

1 **ComM is a hexameric helicase that promotes branch migration during** 2 **natural transformation in diverse Gram-negative species**

3

4 Thomas M. Nero¹, Triana N. Dalia¹, Joseph Che-Yen Wang², David T. Kysela¹, Matthew L.
5 Bochman^{3,*}, and Ankur B. Dalia^{1,*}

6

7 ¹Department of Biology, ²Electron Microscopy Center, ³Molecular and Cellular Biochemistry
8 Department, Indiana University, Bloomington, IN 47401

9

10 *Authors for correspondence – Ankur B. Dalia, ankdalia@indiana.edu, and Matthew L.
11 Bochman, bochman@indiana.edu

12

13 **ABSTRACT**

14 Acquisition of foreign DNA by natural transformation is an important mechanism of adaptation
15 and evolution in diverse microbial species. Here, we characterize the mechanism of ComM, a
16 broadly conserved AAA+ protein previously implicated in homologous recombination of
17 transforming DNA (tDNA) in naturally competent Gram-negative bacterial species. *In vivo*, we
18 find that ComM was required for efficient comigration of linked genetic markers in *Vibrio*
19 *cholerae* and *Acinetobacter baylyi*, which is consistent with a role in branch migration. Also,
20 ComM was particularly important for integration of tDNA with increased sequence heterology,
21 suggesting that this branch migration promotes the acquisition of novel DNA sequences. *In*
22 *vitro*, we show that purified ComM binds ssDNA, oligomerizes into a hexameric ring, and has 5'
23 to 3' helicase activity. Based on these data, we propose a model for tDNA integration during
24 natural transformation. This study provides mechanistic insight into the enigmatic steps
25 involved in tDNA integration and uncovers the function of a protein required for this conserved
26 mechanism of horizontal gene transfer.

27

28

29

1 INTRODUCTION

2 Natural competence is a physiological state in which some bacterial species can take up free
3 DNA from the environment. Some competent species regulate the genes required for this
4 process and, depending on the organism, competence can be induced in response to the
5 availability of certain nutrients, quorum sensing pathways, or by DNA damage / stress (1). The
6 Gram-negative bacterium *Vibrio cholerae* is activated for competence during growth on chitin,
7 a polymer of β 1,4-linked N-acetyl glucosamine (2). Chitin is the primary constituent of
8 crustacean exoskeletons and is commonly found in the aquatic environment where this
9 facultative pathogen resides. Soluble chitin oligosaccharides indirectly induce expression of
10 TfoX, the master regulator of competence (3,4), which regulates expression of competence-
11 related genes in concert with HapR, the master regulator of quorum sensing (5,6).

12 *Acinetobacter baylyi* ADP1, on the other hand, is a naturally competent Gram-negative microbe
13 that is constitutively active for competence throughout exponential growth (7).

14

15 During competence, dsDNA is bound extracellularly, however, only a single strand of this DNA is
16 transported into the cytoplasm. Competent bacterial species may use this ingested DNA as a
17 source of nutrients, however, if this DNA has sufficient homology to the host chromosome, the
18 incoming DNA can also be integrated into the bacterial genome by homologous recombination
19 (8). This process of DNA uptake and integration is referred to as natural transformation. As a
20 result, natural transformation is an important mechanism of horizontal gene transfer and can
21 lead to the repair of damaged DNA or facilitate acquisition of novel genetic information.

22 Homologous recombination of single-stranded transforming DNA (tDNA) with the host
23 chromosome requires the function of RecA, which facilitates homology searching and initiates
24 strand invasion of tDNA through the formation of a displacement loop (D-loop). Following RecA
25 mediated strand invasion, DNA junctions of this D-loop can then be moved in a process known
26 as branch migration to increase or decrease the amount of tDNA integrated. Then, by a
27 presently unresolved mechanism, this intermediate is resolved to stably integrate tDNA into the
28 host chromosome. The molecular details involved in the integration of tDNA downstream of
29 RecA strand invasion, however, are poorly understood.

1
2 One previously studied gene from the competent species *Haemophilus influenzae*, *comM*, is not
3 required for DNA uptake but is required for the integration of tDNA into the host chromosome,
4 a phenotype consistent with ComM playing a role in homologous recombination during natural
5 transformation (9). The function of ComM, however, has remained unclear. Here, through both
6 *in vivo* and *in vitro* characterization of ComM in *V. cholerae* and *A. baylyi*, we uncover that this
7 protein functions as a hexameric helicase to aid in branch migration during this conserved
8 mechanism of horizontal gene transfer.

9

10 **MATERIALS AND METHODS**

11 Bacterial strains and growth conditions

12 The parent *V. cholerae* strain used throughout this study is E7946 (10), while the *A. baylyi* strain
13 used is ADP1 (also known as BD413) (11). A description of all strains used in this study is listed
14 in Table S1. Strains were routinely grown in LB broth and plated on LB agar. When required,
15 media was supplemented with 200 µg/mL spectinomycin, 50 µg/mL kanamycin, 100 µg/mL
16 carbenicillin, 10 µg/mL trimethoprim, or 40 µg/mL X-Gal.

17

18 Construction of mutants and transforming DNA

19 Linear PCR product was constructed using splicing-by-overlap extension PCR exactly as
20 previously described (12). All primers used to construct and detect mutant alleles are described
21 in **Table S2**. The pBAD18 Kan plasmid used as tDNA was purified from TG1, a *recA+* *E. coli* host.
22 Mutants were made by cotransformation exactly as previously described (13).

23

24 Transformation assays

25 Transformation assays of *V. cholerae* on chitin were performed exactly as previously described
26 (3). Briefly, $\sim 10^8$ CFU of mid-log *V. cholerae* were incubated statically in instant ocean medium
27 (7g/L; aquarium systems) containing chitin from shrimp shells (Sigma) for 16-24 hours at 30°C.
28 Then, tDNA was added (500 ng for linear products containing an antibiotic resistance cassette
29 inserted at the non-essential locus VC1807 and 1500 ng for plasmid DNA), and reactions were

1 incubated at 30°C for 5-16 hours. Reactions were outgrown with the addition of LB medium to
2 each reaction by shaking at 37°C for 1-3 hours and then plated for quantitative culture onto
3 medium to select for the tDNA (transformants) or onto plain LB (total viable counts).
4 Transformation efficiency is shown as the number of transformants / total viable counts. In
5 cases where no colonies were observed, efficiencies were denoted as below the limit of
6 detection (LOD) for the assay.

7
8 Chitin-independent transformation assays were performed exactly as previously described
9 using strains that contain an IPTG inducible P_{tac} promoter upstream of the native TfoX gene
10 (14). Briefly, strains were grown overnight with 100 μ M IPTG. Then, $\sim 10^8$ cells were diluted into
11 instant ocean medium, and tDNA was added. Reactions were incubated statically for 5 hours
12 and then outgrown by adding LB and shaking at 37°C for 1-3 hours. Reactions were plated for
13 quantitative culture as described above.

14
15 For transformation of *A. baylyi* ADP1, strains were grown overnight in LB media. Then, $\sim 10^8$
16 cells were diluted into fresh LB medium, and tDNA was added (~ 100 ng). Reactions were
17 incubated at 30°C with agitation for 5 hours and then plated for quantitative culture as
18 described above.

19 20 Protein expression and purification

21 ComM and Pif1 were cloned into StrepII expression vectors, expressed in Rosetta 2(DE3) pLysS
22 cells using autoinduction medium (15), and purified using Strep-Tactin Sepharose (IBA). For
23 details please see **Supplementary Methods**.

24 25 Electrophoretic mobility shift assay

26 The ssDNA probe (BBC742) and dsDNA probe (annealed BBC742 and BBC743) were labeled
27 using Cy5 dCTP (GE Healthcare) and Terminal deoxynucleotidyl Transferase (TdT; Promega).
28 Reactions were composed of 10 mM Tris-HCl pH 7.5, 20 mM KCl, 1 mM DTT, 10% Glycerol, 100
29 μ g/ml BSA, 9 nM Cy5 labeled probe, purified ComM protein at the indicated concentration, and

1 0.1 mM ATP where indicated. Reactions were incubated at RT for 30 minutes and run on a 13%
2 Tris-borate acrylamide gel at 150 V for 45 min. Probes were detected using a Chemidoc MP
3 imaging system (Biorad).

4

5 Blue native PAGE

6 Blue native PAGE was performed essentially as previously described (16). Purified ComM (2.5
7 μM) was incubated for 30 min at room temperature in reaction buffer [10 mM Tris-HCl pH 7.5,
8 20 mM KCl, 1 mM DTT, 10% Glycerol] with 5 mM ATP and/or 5 μM ssDNA (oligo ABD363) where
9 indicated. For additional details see **Supplementary Methods**.

10

11 Negative stain electron microscopy

12 For negative stain electron microscopy, sample was prepared by applying 4 μL of ComM-ATP-
13 ssDNA solution onto a glow-discharged continuous carbon film coated copper grid (EMS) and
14 stained with 0.75% (w/v) uranyl formate. EM micrographs were acquired using a 300 kV JEM-
15 3200FS electron microscopy with 20-eV energy slit under low dose conditions ($\leq 20 \text{ e}^-/\text{\AA}^2$) on a
16 Gatan UltraScan 4000 4k x 4k CCD camera. Additional details for EM image collection and
17 analysis are provided in the **Supplementary Methods**. The EM electron density map has been
18 deposited to EMDataBank.org with the accession number EMD-8575.

