Single-wavelength anomalous diffraction (SAD) data were collected remotely on the tunable NE-CAT 24-ID-C beamline at the Advanced Photon Source at the selenium edge energy of 12.663 keV (0.9791 Å; Supplementary Table 3). Data were integrated and scaled using the NE-CAT RAPD pipeline, which utilizes LABELIT⁸¹, RADDOSE⁸², BEST⁸³, MOSFLM⁸⁴, Xtriage from PHENIX⁸⁵, XDS⁸⁶, and AIMLESS⁸⁷. Strong anomalous signal was obtained from a single crystal diffracting to 2.9 Å (space group $P2_1$; unit cell dimensions: $a = 100.24$ Å, $b = 108.87$ Å, $c = 118.67$ Å and $\alpha = 90.00^\circ$, $\beta = 107.41^\circ$, $\gamma = 90.00^\circ$), and multiple SeMet SAD datasets were collected from different positions of this crystal. All possible combinations of datasets were tested and merged using the program BLEND in the CCP4 suite^{88,89}. Experimental phases were obtained from the combination with the strongest anomalous signal. Heavy-atom sites were located using SHELX C/D/E⁹⁰ in the CCP4 suite and phasing, density modification, and initial model building were carried out using the CRANK2 pipeline⁹¹ in the CCP4 suite. Iterative rounds of refinement and model building were carried out using the programs COOT⁹² and REFMAC⁹³ in the CCP4 suite to improve the initial model, which resolved most regions of four TgMcrB^{AAA} monomers. Native TgMcrB^{AAA} with 2.5 mM GTP γ S was crystallized at a concentration of 16 mg/mL by sitting drop vapor diffusion in 0.1 M sodium acetate, pH 6.5, 16.75% (v/v) 2-methyl-2,4-pentanediol with a drop size of 2 μ L, and a reservoir volume of 650 μ L. Crystals were frozen as described above. The partial model from the SeMet SAD datasets was used as the search model to perform molecular replacement, using PHASER⁹⁴ in the CCP4 suite on a native dataset diffracting to 2.83 Å (space group $P2_1$; unit cell dimensions: $a = 100.02$ Å, $b = 108.55$ Å, $c = 118.43$ Å and $\alpha = 90.00^\circ$, $\beta = 106.94^\circ$, $\gamma = 90.00^\circ$), which was collected at the NE-CAT 24-ID-E beamline at the selenium edge energy of 12.663 keV (0.9791 Å). Further model building and refinement was carried out manually in COOT and PHENIX, respectively^{85,92}. Non-crystallographic symmetry was enforced during the refinement with no additional constraints imposed. The final model contained four well-resolved and two poorly resolved molecules in the asymmetric unit and was refined to 2.95 Å resolution with R_{work}/R_{free} values of 0.345/0.364 (Supplementary Table 3). Poorly resolved electron density surrounding the small subdomain of subunit F—near where each hexamer contacts its neighbor in the crystal lattice—and our inability to properly build and refine this portion of the model both contributed to the high R values.

Cryo-EM sample preparation and data collection. For TgMcrB^{AAA}, thawed protein was diluted to 10 mg/mL with SEC₁₅₀ buffer containing 2.5 mM GTP γ S. Samples were mixed with 20 \times digitonin (Calbiochem) stock to a final concentration of 0.05%, and 3.5 μ L aliquots were applied to C-flat thick holey carbon grids (CF-1.2/1.3-4C-T, Protochips), blotted for 7 s at 4 °C and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific).

For the TgMcrB^{AAAC} complex, thawed TgMcrB^{AAA} was mixed with freshly purified TgMcrC at a molar ratio of 4:1. The sample was concentrated using a 2-mL centrifugal concentrator (100 kDa cutoff, Millipore). Concentrated protein was buffer-exchanged to SEC₁₅₀ buffer in the concentrator, and the final concentration was estimated to be ~14 mg/mL by measuring the absorbance at 280 nm. The complex was then mixed with 50 \times GTP γ S stock solution to a final GTP γ S concentration of 2.5 mM and incubated for 30 min at 4 °C. Samples were mixed with 20 \times digitonin to a final concentration of 0.05%, and 3.5- μ L aliquots were

applied to C-flat thick holey carbon grids (CF-1.2/1.3-4C-T), blotted for 8–10 s at 4 °C and plunge-frozen in liquid ethane, using a Vitrobot Mark IV.