19

20 Phylogenetic Trees

21 ComM homologs were identified using a protein BLAST search of diverse bacterial genomes
22 followed by phylogenetic analysis. Starting with a broadly representative set of genomes
23 adapted from Wu and Eisen (17), an initial ComM candidate pool was generated from a
24 comprehensive protein BLAST search of all predicted CDS translations against the *V. cholerae*
25 ComM allele with a 0.001 e-value cutoff. Subsequent sequence alignment with MUSCLE (18)
26 and maximum likelihood phylogenetic reconstruction with FastTree (19) identified a single
27 clade of alleles with high sequence similarity (generally well above 50%) to *V. cholerae* ComM
28 (data not shown), with the remaining alleles excluded from further analysis. After manually
29 pruning divergent alleles with alignments covering <70% of ComM, the retained sequences

1 comprised the set of true ComM orthologs used in the final sequence alignment and ComM
2 phylogeny reconstruction. For whole genome phylogenetic reconstruction, Phylosift (20)
3 identified and aligned a default set of 36 highly conserved marker genes. FastTree was used for
4 initial reconstruction, whereas RAxML (21) subsequently estimated the maximum likelihood
5 phylogeny for a reduced set of representative genomes under the LG (22) substitution model
6 with gamma-distributed rate variation.

7

8 Helicase assays

9 Helicase assay substrates were 5' end-labeled with T4 polynucleotide kinase (T4 PNK; NEB) and
10 γ [³²P]-ATP. DNA unwinding was assessed by incubating the indicated concentrations of helicase
11 with 5 mM ATP and 0.1 nM radiolabelled fork in resuspension buffer (25 mM Na-HEPES (pH
12 7.5), 5% (v/v) glycerol, 300 mM NaOAc, 5 mM MgOAc, and 0.05% Tween-20). Reactions were
13 incubated at 37°C for 30 min and stopped with the addition of 1x Stop-Load dye (5% glycerol,
14 20 mM EDTA, 0.05% SDS, and 0.25% bromophenol blue) supplemented with 400 μ g/mL SDS-
15 Proteinase K followed by a 10-min incubation at 37°C. Unwound DNA was separated on 8% 19:1
16 acrylamide:bis-acrylamide gels in TBE buffer at 10 V/cm. Gels were dried under vacuum and
17 imaged using a Typhoon 9210 Variable Mode Imager. DNA unwinding was quantified using
18 ImageQuant 5.2 software. Primers used for probes and additional details can be found in the
19 **Supplementary Methods.**

20

21 DNA damage assay

22 For DNA damage assays, $\sim 10^8$ cells of a midlog culture in instant ocean medium were treated
23 with the indicated concentration of MMS or MMC for 1 hour at 30°C. To determine viable
24 counts, reactions were plated for quantitative culture on LB agar.

25

26 GFP-ComM western blots

27 Strains were grown with or without 100 μ M IPTG. Cell lysates were run on 10% SDS PAGE gels
28 and electrophoretically transferred to PVDF. This membrane was then blotted with primary
29 rabbit anti-GFP (Invitrogen) or mouse anti-RpoA (Biolegend) antibodies and a goat anti-rabbit

1 or anti-mouse IRDye 800CW (LI-COR) secondary as appropriate. Bands were detected using a LI-
2 COR Odyssey classic infrared imaging system.

3

4 **RESULTS**

5 *ComM is required for integration of tDNA during natural transformation*

6 Previously, we performed an unbiased transposon-sequencing screen (Tn-seq) in *V. cholerae* to
7 identify genes involved in natural transformation (3). One gene identified in that screen was
8 VC0032, which encodes a homolog of *comM* from *H. influenzae*. ComM was previously
9 implicated in the integration of tDNA during natural transformation in *H. influenzae* (9). To
10 determine if this was also the case in *V. cholerae*, we performed chitin-dependent natural
11 transformation assays using two distinct sources of tDNA. One was a linear PCR product that
12 inserts an antibiotic resistance cassette at a non-essential locus, while the other was a
13 replicating plasmid that lacks any homology to the host genome. We hypothesized that
14 transformation with linear product requires both DNA uptake and chromosomal integration,
15 while transformation with the plasmid only requires DNA uptake. To test this, we transformed a
16 recombination deficient $\Delta recA$ strain of *V. cholerae*, and found that, as expected, this strain
17 could not be transformed with linear product but could be transformed with a replicating
18 plasmid, consistent with plasmid transformation being recombination-independent (**Fig. S1**).
19 Additionally, plasmid transformation in this assay is dependent on natural competence, as
20 mutants in genes required for uptake ($\Delta pilA$) (5) and cytoplasmic protection ($\Delta dprA$) of tDNA
21 are not transformed (5,23,24) (**Fig. S1**). Using this assay, we find that a *comM* mutant
22 ($\Delta VC0032$) in *V. cholerae* displays a ~100-fold reduction for transformation with linear product,
23 while rates of plasmid transformation were equal to the WT (**Fig. 1A**). This is consistent with
24 ComM playing a role downstream of DNA uptake, and potentially during recombination. These
25 assays were performed on chitin to induce the natural competence of this organism. To
26 determine if ComM is playing a role specifically downstream of competence induction, we
27 performed a chitin-independent transformation assay by overexpressing the competence
28 regulator TfoX. Under these conditions, a *comM* mutant still had reduced rates of
29 transformation when transformed with a linear PCR product (**Fig. 1B**). Cumulatively, these

1 results are consistent with ComM playing a role in the late steps of transformation downstream
2 of DNA uptake.

3

4 To confirm that the phenotypes observed are due to mutation of *comM*, we complemented
5 strains by integrating an IPTG inducible P_{tac} -*comM* construct at a heterologous site on the
6 chromosome (at the *lacZ* locus). Our previous work has indicated that this expression construct
7 is leaky (14) (**Fig. S2**), and consistent with this, we observe complementation even in the
8 absence of inducer (**Fig. 1C**).

9

10 *ComM promotes branch migration through heterologous sequences in vivo*

11 *V. cholerae* ComM is a predicted AAA+ ATPase, and members of this family have diverse
12 functions (25). To determine if ComM had structural similarity to any AAA+ ATPase of known
13 function, we submitted the primary sequence of this protein to the Phyre2 server (26). Despite
14 a lack of significant homology by BLAST, this analysis revealed structural similarity to MCM2-7,
15 the replicative helicase of eukaryotes (27,28). As a result, we explored whether ComM
16 functions as a helicase to promote branch migration during natural transformation.

17

18 To test this *in vivo*, we assessed comigration of linked genetic markers on a linear tDNA product
19 (**Fig. 2A**). If branch migration during transformation is efficient, we hypothesized that both
20 markers would be integrated into the host chromosome. However, if branch migration is
21 inefficient, we hypothesized that we may observe integration of one marker but not the other.

22 For this assay, we generated two tDNA constructs that contained a Spec^R marker upstream of
23 *lacZ* as well as a genetically linked point mutation in the *lacZ* gene that was either 820bp or
24 245bp downstream of the Spec^R marker (**Fig 2A**). We selected for the Spec^R marker and
25 screened for integration of the linked *lacZ* mutation as an indirect measure of branch migration.

26 To prevent post-recombination repair of the *lacZ* allele by the mismatch repair (MMR) system,
27 these experiments were performed in *mutS* mutant backgrounds. As with previous
28 experiments, a *comM* mutant is severely reduced for transformation efficiency for the Spec^R
29 marker compared to the parent (**Fig. 2A, left and 1A**). Of those cells that integrated the Spec^R

1 marker, the comigration efficiency of the *lacZ* mutation was higher than 90% for both products
2 (820bp and 245bp) in the parent strain background (**Fig. 2A, right**), which is consistent with
3 highly efficient branch migration in this background. When *comM* is deleted, however, the
4 comigration efficiency for the *lacZ* mutation drops significantly for both products compared to
5 the parent strain, and the reduction is more severe for the product where the *lacZ* mutation is
6 farther away (**Fig. 2A, right**). Cumulatively, these data suggest that *comM* may play a role in
7 branch migration during natural transformation to increase the amount of tDNA integrated into
8 the host chromosome.

9
10 To promote adaptation and evolution, tDNA integrated during natural transformation must be
11 heterologous to the host chromosome. So next, we decided to test whether ComM promotes
12 integration of heterologous tDNA. To that end, we created strains that contain an inactivated
13 Tm^R marker integrated in the chromosome. The marker was inactivated with either a 29bp
14 deletion or a nonsense point mutation. We then transformed these strains with tDNA that
15 would restore the Tm^R marker. Again, to eliminate any confounding effects of MMR, we
16 performed these experiments in *mutS* mutant backgrounds. First, we find that integration of a
17 point mutation is similar to a 29bp insertion in the parent strain background, indicating that in
18 the presence of ComM, tDNA is efficiently integrated regardless of sequence heterology. In the
19 *comM* mutant, however, we find that a point mutation is significantly easier to integrate
20 compared to the 29bp insertion (**Fig. 2B**). This finding is consistent with ComM promoting
21 branch migration through heterologous sequences during natural transformation; however, in
22 its absence, the integration of heterologous sequences is unfavored.

23 24 *ComM* hexamerizes in the presence of ATP and ssDNA

25 Because our *in vivo* data suggested that ComM acts as a branch migration factor, we next
26 decided to test the biochemical activity of this protein *in vitro*. First, we determined that N-
27 terminally tagged ComM was functional *in vivo* while a C-terminally tagged fusion was not (**Fig.**
28 **S2**). Furthermore, recent studies indicate that *comM* is part of the competence regulon (29).
29 Using a strain where *comM* was N-terminally tagged at the native locus, we found that *comM*

1 protein levels are increased under competence inducing conditions (TfoX overexpression),
2 indicating that native regulation of tagged *comM* is maintained (**Fig. S2**).
3
4 To characterize ComM *in vitro*, we expressed StrepII-ComM (N-terminal tag) in *E. coli* and
5 purified it to homogeneity. The peak of recombinant ComM eluting from preparative gel
6 filtration chromatography had a calculated molecular weight of 57 kDa (**Fig. S3**). As the
7 predicted mass of ComM is 61 kDa, this suggests that ComM exists as a monomer in solution.
8 Many helicases oligomerize in their active state, including all known MCM family helicases (30).
9 So, we next tested whether purified ComM oligomerizes *in vitro*. Because ComM is a predicted
10 ATPase and may interact with DNA, we hypothesized that these factors may be required for its
11 oligomerization and activity. To assess oligomerization, we performed blue-native PAGE (16). In
12 this assay, ComM appears to oligomerize robustly in the presence of ATP and ssDNA, with some
13 oligomerization also observed in the presence of ATP alone (**Fig. 3A**). This latter observation,
14 however, may be due to a small amount of contaminating ssDNA that remains bound to ComM
15 during purification. Defining the number of subunits in this oligomer was unreliable by blue-
16 native PAGE due to lack of resolving power by the gel. However, the higher molecular weight
17 species generated in the presence of ATP and ssDNA was likely larger than a dimer. Therefore,
18 we attempted to visualize ComM oligomers by negative stain transmission electron microscopy
19 (TEM). In the absence of ligands, ComM particles were small and uniform, consistent with our
20 gel filtration results and demonstrated the purity of our protein preparations. In the presence
21 of ATP and ssDNA, we observed ring-like densities for ComM that upon 2-D averaging revealed
22 that this protein forms a hexameric ring (**Fig. 3B and 3C**). Furthermore, we generated a 3-D
23 reconstruction from the negative stain TEM images of ComM in the presence of ATP and ssDNA
24 (~ 13.8 Å resolution, see **Supplementary Methods**), which revealed that ComM forms a three-
25 tiered barrel-like structure with a large opening on both ends (**Fig. 3D**). The pore on the bottom
26 of this barrel is ~ 18 Å, which can accommodate ssDNA but not dsDNA, while the 26 Å pore is
27 able to accommodate dsDNA. The size of these pores, however, may be an underestimate due
28 to the stain used during EM and/or ssDNA that may be bound and averaged into the 3-D
29 construction. Regardless, these data suggest that ComM forms a hexameric ring, consistent

1 with structures adopted by many AAA+ helicases (25). ComM also oligomerized in the presence
2 of ADP and the non-hydrolysable ATP analog AMP-PNP, indicating that ATP is not a strict
3 requirement for hexamer formation (**Fig. S4**).

4

5 ComM binds ssDNA and dsDNA in the presence of ATP

6 Because ATP was required for oligomerization in the presence of ssDNA, we hypothesized that
7 it would also be required for DNA binding. To test this, we mutated the conserved lysine in the
8 Walker A motif of ComM, which in other AAA+ ATPases, promotes ATP binding (25). *In vivo*, we
9 find that this mutation abrogates ComM function (**Fig. S5A**), while protein stability is
10 maintained (**Fig. S2**). To determine if ATP binding was required for ComM to bind ssDNA, we
11 performed electrophoretic mobility shift assays (EMSAs) using purified ComM and ComM^{K224A} *in*
12 *vitro*. We find that WT ComM binds both ssDNA and dsDNA in an ATP-dependent manner (**Fig.**
13 **S5B and C**). Consistent with this, ComM^{K224A} does not bind ssDNA even in the presence of ATP
14 (**Fig. S5B**). Taken together, these results indicate that ATP binding is necessary for ComM to
15 bind DNA and function during natural transformation.

16

17 ComM has helicase activity in vitro

18 Above, we demonstrate that ComM may play a role in branch migration *in vivo*. Some branch
19 migration factors (e.g., RuvABC and RecG (31,32)) display helicase activity. So next, we tested
20 the helicase activity of purified ComM *in vitro*. We observed enzymatic unwinding of a forked
21 DNA substrate with increasing concentrations of ComM. Assuming that ComM hexamers are
22 the active oligomeric form, the helicase activity had an apparent K_M of 50.8 nM (**Fig. 4A**). As
23 expected, the purified ATP binding mutant ComM^{K224A} did not display helicase activity in this
24 assay compared to WT ComM and Pif1, a previously characterized helicase that served as a
25 positive control in these assays (**Fig. 4B**). Furthermore, the non-hydrolysable ATP analog ATP γ S
26 inhibited helicase activity, which is consistent with ATP hydrolysis being required for ComM
27 function (**Fig. 4B**).

28

1 Most helicases exhibit a preferred directionality (either 5' to 3' or 3' to 5'). A forked DNA
2 substrate, however, does not distinguish between these activities. To determine whether
3 ComM had a preference, we tested DNA substrates containing a 5' overhang (5' to 3'), a 3'
4 overhang (3' to 5'), or blunt ends in our helicase assay. We only observed robust helicase
5 activity with the substrates containing a 5' overhang (Fork and 5' Tail), suggesting a strong 5' to
6 3' preference (**Fig. 4C**). Taken together, these data suggest that ComM is an ATP-dependent 5'
7 to 3' helicase.

8

9 *ComM is broadly conserved*

10 Next, we assessed how broadly conserved ComM was among bacterial species. Homologs of
11 ComM are found almost ubiquitously among Gram-negative species, including all known Gram-
12 negative naturally competent microbes (**Fig. 5** and **Fig. S6**). Among this group, only select
13 species lacked a ComM homolog, suggesting that loss of ComM occurred relatively recently
14 (**Fig. 5** and **Fig. S6**). By contrast, we see a pervasive lack of ComM homologs among species
15 within the Bacilli and Mollicutes, (**Fig. 5** and **Fig. S6**), suggesting that ComM may have been lost
16 in a common ancestor for these two Classes. Interestingly, all known naturally competent
17 Gram-positive species fall within the Bacilli group, indicating that branch migration during
18 natural transformation must occur via a ComM-independent mechanism in these microbes.

19

20 *ComM promotes branch migration in Acinetobacter baylyi*

21 Because ComM is broadly conserved among Gram-negative naturally competent microbes,
22 next, we tested its role during natural transformation in the model competent species *A. baylyi*
23 ADP1. As in *V. cholerae*, a *comM* mutant of *A. baylyi* displayed greatly reduced rates of natural
24 transformation when using linear tDNA (**Fig. 6A, left**). Also, this mutant displayed reduced
25 comigration of linked genetic markers, consistent with *comM* playing a role in branch migration
26 (**Fig. 6A, right**). Additionally, we observed that ComM is required for integration of tDNA
27 containing larger regions of heterologous sequence (**Fig. 6B**). These data are consistent with
28 what was observed in *V. cholerae* (**Fig. 2**) and suggests that ComM is a conserved branch

1 migration factor important for natural transformation in diverse Gram-negative bacterial
2 species.

3

4 *RadA is not required for natural transformation in V. cholerae.*

5 While ComM is broadly conserved, it is absent in all of the known Gram-positive naturally
6 competent species (**Fig. 5**). In the Gram-positive *Streptococcus pneumoniae*, mutants of *radA*
7 (also known as *sms*) display reduced rates of natural transformation but are not affected at the
8 level of tDNA uptake, similar to what is observed for *V. cholerae comM* mutants in this study
9 (33) and in *H. influenzae* (9). Similar results are also seen in *radA* mutants in *Bacillus subtilis*
10 (34). *E. coli* RadA has recently been shown to promote branch migration during RecA-mediated
11 strand exchange (35). Also, it was very recently demonstrated that RadA is a hexameric helicase
12 that promotes bidirectional branch migration during natural transformation in *S. pneumoniae*
13 (36). Interestingly, *radA* is broadly conserved and both *V. cholerae* and *A. baylyi* contain *radA*
14 homologs (VC2343 and ACIAD2664, respectively). RadA, however, is not required for natural
15 transformation in *V. cholerae* (**Fig. S7**). Also, preliminary Tn-seq data from our lab indicates that
16 RadA is not important for natural transformation in *A. baylyi*. Thus, it is tempting to speculate
17 that in Gram-positive competent species, RadA carries out the same function that ComM plays
18 during natural transformation in Gram-negative species.