For the TgMcrBC and EcMcrBC complexes, thawed McrB was mixed with freshly purified McrC at a molar ratio of 4:1. The samples were concentrated using 2-mL centrifugal concentrators (100 kDa cutoff, Millipore). Concentrated proteins were then buffer-exchanged into SEC₂₅₀ buffer (for TgMcrBC) or SEC₁₅₀ buffer (for EcMcrBC) in the concentrators and finally concentrated to ~16 mg/mL. The prepared complexes were mixed with 50 \times GTP γ S stock solution to a final GTP γ S concentration of 2.5 mM and incubated for 30 min at 4 °C. Samples were mixed with 20 \times digitonin to a final concentration of 0.05% digitonin, and 3.5 μ L aliquots were applied to Quantifoil R1.2/1.3 400 mesh Au grids, blotted for 8–10 s at 4 °C, and plunge-frozen in liquid ethane, using a Vitrobot Mark IV.

Cryo-EM data were collected on a 300-kV Titan Krios electron microscope (Thermo Fisher Scientific) equipped with a K2 Summit direct electron detector at a nominal magnification of 29,000 \times in super-resolution counting mode using SerialEM^{95,96}. After binning over 2×2 pixels, the calibrated pixel size was 1.0 Å on the specimen level. For all specimens other than EcMcrBC, exposures of 10 s were dose-fractionated into 40 frames with a dose rate of 8 electrons per pixel per second, resulting in a total dose of 80 electrons per Å². For EcMcrBC, exposures of 20 s were dose-fractionated into 40 frames with a dose rate of 4 electrons per pixel per second, resulting in a total dose of 80 electrons per Å². Cryo-EM data collection statistics are summarized in Supplementary Table 4.

Cryo-EM data processing. For TgMcrB^{AAA} and the TgMcrB^{AAAC} complex, image processing was done in RELION-3.0-beta^{97–99}, and images for TgMcrBC and EcMcrBC were processed in both CryoSPARC-2.4.0 (Structura Biotechnology)¹⁰⁰ and RELION-3.0. All movie frames were corrected with a gain reference collected during the same EM session, and specimen movement was corrected using RELION's implementation of motion correction (for TgMcrB^{AAA} and TgMcrB^{AAAC}) or MotionCorr2 (for TgMcrBC and EcMcrBC) with dose weighting^{97,101}. The contrast transfer function (CTF) parameters were estimated using CTFFIND-4.1.8 (ref. ¹⁰²) for TgMcrB^{AAA} and TgMcrB^{AAAC} or Gctf-1.0.6 (ref. ¹⁰³) for TgMcrBC and EcMcrBC. Images showing substantial ice contamination, abnormal background, thick ice, low contrast, or poor Thon rings were discarded.

For TgMcrB^{AAA}, 1599 micrographs were collected, of which 1517 micrographs were selected for further processing. Particles were picked with Gautomatch (<https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/>) without templates, which identified 277,503 particles that were windowed into 320 \times 320 pixel images. The particle images were binned four times and subjected to two rounds of 2D classification. Classes that produced averages with fine structural detail and showed no overlap with neighboring particles were combined and used to generate an initial reference map. The selected 153,891 particles were subjected to 3D classification into four classes, three of which were selected and used to re-extract the corresponding particles into 320 \times 320 pixel images that were then rescaled into 256 \times 256 pixel images. The centered, re-extracted particles were refined with C1 symmetry to a resolution of 3.4 Å according to the Fourier shell correlation (FSC) = 0.143 criterion¹⁰⁴, which was used for all resolution estimates. Subsequent CTF refinement and Bayesian polishing improved the overall resolution of the map to 3.1 Å.