19

20 *ComM is not required for DNA repair*

21 ComM homologs are also found among diverse non-competent species (**Fig. 5**). Moreover,
22 many bacterial helicases are implicated in promoting branch migration during other types of
23 homologous recombination, including during DNA repair. So next, we wanted to determine if
24 ComM also plays a role in DNA repair independent of its role during natural transformation.
25 ComM is poorly expressed in the absence of competence induction in *V. cholerae* (**Fig. S2**).
26 Thus, to test the role of ComM in DNA damage, we tested survival of P_{tac} -*tfoX* and P_{tac} -*tfoX*
27 $\Delta comM$ strains under competence inducing conditions (i.e. ectopic expression of TfoX). A WT
28 strain and a *recA* mutant were also included in these assays as controls. DNA damage was
29 tested using methyl methanesulfonate (MMS - methylates DNA / stalls replication forks) and

1 mitomycin C (MMC – alkylates DNA / generates interstrand DNA crosslinks)(37,38). As
2 expected, the *recA* mutant was more sensitive to these treatments compared to the WT,
3 consistent with a critical role for homologous recombination in DNA repair (**Fig. S8**) (39,40). The
4 P_{tac} -*tfoX* Δ *comM* mutant, however, was as resistant to these DNA damaging agents as the P_{tac} -
5 *tfoX* strain, indicating that this branch migration factor either does not play a role during DNA
6 repair or that the activity of this protein is redundant with other branch migration factors in the
7 context of repair (**Fig. S8**). Furthermore, induction of competence (via ectopic expression of
8 TfoX) showed little to no difference in DNA damage repair compared to the WT, indicating that
9 natural competence plays a limited role in DNA repair in *V. cholerae* (**Fig. S8**).

10

11 **DISCUSSION**

12 Natural transformation is an important mechanism of horizontal gene transfer in bacterial
13 species. It is dependent upon activation of bacterial competence or the ability to bind and take
14 up exogenous DNA. Altogether, our *in vivo* and *in vitro* data elucidate a role for ComM as a
15 branch migration factor that promotes the integration of tDNA during natural transformation.
16 In our model, integration is initiated by RecA-mediated strand invasion and formation of a D-
17 loop that generates a three-stranded intermediate structure (**Fig. 7**). If this D-loop is formed on
18 the 5' end of the invading tDNA, we propose that ComM loads onto the 5' end of the displaced
19 genomic strand. Through its 5' to 3' ATP-dependent helicase activity, ComM then proceeds to
20 displace the native strand, promoting branch migration of the D-loop and concomitantly
21 enhancing the integration of tDNA by exposing the complementary genomic sequence (**Fig. 7**).
22 Enhanced integration of tDNA during this process could be facilitated by the 5' to 3' polarity for
23 RecA-mediated strand exchange (41). Following this branch migration, the junctions are
24 resolved to mediate stable integration of tDNA by an unresolved mechanism. Inherent to this
25 model is a 5' end bias for the initial strand invasion of tDNA. Alternatively, ComM directionality
26 or loading may be biased by other as yet unidentified proteins that may facilitate 3' to 5'
27 helicase activity. Dissecting these questions will be the focus of future work.

28

1 Our data further suggests that ComM's role as a branch migration factor is important for the
2 incorporation of heterologous sequences. The main drivers for evolution and maintenance of
3 natural transformation in bacterial species are heavily debated. One model suggests that this
4 process is largely for enhancing genetic diversity, while another hypothesis is that natural
5 transformation evolved as a mechanism for acquisition of DNA as a nutrient (42,43). These
6 processes are not mutually exclusive. Because ComM affects only the integration of
7 heterologous tDNA (and not its uptake), however, the activity of this protein supports a role for
8 natural transformation in adaptation and evolution through the acquisition of novel genetic
9 material. Other competence genes that are involved specifically in homologous recombination
10 (e.g., *dprA*) also support this hypothesis.

11
12 Our *in vitro* data suggests that ComM forms a hexameric ring structure in its active state similar
13 to that of eukaryotic MCM2-7 and bacterial DnaB (30,44). Our 3D reconstruction reveals a
14 three-tiered barrel-like complex with a 26-Å pore and 18-Å pore on the top and bottom,
15 respectively. We propose that ComM either forms around, or is loaded onto the displaced
16 single strand of the native genomic DNA and acts as a wedge to dissociate these strands.
17 Alternatively, dsDNA may enter the 26-Å pore, be unwound in the central channel of the
18 complex with single strands being extruded through the side channels that are evident between
19 the tiers of the barrel structure. Such side channels are not uncommon among hexameric AAA+
20 helicases (e.g., SV40 T-antigen (45,46) and the archaeal MCM (27,47)) and have been
21 hypothesized to be exit channels for extruded ssDNA. The pore sizes observed in our
22 reconstruction appear to be consistent with those found in other ring helicases, however, it has
23 been shown that the size of the opening and channel can change depending on the nucleotide
24 bound (ATP vs. ADP) (45,48). Future work will focus on characterizing the ComM structure
25 bound to ADP and AMP-PNP, which may help inform the structural changes associated with the
26 catalytic cycle of this hexameric helicase.

27
28 Our phylogenetic analysis indicates that ComM is broadly conserved in bacterial species, and is
29 largely, only excluded from the Bacilli and Mollicutes. There is also, however, evidence for

1 isolated examples for ComM loss in species that fall outside of these two groups, which
2 suggests that these have occurred relatively recently. Some of these represent obligate
3 intracellular pathogens or endosymbionts, which commonly contain highly reduced genomes
4 (e.g. *Buchnera* and *Chlamydia*) (49,50). Another example of recent ComM loss is among
5 *Prochlorococcus* species, which are related to other naturally competent cyanobacteria (e.g.
6 *Synechococcus* and *Synechocystis spp.*). Interestingly, we found that *Prochlorococcus* lacked
7 many of the genes required for natural transformation (e.g. *dprA*, *comEA*, *comEC*, etc.) (51),
8 indicating that competence may have been lost in this lineage of cyanobacteria. Another
9 notable example of ComM loss is in *Acinetobacter baumannii*, which is an opportunistic
10 pathogen that is closely related to *A. baylyi*. Many strains of *A. baumannii*, including the one
11 analyzed here (ATCC 17978), contain an AbaR-type genomic island integrated into *comM*
12 (52,53). Strains that lack this horizontally transferred genomic island display low rates of natural
13 transformation (54), indicating that AbaR island-dependent inactivation of ComM may inhibit
14 this mechanism of HGT in *A. baumannii*. This represents another example in a growing list of
15 horizontally acquired genomic elements that inhibit natural transformation (55-57). Our
16 phylogenetic analysis also indicates that ComM is also highly conserved among non-competent
17 bacterial species. This suggests that ComM may have a function outside of natural
18 transformation. Alternatively, it is possible that many of these species are capable of natural
19 transformation, however, the conditions required for competence induction have not yet been
20 identified. Indeed, the inducing cue for natural transformation in *V. cholerae* was only
21 discovered in 2006 (2).

22
23 Our data suggest that branch migration in Gram-positive and Gram-negative species has
24 diverged in their dependence on distinct helicases (RadA and ComM, respectively). The tract
25 length of DNA recombined into *S. pneumoniae* during natural transformation is ~2.5 kb on
26 average (58). In *H. influenzae*, the mean recombination tract length is ~14 kb (59). Thus,
27 compared to RadA-dependent branch migration in Gram-positive species, ComM may facilitate
28 the integration of more tDNA in Gram-negative species. Other proteins that impact DNA

1 integration during natural transformation, however, may confound this overly simplified
2 comparison.

3
4 ComM is not essential for natural transformation because transformants are still observed in a
5 *comM* mutant. This suggests that other proteins may be involved in promoting the integration
6 of tDNA in the absence of this branch migration factor. Also, our data suggest that these
7 alternative branch migration factors are less efficient at incorporating tDNA with sequence
8 heterology compared to ComM in both *V. cholerae* and *A. baylyi*. This role could be carried out
9 by another helicase. Candidates include RuvAB, RecG, and PriA, which have all previously been
10 implicated in branch migration (31,32,60). In fact, the primosome helicase PriA is essential for
11 natural transformation in *Neisseria gonorrhoeae*. This may be due to a role in branch migration,
12 however, PriA is also essential for restarting stalled replication forks and *priA* mutants have a
13 severe growth defect (61). Also, it has been proposed that PriA helicase activity may facilitate
14 the uptake of tDNA through the inner membrane (62). A *recG* homologue in *Streptococcus*
15 *pneumoniae* (*mmsA*) was shown to have a mild effect on transformation when deleted (63).
16 RecG reverses stalled replication forks and promotes branch migration opposite to the direction
17 of RecA-mediated strand exchange in *E. coli* (64,65). Also, RecG acts on three-strand
18 intermediates *in vitro*, which could give credence to involvement in natural transformation
19 where such a structure is formed (66). Because RecG works counter to the direction of RecA-
20 mediated strand exchange, however, it is currently unclear how RecG may play a role in natural
21 transformation in the presence of ComM. Another possibility is that RecA, which has inherent
22 ATP-dependent unidirectional branch migration activity (67), could promote branch migration
23 independent of other canonical branch migration factors. Future work will focus on identifying
24 genetically interacting partners of ComM to uncover the role of additional factors required for
25 efficient integration of tDNA during natural transformation.