For the TgMcrB^{AAAC} complex, 1795 of the 2070 collected micrographs were selected for further processing. Gautomatch was used to pick the first 200 micrographs without templates, and then ~10,000 picked particles were subjected to 2D classification. Four representative class averages were then selected as templates for Gautomatch to pick particles from all the micrographs. The 264,850 auto-picked particles were cleaned-up by two rounds of 2D classification. The particles from eight classes with well-defined averages (156,149 particles) were used to generate an initial density map in RELION, which was then used as reference for 3D classification of the cleaned-up particles into six classes. Four classes showed good fine structure and were combined (115,774 particles), and subsequent 3D refinement with C1 symmetry, CTF refinement, and Bayesian polishing yielded a map at 4.4-Å resolution. A second dataset collected using the same conditions was processed following the same strategy, yielding a map at 4.3-Å resolution from 88,819 refined particles. The particles from the two datasets were combined and further refined with C1 symmetry to generate an improved map at 4.2 Å (204,593 particles). While this map showed strong density for one half of the complex, the other half was represented by substantially weaker density. 3D refinement was thus repeated with C2 symmetry imposed, which yielded a symmetrized map for the full complex map at 4.2-Å resolution. To overcome the flexibility of the connection between the two half-complexes, particles in the nonsymmetrized map were subjected to automated multibody refinement implemented in RELION-3, using individual masks for the two half-complexes that overlapped in the region of the twofold axis. Signal subtraction was performed for each rigid body using the "relion_flex_analyse" command, which only retains the signal inside the selected rigid body⁹⁹. The signal-subtracted particles for one of the two bodies were used to calculate a reference map for the half-complex, using the "relion_reconstruct" command. The signal-subtracted particles for both bodies were combined (409,186 particles) and subjected to 3D refinement with C1 symmetry and starting with a global search. Subsequent CTF refinement and 3D refinement yielded the final map for the half-complex at an overall resolution of 3.7 Å.

For the TgMcrBC complex, 1936 of 2078 micrographs and for the EcMcrBC complex, 1088 of 1161 micrographs were selected for further processing. Particles were picked with Gautomatch with templates generated from preliminary data collected on a 200-kV Talos Arctica electron microscope (Thermo Fisher Scientific). The auto-picked particles (354,707 for the TgMcrBC complex and 184,487 for the EcMcrBC complex) were extracted into 320 × 320 pixel images that were then rescaled into 256 × 256 pixel images. All particle images were used for ab initio reconstruction in Cryosparc-2.4.0, specifying three output classes. The best of the three maps, including 226,813 particles for the TgMcrBC complex and 106,684 particles for the EcMcrBC complex, were selected for nonuniform refinement, which yielded maps for the half-complexes at 3.0-Å resolution for the TgMcrBC complex and at 4.9-Å resolution for the EcMcrBC complex. The particles were transferred back to RELION using the pyem package (<https://github.com/asarnow/pyem>), re-extracted into 320 × 320 pixel images, and further refined without imposing symmetry to generate maps for the TgMcrBC complex at 2.9-Å resolution and for the EcMcrBC complex at 4.1-Å resolution. CTF refinement and Bayesian polishing improved the maps to resolutions of 2.7 and 3.5 Å, respectively. Finally, the particles were re-extracted into 400 × 400 pixel images. Refinement, CTF refinement and Bayesian polishing yielded the final maps at 2.4-Å resolution for TgMcrBC and at 3.3-Å resolution for EcMcrBC.

Model building and refinement. For the TgMcrB^{AAA} hexamer, the best refined monomer from the X-ray model was used and fit into each subunit density of the 3.1-Å resolution cryo-EM map using UCSF Chimera¹⁰⁵; no other information from the X-ray structure was used. Further iterative refinement cycles between the phenix.real_space_refine command in PHENIX with secondary structure restraints and manual adjustments in COOT yielded the final model for the TgMcrB^{AAA} hexamer.

For the TgMcrB^{AAA}C complex, the final cryo-EM model of the TgMcrB^{AAA} hexamer was manually fit into the 3.7-Å resolution cryo-EM map of the half-complex and refined using the phenix.real_space_refine command in PHENIX with morphing, simulated annealing, and secondary structure restraints. Ab initio model building for TgMcrC was carried out in COOT⁹², guided by secondary structure predictions from SPIDER2¹⁰⁶ and PSIPRED¹⁰⁷. The density for TgMcrC was good up to residue 312, but the remaining C-terminal endonuclease domain was poorly resolved. Therefore, a homology search was performed in I-TASSER^{108–110} for TgMcrC residues 312–458, and a homology model was generated based on the *Saccharolobus solfataricus* Holliday junction resolving enzyme⁵⁷ (PDB: 1OB8; sequence identity: 17%). This homology model was fit into the corresponding density and manual adjustments were performed in COOT. Finally, all built models were combined and iterative cycles of real-space refinement in PHENIX with secondary structure restraints and manual adjustments in COOT were performed, yielding the final model for the TgMcrB^{AAA}C half-complex.

For the TgMcrBC complex, the final cryo-EM model of the TgMcrB^{AAA}C half-complex was manually fit into the 2.4-Å resolution cryo-EM map of the half-complex and refined using the phenix.real_space_refine command in PHENIX with morphing, simulated annealing, and secondary structure restraints. Because the nuclease domain of TgMcrC was poorly resolved, most regions were removed from the model. Finally, iterative cycles of real-space refinement in PHENIX with secondary structure restraints and manual adjustments in COOT were performed, yielding the final model of the TgMcrBC half-complex.

For the EcMcrBC complex, SWISS-MODEL¹¹¹ was used to generate a homology model of EcMcrB based on the TgMcrB^{AAA} structure (sequence identity: 20% across the whole protein, 24% across the AAA+ domain alone), which was manually fit into one subunit in the 3.3-Å resolution cryo-EM map using UCSF Chimera. After manual adjustments in COOT, the corrected model was fit into each subunit of the hexameric EcMcrB density. To build the EcMcrC model, the unique N-terminal domain was removed from the TgMcrC model, and all the residues were mutated to alanine except for the highly conserved residues (based on the sequence alignment of McrC homologs; sequence identity: 13% across the whole protein, 20% across the finger domain alone). This model was manually fit into the corresponding density of the 3.3-Å resolution cryo-EM map in UCSF Chimera, followed by manual adjustment of each residue in COOT. Manual adjustment was guided by secondary structure predictions from SPIDER2 and PSIPRED. Finally, all built models were combined, and iterative cycles of real-space refinement in PHENIX with secondary structure restraints and manual adjustments in COOT yielded the final model of the EcMcrBC half-complex.

All the refinement statistics are summarized in Supplementary Table 4. For model validation, the final model for each map was refined against one of the independent half maps (map 1) of the corresponding map. FSC curves were then calculated between the refined model and half map 1 (work), half map 2 (free), as well as the combined map (Supplementary Figs. 1h, 4f, 7f and 10g).

Statistics and reproducibility. Several cryo-EM data collection sessions including ones for grid screening were performed for each protein complex, but only the best datasets (usually one dataset for each structure except for the TgMcrB^{AAA}C complex) are reported and were processed as described in “Methods” section. The freezing conditions of cryo-EM grids were always duplicated or triplicated with a small variation at the blotting times (7–10 s), providing us reproducible micrograph images and more chances to yield the best ice thickness over one grid for

each data collection. The cryo-EM images shown in Supplementary Figs. 1b, 4a and 7a are representative micrographs in >1000 similar micrographs. The 2D-class averages shown in Supplementary Figs. 1c, 4b, 7b and 10a represent eight classes with the most particles. The negative-stain EM images shown in Fig. 2e are representative micrographs from the reproducible EM data taken every time together with the three independent experiments for GTPase activity assays.

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

Data availability

Protein sequences are available from either Uniprot (<https://www.uniprot.org/>) or the Joint Genome Institute Integrated Microbial Genomes and Microbiomes (JGI IMG/M) database (<https://img.jgi.doe.gov/>). Atomic coordinates and structure factors for TgMcrB^{AAA} have been deposited in the Protein Data Bank (<https://www.rcsb.org/>) with the accession code 6UT3. The B-factor sharpened 3D cryo-EM density maps and atomic coordinates of EcMcrBC, TgMcrB^{AAA}, TgMcrBC, TgMcrB^{AAA}C (full mask), and TgMcrB^{AAA}C (combined) have been deposited in the Worldwide Protein Data Bank (<http://www.wwpdb.org/>) under accession numbers EMD-20867 and 6UT6, EMD-20865 and 6UT4, EMD-20866 and 6UT5, EMD-20868 and 6UT7, EMD-20871 and 6UT8, respectively. The B-factor sharpened 3D cryo-EM density maps of the two multibody-refined body 1 and body 2 have been deposited in the Worldwide Protein Data Bank (wwPDB) under accession number EMD-20869 (body 1) and EMD-20870 (body 2), respectively. Other data are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