26
27 In addition to a AAA+ ATPase domain, ComM also contains two magnesium (Mg) chelatase
28 domains. Mg chelatase domain-containing proteins have only previously been implicated in
29 inserting Mg into protoporphyrin rings in photosynthetic organisms (68). While it is currently

1 unclear how these domains participate in ComM function, to our knowledge, this study is the
2 first to indicate a function for a Mg-chelatase domain-containing protein in a non-
3 photosynthetic organism.

4

5 In conclusion, the results from this study strongly support that ComM enhances natural
6 transformation by promoting ATP-dependent 5' to 3' helicase activity, which allows for the
7 efficient integration of heterologous tDNA. Furthermore, our data in *V. cholerae* and *A. baylyi*
8 as well as previous work from *H. influenzae* indicate that this helicase is a broadly conserved
9 mechanism for integration of tDNA in diverse naturally transformable Gram-negative species.

10

11 **ACKNOWLEDGEMENTS**

12 This work was supported by US National Institutes of Health Grant AI118863 to A.B.D., an
13 Indiana University Collaborative Research Grant to M.L.B., startup funds from the Indiana
14 University College of Arts and Sciences to ABD, and DTK was supported by NIH grant R35
15 GM122556 to Yves Brun.

16

17 **REFERENCES**

18

- 19 1. Seitz, P. and Blokesch, M. (2013) Cues and regulatory pathways involved in natural
20 competence and transformation in pathogenic and environmental Gram-negative
21 bacteria. *FEMS Microbiol Rev*, **37**, 336-363.
- 22 2. Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.Y. and Schoolnik, G.K. (2005) Chitin
23 induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824-1827.
- 24 3. Dalia, A.B., Lazinski, D.W. and Camilli, A. (2014) Identification of a membrane-bound
25 transcriptional regulator that links chitin and natural competence in *Vibrio*
26 *cholerae*. *MBio*, **5**, e01028-01013.
- 27 4. Yamamoto, S., Mitobe, J., Ishikawa, T., Wai, S.N., Ohnishi, M., Watanabe, H. and
28 Izumiya, H. (2014) Regulation of natural competence by the orphan two-component
29 system sensor kinase ChiS involves a non-canonical transmembrane regulator in
30 *Vibrio cholerae*. *Mol Microbiol*, **91**, 326-347.

- 1 5. Lo Scrudato, M. and Blokesch, M. (2012) The regulatory network of natural
2 competence and transformation of *Vibrio cholerae*. *PLoS Genet*, **8**, e1002778.
- 3 6. Lo Scrudato, M. and Blokesch, M. (2013) A transcriptional regulator linking quorum
4 sensing and chitin induction to render *Vibrio cholerae* naturally transformable.
5 *Nucleic Acids Res*, **41**, 3644-3658.
- 6 7. Metzgar, D., Bacher, J.M., Pezo, V., Reader, J., Doring, V., Schimmel, P., Marliere, P. and
7 de Crecy-Lagard, V. (2004) *Acinetobacter* sp. ADP1: an ideal model organism for
8 genetic analysis and genome engineering. *Nucleic Acids Res*, **32**, 5780-5790.
- 9 8. Lorenz, M.G. and Wackernagel, W. (1994) Bacterial gene transfer by natural genetic
10 transformation in the environment. *Microbiol Rev*, **58**, 563-602.
- 11 9. Gwinn, M.L., Ramanathan, R., Smith, H.O. and Tomb, J.F. (1998) A new
12 transformation-deficient mutant of *Haemophilus influenzae* Rd with normal DNA
13 uptake. *J Bacteriol*, **180**, 746-748.
- 14 10. Miller, V.L., DiRita, V.J. and Mekalanos, J.J. (1989) Identification of *toxS*, a regulatory
15 gene whose product enhances *toxR*-mediated activation of the cholera toxin
16 promoter. *J Bacteriol*, **171**, 1288-1293.
- 17 11. Juni, E. and Janik, A. (1969) Transformation of *Acinetobacter calco-aceticus*
18 (*Bacterium anitratum*). *J Bacteriol*, **98**, 281-288.
- 19 12. Dalia, A.B., Lazinski, D.W. and Camilli, A. (2013) Characterization of
20 undermethylated sites in *Vibrio cholerae*. *J Bacteriol*, **195**, 2389-2399.
- 21 13. Dalia, A.B., McDonough, E. and Camilli, A. (2014) Multiplex genome editing by
22 natural transformation. *Proc Natl Acad Sci U S A*, **111**, 8937-8942.
- 23 14. Dalia, A.B. (2016) *RpoS* is required for natural transformation of *Vibrio cholerae*
24 through regulation of chitinases. *Environ Microbiol*, **18**, 3758-3767.
- 25 15. Studier, F.W. (2005) Protein production by auto-induction in high density shaking
26 cultures. *Protein Expr Purif*, **41**, 207-234.
- 27 16. Wittig, I., Braun, H.P. and Schagger, H. (2006) Blue native PAGE. *Nat Protoc*, **1**, 418-
28 428.
- 29 17. Wu, M. and Eisen, J.A. (2008) A simple, fast, and accurate method of phylogenomic
30 inference. *Genome Biol*, **9**, R151.

- 1 18. Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and
2 high throughput. *Nucleic Acids Res*, **32**, 1792-1797.
- 3 19. Price, M.N., Dehal, P.S. and Arkin, A.P. (2010) FastTree 2--approximately maximum-
4 likelihood trees for large alignments. *PLoS One*, **5**, e9490.
- 5 20. Darling, A.E., Jospin, G., Lowe, E., Matsen, F.A.t., Bik, H.M. and Eisen, J.A. (2014)
6 PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, **2**, e243.
- 7 21. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-
8 analysis of large phylogenies. *Bioinformatics*, **30**, 1312-1313.
- 9 22. Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix.
10 *Molecular biology and evolution*, **25**, 1307-1320.
- 11 23. Mortier-Barriere, I., Velten, M., Dupaigne, P., Mirouze, N., Pietrement, O., McGovern,
12 S., Fichant, G., Martin, B., Noirot, P., Le Cam, E. *et al.* (2007) A key presynaptic role in
13 transformation for a widespread bacterial protein: DprA conveys incoming ssDNA
14 to RecA. *Cell*, **130**, 824-836.
- 15 24. Suckow, G., Seitz, P. and Blokesch, M. (2011) Quorum sensing contributes to natural
16 transformation of *Vibrio cholerae* in a species-specific manner. *J Bacteriol*, **193**,
17 4914-4924.
- 18 25. Hanson, P.I. and Whiteheart, S.W. (2005) AAA+ proteins: have engine, will work. *Nat*
19 *Rev Mol Cell Biol*, **6**, 519-529.
- 20 26. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The
21 Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, **10**,
22 845-858.
- 23 27. Brewster, A.S., Wang, G., Yu, X., Greenleaf, W.B., Carazo, J.M., Tjajadi, M., Klein, M.G.
24 and Chen, X.S. (2008) Crystal structure of a near-full-length archaeal MCM:
25 functional insights for an AAA+ hexameric helicase. *Proc Natl Acad Sci U S A*, **105**,
26 20191-20196.
- 27 28. Kaplan, D.L., Davey, M.J. and O'Donnell, M. (2003) Mcm4,6,7 uses a "pump in ring"
28 mechanism to unwind DNA by steric exclusion and actively translocate along a
29 duplex. *J Biol Chem*, **278**, 49171-49182.

- 1 29. Borgeaud, S., Metzger, L.C., Scignari, T. and Blokesch, M. (2015) The type VI
2 secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science*, **347**,
3 63-67.
- 4 30. Bochman, M.L. and Schwacha, A. (2009) The Mcm complex: unwinding the
5 mechanism of a replicative helicase. *Microbiol Mol Biol Rev*, **73**, 652-683.
- 6 31. Lloyd, R.G. and Sharples, G.J. (1993) Processing of recombination intermediates by
7 the RecG and RuvAB proteins of *Escherichia coli*. *Nucleic Acids Res*, **21**, 1719-1725.
- 8 32. West, S.C. (1997) Processing of recombination intermediates by the RuvABC
9 proteins. *Annu Rev Genet*, **31**, 213-244.
- 10 33. Burghout, P., Bootsma, H.J., Kloosterman, T.G., Bijlsma, J.J., de Jongh, C.E., Kuipers,
11 O.P. and Hermans, P.W. (2007) Search for genes essential for pneumococcal
12 transformation: the RADA DNA repair protein plays a role in genomic
13 recombination of donor DNA. *J Bacteriol*, **189**, 6540-6550.
- 14 34. Carrasco, B., Fernandez, S., Asai, K., Ogasawara, N. and Alonso, J.C. (2002) Effect of
15 the recU suppressors sms and subA on DNA repair and homologous recombination
16 in *Bacillus subtilis*. *Mol Genet Genomics*, **266**, 899-906.
- 17 35. Cooper, D.L. and Lovett, S.T. (2016) Recombinational branch migration by the
18 RadA/Sms paralog of RecA in *Escherichia coli*. *eLife*, **5**.
- 19 36. Marie, L., Rapisarda, C., Morales, V., Berge, M., Perry, T., Soulet, A.L., Gruget, C.,
20 Remaut, H., Fronzes, R. and Polard, P. (2017) Bacterial RadA is a DnaB-type helicase
21 interacting with RecA to promote bidirectional D-loop extension. *Nat Commun*, **8**,
22 15638.
- 23 37. Weng, M.W., Zheng, Y., Jasti, V.P., Champeil, E., Tomasz, M., Wang, Y., Basu, A.K. and
24 Tang, M.S. (2010) Repair of mitomycin C mono- and interstrand cross-linked DNA
25 adducts by UvrABC: a new model. *Nucleic Acids Res*, **38**, 6976-6984.
- 26 38. Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A.S. and Helleday,
27 T. (2005) Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no
28 detectable in vivo DNA double-strand breaks. *Nucleic Acids Res*, **33**, 3799-3811.
- 29 39. Goranov, A.I., Kuester-Schoeck, E., Wang, J.D. and Grossman, A.D. (2006)
30 Characterization of the global transcriptional responses to different types of DNA