Received: 25 February 2020; Accepted: 26 October 2020;

Published online: 20 November 2020

References

- Ro, A. *Tackling Drug-resistant Infections Globally: Final Report and Recommendations* (Review on antimicrobial resistance, London, UK, 2016).
- Solomon, S. L. & Oliver, K. B. Antibiotic resistance threats in the United States: stepping back from the brink. *Am. Fam. Physician* **89**, 938–941 (2014).
- Theuretzbacher, U. & Piddock, L. J. V. Non-traditional antibacterial therapeutic options and challenges. *Cell Host Microbe* **26**, 61–72 (2019).
- Kortright, K. E., Chan, B. K., Koff, J. L. & Turner, P. E. Phage therapy: a renewed approach to combat antibiotic-resistant bacteria. *Cell Host Microbe* **25**, 219–232 (2019).
- Wittebole, X., De Roock, S. & Opal, S. M. A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. *Virulence* **5**, 226–235 (2014).
- Dedrick, R. M. et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. *Nat. Med.* **25**, 730–733 (2019).
- Schooley, R. T. et al. Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant acinetobacter baumannii infection. *Antimicrob. Agents Chemother.* **61**, e00954-17 (2017).
- Hille, F. et al. The biology of CRISPR-Cas: backward and forward. *Cell* **172**, 1239–1259 (2018).
- Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).
- Samson, J. E., Magadan, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
- Stanley, S. Y. & Maxwell, K. L. Phage-encoded Anti-CRISPR defenses. *Annu. Rev. Genet.* **52**, 445–464 (2018).
- Luria, S. E. & Human, M. L. A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.* **64**, 557–569 (1952).
- Weigele, P. & Raleigh, E. A. Biosynthesis and function of modified bases in bacteria and their viruses. *Chem. Rev.* **116**, 12655–12687 (2016).
- Gast, F. U. et al. The recognition of methylated DNA by the GTP-dependent restriction endonuclease McrBC resides in the N-terminal domain of McrB. *Biol. Chem.* **378**, 975–982 (1997).
- Kruger, T., Wild, C. & Noyer-Weidner, M. McrB: a prokaryotic protein specifically recognizing DNA containing modified cytosine residues. *EMBO J.* **14**, 2661–2669 (1995).
- Pieper, U., Schweitzer, T., Groll, D. H., Gast, F.-U. & Pingoud, A. The GTP-binding domain of McrB: more than just a variation on a common theme? *J. Mol. Biol.* **292**, 547–556 (1999).
- Sukackaite, R., Grazulis, S., Tamulaitis, G. & Siksnys, V. The recognition domain of the methyl-specific endonuclease McrBC flips out 5-methylcytosine. *Nucleic Acids Res.* **40**, 7552–7562 (2012).

18. Sutherland, E., Coe, L. & Raleigh, E. A. McrBC: a multisubunit GTP-dependent restriction endonuclease. *J. Mol. Biol.* **225**, 327–348 (1992).
19. Zagorskaite, E., Manakova, E. & Sasnauskas, G. Recognition of modified cytosine variants by the DNA-binding domain of methyl-directed endonuclease McrBC. *FEBS Lett.* **592**, 3335–3345 (2018).
20. Nirwan, N. et al. Hexameric assembly of the AAA+ protein McrB is necessary for GTPase activity. *Nucleic Acids Res.* **47**, 868–882 (2019).
21. Panne, D., Muller, S. A., Wirtz, S., Engel, A. & Bickle, T. A. The McrBC restriction endonuclease assembles into a ring structure in the presence of G nucleotides. *EMBO J.* **20**, 3210–3217 (2001).
22. Panne, D., Raleigh, E. A. & Bickle, T. A. The McrBC endonuclease translocates DNA in a reaction dependent on GTP hydrolysis. *J. Mol. Biol.* **290**, 49–60 (1999).
23. Pieper, U. et al. The GTP-dependent restriction enzyme McrBC from *Escherichia coli* forms high-molecular mass complexes with DNA and produces a cleavage pattern with a characteristic 10-base pair repeat. *Biochemistry* **41**, 5245–5254 (2002).
24. Stewart, F. J., Panne, D., Bickle, T. A. & Raleigh, E. A. Methyl-specific DNA binding by McrBC, a modification-dependent restriction enzyme 11Edited by J. Karn. *J. Mol. Biol.* **298**, 611–622 (2000).
25. O'Driscoll, J. et al. A genetic dissection of the LlaJI restriction cassette reveals insights on a novel bacteriophage resistance system. *BMC Microbiol.* **6**, 40 (2006).
26. O'Sullivan, D. J., Zagula, K. & Klaenhammer, T. R. In vivo restriction by LlaI is encoded by three genes, arranged in an operon with llaIM, on the conjugative Lactococcus plasmid pTR2030. *J. Bacteriol.* **177**, 134–143 (1995).
27. Ohshima, H., Matsuoka, S., Asai, K. & Sadaie, Y. Molecular organization of intrinsic restriction and modification genes BsuM of *Bacillus subtilis* Marburg. *J. Bacteriol.* **184**, 381–389 (2002).
28. Hosford, C. J. & Chappie, J. S. The crystal structure of the *Helicobacter pylori* LlaJLR1 N-terminal domain provides a model for site-specific DNA binding. *J. Biol. Chem.* **293**, 11758–11771 (2018).
29. Hosford, C. J., Bui, A. Q. & Chappie, J. S. The structure of the *Thermococcus gammatolerans* McrB N-terminal domain reveals a new mode of substrate recognition and specificity among McrB homologs. *J. Biol. Chem.* **295**, 743–756 (2020).
30. Yang, X., Xu, M. & Yang, S. T. Restriction modification system analysis and development of in vivo methylation for the transformation of *Clostridium cellulovorans*. *Appl. Microbiol. Biotechnol.* **100**, 2289–2299 (2016).
31. Snider, J., Thibault, G. & Houry, W. A. The AAA+ superfamily of functionally diverse proteins. *Genome Biol.* **9**, 216 (2008).
32. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43 (1999).
33. Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).
34. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the AAA+ nucleotide binding pocket. *Biochim. Biophys. Acta* **1823**, 2–14 (2012).
35. Pieper, U., Brinkmann, T., Krüger, T., Noyer-Weidner, M. & Pingoud, A. Characterization of the interaction between the restriction endonuclease McrBC from *E. coli* and its cofactor GTP11Edited by J. Karn. *J. Mol. Biol.* **272**, 190–199 (1997).
36. Pieper, U., Schweitzer, T., Groll, D. H. & Pingoud, A. Defining the location and function of domains of McrB by deletion mutagenesis. *Biol. Chem.* **380**, 1225–1230 (1999).
37. Bourne, H. R., Sanders, D. A. & McCormick, F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117–127 (1991).
38. Azmi, I. F. et al. ESCRT-III family members stimulate Vps4 ATPase activity directly or via Vta1. *Dev. Cell* **14**, 50–61 (2008).
39. Jessop, M. et al. Structural insights into ATP hydrolysis by the MoxR ATPase RavA and the LdcI-RavA cage-like complex. *Commun. Biol.* **3**, 46 (2020).
40. Paduch, M., Jelen, F. & Otlewski, J. Structure of small G proteins and their regulators. *Acta Biochim. Pol.* **48**, 829–850 (2001).
41. Nirwan, N. et al. Structure-based mechanism for activation of the AAA+ GTPase McrB by the endonuclease McrC. *Nat. Commun.* **10**, 3058 (2019).
42. Gai, D., Zhao, R., Li, D., Finkielstein, C. V. & Chen, X. S. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**, 47–60 (2004).
43. Monroe, N., Han, H., Shen, P. S., Sundquist, W. I. & Hill, C. P. Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. *Elife* **6**, e24487 (2017).
44. Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **146**, 11–31 (2004).
45. Vetter, I. R. & Wittinghofer, A. Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Q. Rev. Biophys.* **32**, 1–56 (1999).
46. Chappie, J. S., Acharya, S., Leonard, M., Schmid, S. L. & Dyda, F. G domain dimerization controls dynamin's assembly-stimulated GTPase activity. *Nature* **465**, 435–440 (2010).
47. Wittinghofer, A. GTP and ATP hydrolysis in biology. *Biopolymers* **105**, 419–421 (2016).
48. Miller, J. M. & Enemark, E. J. Fundamental characteristics of AAA+ protein family structure and function. *Archaea* **2016**, 9294307 (2016).
49. Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* **317**, 41–72 (2002).
50. Tsai, J. T. et al. Crystal structure of the YcjX stress protein reveals a Ras-like GTP-binding protein. *J. Mol. Biol.* **431**, 3179–3190 (2019).
51. Pieper, U., Schweitzer, T., Groll, D. H., Gast, F. U. & Pingoud, A. The GTP-binding domain of McrB: more than just a variation on a common theme? *J. Mol. Biol.* **292**, 547–556 (1999).
52. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
53. Nakae, S. et al. Structure of the EndoMS-DNA complex as mismatch restriction endonuclease. *Structure* **24**, 1960–1971 (2016).
54. Knizewski, L., Kinch, L. N., Grishin, N. V., Rychlewski, L. & Ginalski, K. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. *BMC Struct. Biol.* **7**, 40 (2007).
55. Pieper, U. & Pingoud, A. A mutational analysis of the PD...D/EXK motif suggests that McrC harbors the catalytic center for DNA cleavage by the GTP-dependent restriction enzyme McrBC from *Escherichia coli*. *Biochemistry* **41**, 5236–5244 (2002).
56. Perona, J. J. & Martin, A. M. Conformational transitions and structural deformability of EcoRV endonuclease revealed by crystallographic analysis. *J. Mol. Biol.* **273**, 207–225 (1997).
57. Middleton, C. L., Parker, J. L., Richard, D. J., White, M. F. & Bond, C. S. Crystallization and preliminary X-ray diffraction studies of Hje, a Holliday junction resolving enzyme from *Sulfolobus solfataricus*. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 171–173 (2003).
58. de la Pena, A. H., Goodall, E. A., Gates, S. N., Lander, G. C. & Martin, A. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. *Science* **362**, eaav0725 (2018).
59. Enemark, E. J. & Joshua-Tor, L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**, 270–275 (2006).
60. Gates, S. N. et al. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. *Science* **357**, 273–279 (2017).
61. Puchades, C. et al. Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* **358**, eaao0464 (2017).
62. Ripstein, Z. A., Huang, R., Augustyniak, R., Kay, L. E. & Rubinstein, J. L. Structure of a AAA+ unfoldase in the process of unfolding substrate. *Elife* **6**, e25754 (2017).
63. Twomey, E. C. et al. Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* **365**, eaax1033 (2019).
64. Zehr, E. et al. Katanin spiral and ring structures shed light on power stroke for microtubule severing. *Nat. Struct. Mol. Biol.* **24**, 717–725 (2017).
65. Zhao, M. et al. Mechanistic insights into the recycling machine of the SNARE complex. *Nature* **518**, 61–67 (2015).
66. Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865–877 (2007).
67. Mishra, A. K. & Lambright, D. G. Invited review: small GTPases and their GAPs. *Biopolymers* **105**, 431–448 (2016).
68. Scheffzek, K. et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338 (1997).
69. Tesmer, J. J., Berman, D. M., Gilman, A. G. & Sprang, S. R. Structure of RGS4 bound to ALF4-activated G(i alpha1): stabilization of the transition state for GTP hydrolysis. *Cell* **89**, 251–261 (1997).
70. Scrima, A., Thomas, C., Deaconescu, D. & Wittinghofer, A. The Rap-RapGAP complex: GTP hydrolysis without catalytic glutamine and arginine residues. *EMBO J.* **27**, 1145–1153 (2008).
71. Forgac, M. Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. *Nat. Rev. Mol. Cell Biol.* **8**, 917–929 (2007).
72. Iwata, M. et al. Crystal structure of a central stalk subunit C and reversible association/dissociation of vacuole-type ATPase. *Proc. Natl Acad. Sci. USA* **101**, 59–64 (2004).
73. Stock, D., Gibbons, C., Arechaga, I., Leslie, A. G. & Walker, J. E. The rotary mechanism of ATP synthase. *Curr. Opin. Struct. Biol.* **10**, 672–679 (2000).
74. Yasuda, R., Noji, H., Yoshida, M., Kinoshita, K. Jr. & Itoh, H. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. *Nature* **410**, 898–904 (2001).
75. Meagher, M., Epling, L. B. & Enemark, E. J. DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nat. Commun.* **10**, 3117 (2019).