- 1 damage and disruption of replication in *Bacillus subtilis*. *J Bacteriol*, **188**, 5595-
2 5605.
- 3 40. Better, M. and Helinski, D.R. (1983) Isolation and characterization of the *recA* gene
4 of *Rhizobium meliloti*. *J Bacteriol*, **155**, 311-316.
- 5 41. Cox, M.M. and Lehman, I.R. (1981) Directionality and polarity in *recA* protein-
6 promoted branch migration. *Proc Natl Acad Sci U S A*, **78**, 6018-6022.
- 7 42. Michod, R.E., Wojciechowski, M.F. and Hoelzer, M.A. (1988) DNA repair and the
8 evolution of transformation in the bacterium *Bacillus subtilis*. *Genetics*, **118**, 31-39.
- 9 43. Redfield, R.J. (1993) Genes for breakfast: the have-your-cake-and-eat-it-too of
10 bacterial transformation. *J Hered*, **84**, 400-404.
- 11 44. San Martin, M.C., Stamford, N.P., Dammerova, N., Dixon, N.E. and Carazo, J.M. (1995)
12 A structural model for the *Escherichia coli* DnaB helicase based on electron
13 microscopy data. *J Struct Biol*, **114**, 167-176.
- 14 45. Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X.S. (2004) Mechanisms of
15 conformational change for a replicative hexameric helicase of SV40 large tumor
16 antigen. *Cell*, **119**, 47-60.
- 17 46. Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E.,
18 Jochimiak, A., Szakonyi, G. and Chen, X.S. (2003) Structure of the replicative helicase
19 of the oncoprotein SV40 large tumour antigen. *Nature*, **423**, 512-518.
- 20 47. Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006)
21 Structural basis of the *Methanothermobacter thermautotrophicus* MCM helicase
22 activity. *Nucleic Acids Res*, **34**, 5829-5838.
- 23 48. Strycharska, M.S., Arias-Palomo, E., Lyubimov, A.Y., Erzberger, J.P., O'Shea, V.L.,
24 Bustamante, C.J. and Berger, J.M. (2013) Nucleotide and partner-protein control of
25 bacterial replicative helicase structure and function. *Mol Cell*, **52**, 844-854.
- 26 49. Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. (2000) Genome
27 sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature*,
28 **407**, 81-86.
- 29 50. Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W.,
30 Olinger, L., Tatusov, R.L., Zhao, Q. *et al.* (1998) Genome sequence of an obligate
31 intracellular pathogen of humans: *Chlamydia trachomatis*. *Science*, **282**, 754-759.

- 1 51. Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., Arellano, A.,
2 Coleman, M., Hauser, L., Hess, W.R. *et al.* (2003) Genome divergence in two
3 *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature*, **424**, 1042-
4 1047.
- 5 52. Smith, M.G., Gianoulis, T.A., Pukatzki, S., Mekalanos, J.J., Ornston, L.N., Gerstein, M.
6 and Snyder, M. (2007) New insights into *Acinetobacter baumannii* pathogenesis
7 revealed by high-density pyrosequencing and transposon mutagenesis. *Genes Dev*,
8 **21**, 601-614.
- 9 53. Ramirez, M.S., Vilacoba, E., Stietz, M.S., Merkier, A.K., Jeric, P., Limansky, A.S.,
10 Marquez, C., Bello, H., Catalano, M. and Centron, D. (2013) Spreading of AbaR-type
11 genomic islands in multidrug resistance *Acinetobacter baumannii* strains belonging
12 to different clonal complexes. *Curr Microbiol*, **67**, 9-14.
- 13 54. Traglia, G.M., Chua, K., Centron, D., Tolmasky, M.E. and Ramirez, M.S. (2014) Whole-
14 genome sequence analysis of the naturally competent *Acinetobacter baumannii*
15 clinical isolate A118. *Genome Biol Evol*, **6**, 2235-2239.
- 16 55. Dalia, A.B., Seed, K.D., Calderwood, S.B. and Camilli, A. (2015) A globally distributed
17 mobile genetic element inhibits natural transformation of *Vibrio cholerae*. *Proc Natl*
18 *Acad Sci U S A*, **112**, 10485-10490.
- 19 56. Konkol, M.A., Blair, K.M. and Kearns, D.B. (2013) Plasmid-encoded ComI inhibits
20 competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol*, **195**, 4085-
21 4093.
- 22 57. Rabinovich, L., Sigal, N., Borovok, I., Nir-Paz, R. and Herskovits, A.A. (2012) Prophage
23 excision activates *Listeria* competence genes that promote phagosomal escape and
24 virulence. *Cell*, **150**, 792-802.
- 25 58. Croucher, N.J., Harris, S.R., Barquist, L., Parkhill, J. and Bentley, S.D. (2012) A high-
26 resolution view of genome-wide pneumococcal transformation. *PLoS Pathog*, **8**,
27 e1002745.
- 28 59. Mell, J.C., Shumilina, S., Hall, I.M. and Redfield, R.J. (2011) Transformation of natural
29 genetic variation into *Haemophilus influenzae* genomes. *PLoS Pathog*, **7**, e1002151.

- 1 60. Al-Deib, A.A., Mahdi, A.A. and Lloyd, R.G. (1996) Modulation of recombination and
2 DNA repair by the RecG and PriA helicases of Escherichia coli K-12. *J Bacteriol*, **178**,
3 6782-6789.
- 4 61. Kline, K.A. and Seifert, H.S. (2005) Mutation of the priA gene of Neisseria
5 gonorrhoeae affects DNA transformation and DNA repair. *J Bacteriol*, **187**, 5347-
6 5355.
- 7 62. Kruger, N.J. and Stingl, K. (2011) Two steps away from novelty--principles of
8 bacterial DNA uptake. *Mol Microbiol*, **80**, 860-867.
- 9 63. Martin, B., Sharples, G.J., Humbert, O., Lloyd, R.G. and Claverys, J.P. (1996) The mmsA
10 locus of Streptococcus pneumoniae encodes a RecG-like protein involved in DNA
11 repair and in three-strand recombination. *Mol Microbiol*, **19**, 1035-1045.
- 12 64. Whitby, M.C., Ryder, L. and Lloyd, R.G. (1993) Reverse branch migration of Holliday
13 junctions by RecG protein: a new mechanism for resolution of intermediates in
14 recombination and DNA repair. *Cell*, **75**, 341-350.
- 15 65. Azeroglu, B., Mawer, J.S., Cockram, C.A., White, M.A., Hasan, A.M., Filatenkova, M. and
16 Leach, D.R. (2016) RecG Directs DNA Synthesis during Double-Strand Break Repair.
17 *PLoS Genet*, **12**, e1005799.
- 18 66. Whitby, M.C. and Lloyd, R.G. (1995) Branch migration of three-strand recombination
19 intermediates by RecG, a possible pathway for securing exchanges initiated by 3'-
20 tailed duplex DNA. *EMBO J*, **14**, 3302-3310.
- 21 67. Cox, M.M., Morrical, S.W. and Neuendorf, S.K. (1984) Unidirectional branch
22 migration promoted by nucleoprotein filaments of RecA protein and DNA. *Cold*
23 *Spring Harb Symp Quant Biol*, **49**, 525-533.
- 24 68. Papenbrock, J., Mock, H.P., Tanaka, R., Kruse, E. and Grimm, B. (2000) Role of
25 magnesium chelatase activity in the early steps of the tetrapyrrole biosynthetic
26 pathway. *Plant Physiol*, **122**, 1161-1169.

27

28 **FIGURE LEGENDS**

29 **Fig. 1** – *ComM* is required for integration of DNA during natural transformation. (A) Chitin-
30 dependent natural transformation assays in the indicated strains using a linear PCR product or a

1 replicating plasmid as tDNA. **(B)** Chitin-independent transformation assays of the indicated
2 strains with linear PCR product as tDNA. **(C)** Complementation of *comM* *in trans* tested in chitin-
3 dependent transformation assays using a linear PCR product as tDNA. All data are shown as the
4 mean \pm SD and are the result of at least three independent biological replicates. ** = $p < 0.01$,
5 *** = $p < 0.001$, NS = not significant

6
7 **Fig. 2 – *ComM* promotes branch migration through heterologous sequences in vivo.** **(A)** Chitin-
8 dependent transformation assay performed using tDNA that contained linked genetic markers
9 separated by 820 bp or 245 bp. **(B)** Chitin-dependent transformation assay performed in Tm^S
10 strains (Tm^R marker inactivated by a nonsense point mutation or 29-bp deletion) using tDNA
11 that would revert the integrated marker to Tm^R . All strains in **A** and **B** contained a mutation in
12 *mutS* to prevent MMR activity. In schematics above bar graphs, X's denote possible crossover
13 points for homologous recombination. All data are shown as the mean \pm SD and are the result
14 of at least three independent biological replicates. *** = $p < 0.001$

15
16 **Fig. 3 – *ComM* hexamerizes in the presence of ATP and ssDNA.** **(A)** Blue-native PAGE assay of
17 purified ComM in the indicated conditions. **(B)** Negative stain EM of purified ComM under the
18 indicated conditions. Representative ring-like densities observed in the presence of 5 mM ATP
19 and 5 μ M ssDNA are indicated by black arrows. Scale bar = 50 nm. **(C)** A representative 2-D class
20 average of the ring-like densities observed by negative stain EM reveals a hexameric complex.
21 **(D)** 3-D reconstruction (~ 13.8 Å resolution) of the ring complex imposing C6 symmetry.

22
23 **Fig. 4 – *ComM* exhibits helicase activity in vitro.** **(A)** A representative forked substrate helicase
24 assay with increasing concentrations of purified ComM. Concentrations of ComM used (in
25 hexamer) were 0, 10, 25, 37.5, 50, 75, 100, 150, 200, and 250 nM. Images were quantified and
26 plotted as indicated. **(B)** Helicase assay using forked DNA substrate with the indicated purified
27 protein (100 nM Pif1 and 250 nM ComM / ComM^{K224A} hexamer) in the presence of 5 mM ATP
28 (Columns 1, 2, and 3) or ATP γ S (Column 4). **(C)** Helicase assay using the indicated DNA substrate

1 with 250 nM purified ComM hexamer. All data are shown as the mean \pm SD and are the result
2 of at least three independent replicates. *** = $p < 0.001$.

3

4 **Fig. 5 – *ComM* is broadly conserved.** Estimated maximum likelihood phylogeny of diverse
5 species based on a concatenated alignment of 36 conserved proteins identified from whole
6 genome sequences. Species with an identified ComM homolog are highlighted in gray.
7 Competent species are designated by a star next to the species name. Major taxa are labeled
8 along their nodes. Pro: Proteobacteria (Greek letters indicate subdivisions); Bac: Bacilli; Mol:
9 Mollicutes; Cya: Cyanobacteria; Arc: Archaea. Scale bar indicates distance.

10

11

12 **Fig. 6 – *ComM* promotes branch migration in *Acinetobacter baylyi*.** (A) Transformation assay of
13 *A. baylyi* using a linear tDNA product with linked genetic markers. (B) Transformation assay
14 performed in Spec^S strains with tDNA that would revert the strain to Spec^R. Integration of the
15 marker would either repair a point mutation or delete 180 bp of genomic sequence. All strains
16 in A and B contain *mutS* mutations to prevent MMR activity. In schematics above bar graphs,
17 X's denote possible crossover points for homologous recombination. All data are shown as the
18 mean \pm SD and are the result of at least three independent biological replicates. *** = $p < 0.001$.

19

20 **Fig. 7 – Proposed model for the role of *ComM* during natural transformation.** ComM is shown as
21 a hexameric ring that promotes branch migration through its 5' to 3' helicase activity. This can
22 support integration of tDNA with a heterologous region, which is indicated by a gray box.

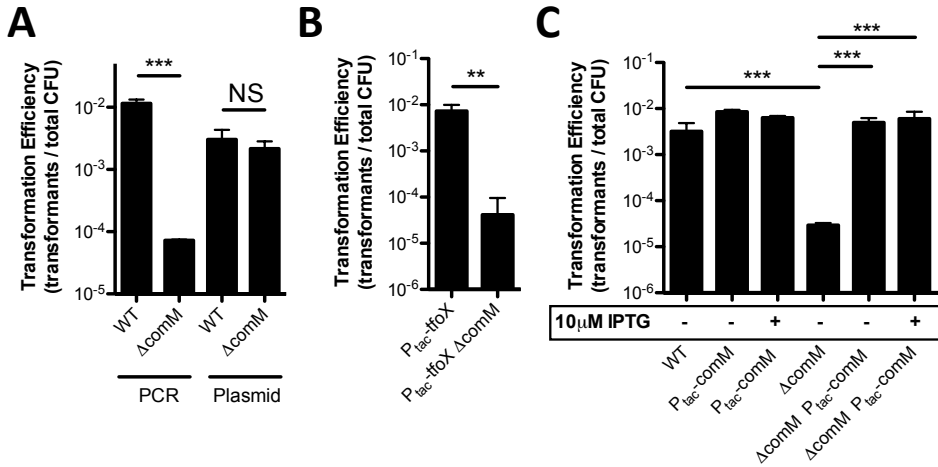


Fig. 1 – *ComM* is required for integration of DNA during natural transformation. (A) Chitin-dependent natural transformation assays in the indicated strains using a linear PCR product or a replicating plasmid as tDNA. **(B)** Chitin-independent transformation assays of the indicated strains with linear PCR product as tDNA. **(C)** Complementation of *comM* in trans tested in chitin-dependent transformation assays using a linear PCR product as tDNA. All data are shown as the mean \pm SD and are the result of at least three independent biological replicates. ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant

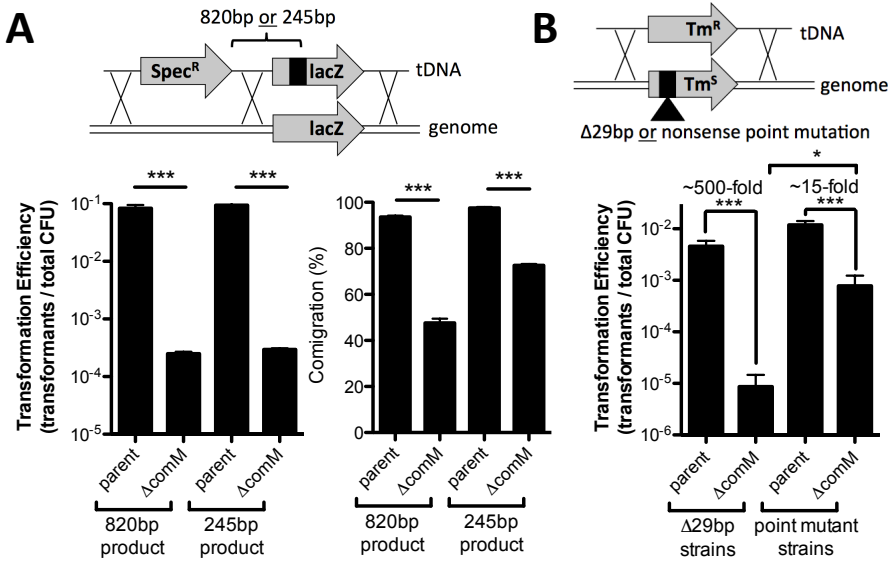


Fig. 2 – *ComM* promotes branch migration through heterologous sequences in vivo. (A) Chitin-dependent transformation assay performed using tDNA that contained linked genetic markers separated by 820 bp or 245 bp. **(B)** Chitin-dependent transformation assay performed in Tm^S strains (Tm^R marker inactivated by a nonsense point mutation or 29-bp deletion) using tDNA that would revert the integrated marker to Tm^R . All strains in **A** and **B** contained a mutation in *mutS* to prevent MMR activity. In schematics above bar graphs, X's denote possible crossover points for homologous recombination. All data are shown as the mean \pm SD and are the result of at least three independent biological replicates. *** = $p < 0.001$

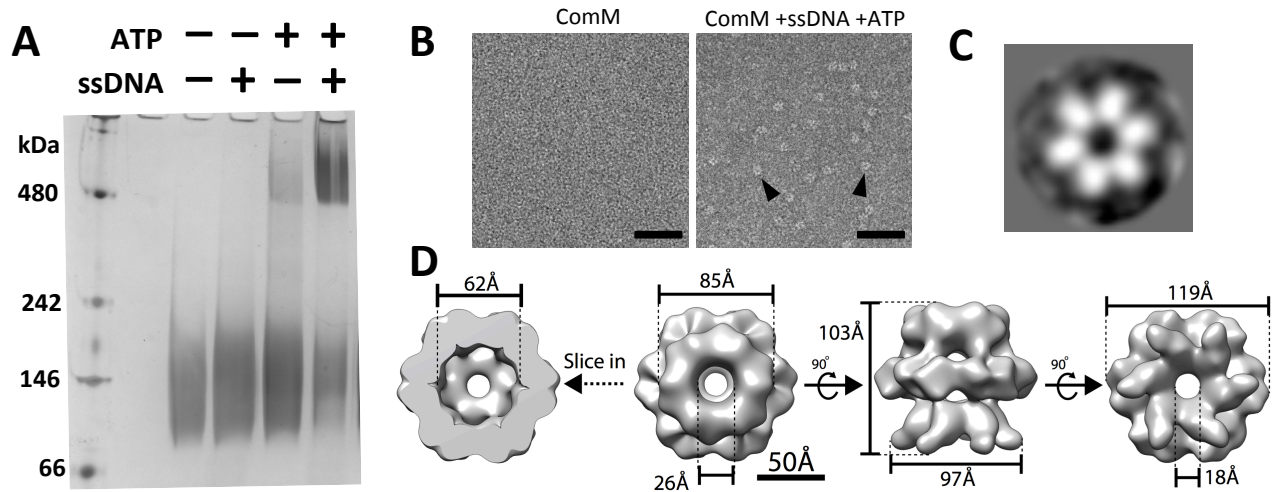


Fig. 3 – ComM hexamerizes in the presence of ATP and ssDNA. (A) Blue-native PAGE assay of purified ComM in the indicated conditions. (B) Negative stain EM of purified ComM under the indicated conditions. Representative ring-like densities observed in the presence of 5 mM ATP and 5 μ M ssDNA are indicated by black arrows. Scale bar = 50 nm. (C) A representative 2-D class average of the ring-like densities observed by negative stain EM reveals a hexameric complex. (D) 3-D reconstruction (\sim 13.8 Å resolution) of the ring complex imposing C6 symmetry.