76. Peters, J. P. 3rd. & Maher, L. J. DNA curvature and flexibility in vitro and in vivo. *Q. Rev. Biophys.* **43**, 23–63 (2010).
77. Bair, C. L. & Black, L. W. A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J. Mol. Biol.* **366**, 768–778 (2007).
78. Dharmalingam, K. & Goldberg, E. B. Mechanism localisation and control of restriction cleavage of phage T4 and lambda chromosomes in vivo. *Nature* **260**, 406–410 (1976).
79. Leonard, M., Song, B. D., Ramachandran, R. & Schmid, S. L. Robust colorimetric assays for dynamin's basal and stimulated GTPase activities. *Methods Enzymol.* **404**, 490–503 (2005).
80. Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification—powerful tools in modern electron microscopy. *Biol. Proced. Online* **6**, 23–34 (2004).
81. Sauter, N. Visualizing the raw diffraction pattern with LABELIT. *Comput. Crystallogr. Newsl.* **2**, 15–24 (2011).
82. Paithankar, K. S., Owen, R. L. & Garman, E. F. Absorbed dose calculations for macromolecular crystals: improvements to RADDOSE. *J. Synchrotron. Radiat.* **16**, 152–162 (2009).
83. Bourenkov, G. P. & Popov, A. N. A quantitative approach to data-collection strategies. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 58–64 (2006).
84. Batty, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 271–281 (2011).
85. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
86. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
87. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
88. Foadi, J. et al. Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1617–1632 (2013).
89. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
90. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 479–485 (2010).
91. Skubák, P. & Pannu, N. S. Automatic protein structure solution from weak X-ray data. *Nat. Commun.* **4**, 2777 (2013).
92. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
93. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).
94. Bunkóczi, G. et al. Phaser. MRage: automated molecular replacement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 2276–2286 (2013).
95. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
96. Mastronarde, D. N. SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc. Microanal.* **9**, 1182–1183 (2003).
97. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
98. Zivanov, J., Nakane, T. & Scheres, S. H. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCr* **6**, 5–17 (2019).
99. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *Elife* **7**, e36861 (2018).
100. 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).
101. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
102. Rohou, A. & Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
103. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
104. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
105. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
106. Yang, Y. et al. In *Prediction of Protein Secondary Structure* Vol. 1484, 55–63 (Springer, 2017).
107. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, 349–357 (2013).
108. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
109. Yang, J. et al. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
110. Yang, J. & Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* **43**, 174–181 (2015).
111. Bienert, S. et al. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* **45**, 313–319 (2016).

Acknowledgements

We are grateful to Mark Ebrahim and Johanna Sotiris for help with grid screening and data collection at the Evelyn Gruss Lipper Cryo-Electron Microscopy Resource Center at The Rockefeller University. We thank Drs. Richard Cerione and Holger Sondermann for critical reading of the manuscript and the Northeastern Collaborative Access Team (NE-CAT) beamline staff at the Advanced Photon Source (APS) for assistance with remote X-ray data collection. This work was supported by National Institutes of Health Grant GM120242 (to J.S.C.) and is based upon research conducted at NE-CAT beamlines (24-ID-C and 24-ID-E) under the general user proposals GUP-51113 and GUP-41829 (PI: J.S.C.). NE-CAT beamlines are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Pilatus 6 M detector on 24-ID-C beamline is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. J.S.C. is a Meinig Family Investigator in the Life Sciences.

Author contributions

Y.N., H.S., C.J.H., T.W., and J.S.C. designed the study and analysed data. C.J.H. and Y.N. cloned purified all constructs and carried out biochemical assays. Y.N. and H.S. collected cryo-EM data, carried out image processing, and built the atomic-resolution cryo-EM models. C.J.H. and Y.N. collected X-ray diffraction data, determined the X-ray structure, and built the X-ray model. Y.N., H.S., T.W., and J.S.C. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information


Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-19735-4>.

Correspondence and requests for materials should be addressed to T.W. or J.S.C.

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

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

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

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020