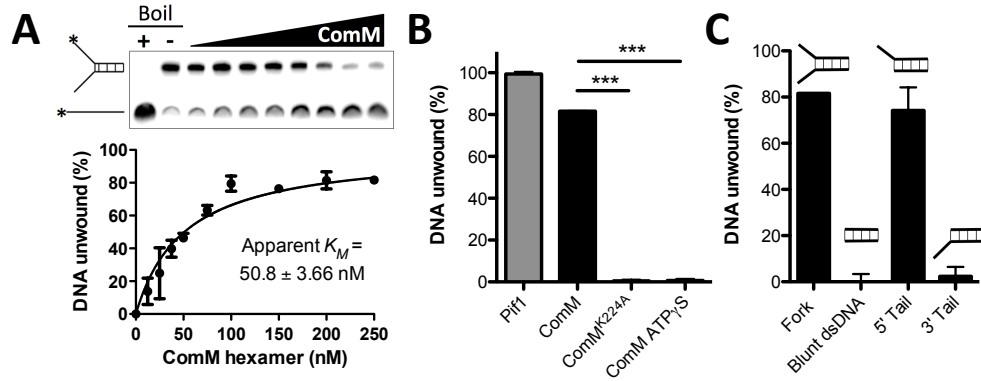


Fig. 4 – *ComM* exhibits helicase activity in vitro. (A) A representative forked substrate helicase assay with increasing concentrations of purified *ComM*. Concentrations of *ComM* used (in hexamer) were 0, 10, 25, 37.5, 50, 75, 100, 150, 200, and 250 nM. Images were quantified and plotted as indicated. (B) Helicase assay using forked DNA substrate with the indicated purified protein (100 nM Pif1 and 250 nM *ComM* / *ComM*^{K224A} hexamer) in the presence of 5 mM ATP (Columns 1, 2, and 3) or ATP γ S (Column 4). (C) Helicase assay using the indicated DNA substrate with 250 nM purified *ComM* hexamer. All data are shown as the mean \pm SD and are the result of at least three independent replicates. *** = $p < 0.001$.

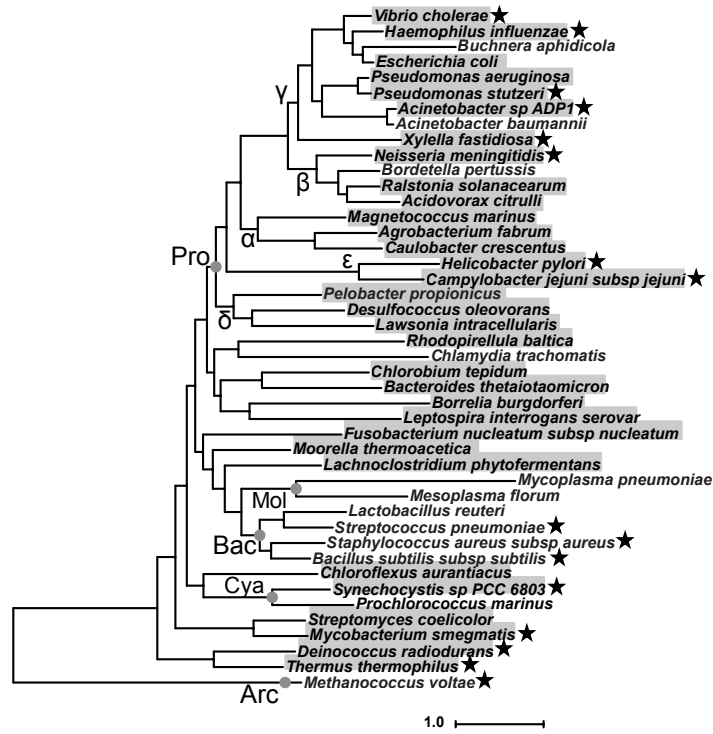


Fig. 5 – *ComM* is broadly conserved. Estimated maximum likelihood phylogeny of diverse species based on a concatenated alignment of 36 conserved proteins identified from whole genome sequences. Species with an identified ComM homolog are highlighted in gray. Competent species are designated by a star next to the species name. Major taxa are labeled along their nodes. Pro: Proteobacteria (Greek letters indicate subdivisions); Bac: Bacilli; Mol: Mollicutes; Cya: Cyanobacteria; Arc: Archaea. Scale bar indicates distance.

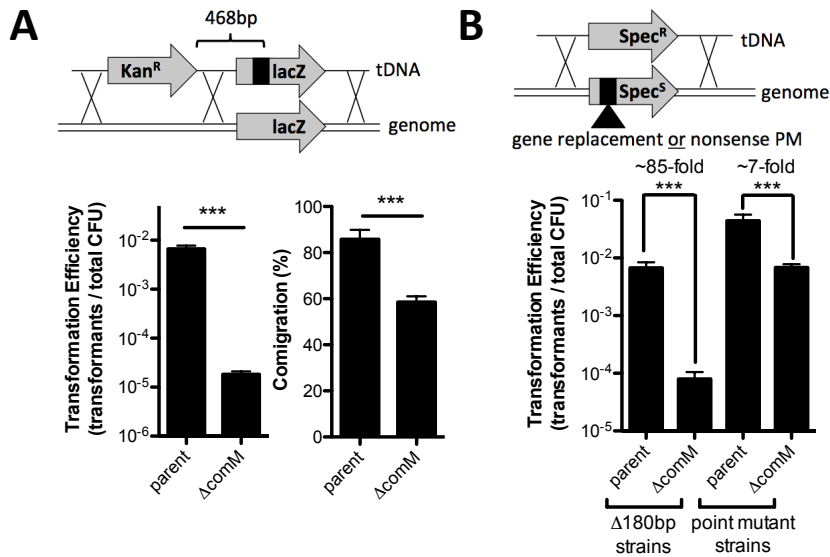


Fig. 6 – *ComM* promotes branch migration in *Acinetobacter baylyi*. (A) Transformation assay of *A. baylyi* using a linear tDNA product with linked genetic markers. (B) Transformation assay performed in Spec^S strains with tDNA that would revert the strain to Spec^R. Integration of the marker would either repair a point mutation or delete 180 bp of genomic sequence. All strains in A and B contain *mutS* mutations to prevent MMR activity. In schematics above bar graphs, X's denote possible crossover points for homologous recombination. All data are shown as the mean ± SD and are the result of at least three independent biological replicates. *** = $p < 0.001$.

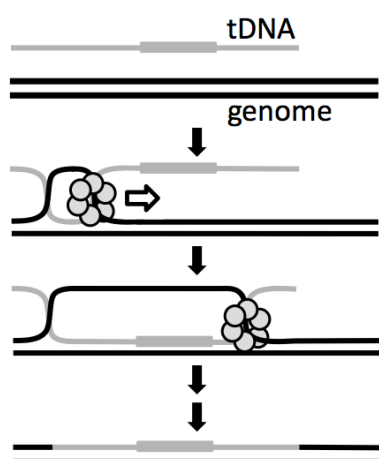


Fig. 7 – Proposed model for the role of *ComM* during natural transformation. *ComM* is shown as a hexameric ring that promotes branch migration through its 5' to 3' helicase activity. This can support integration of tDNA with a heterologous region, which is indicated by a gray box.

SUPPLEMENTARY FIGURES

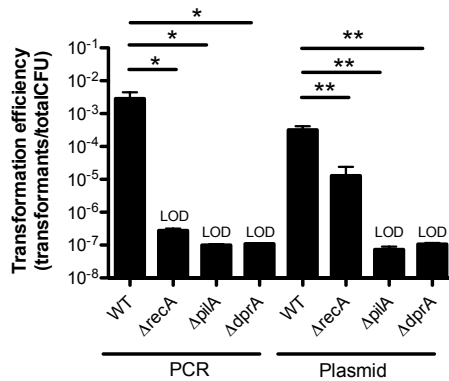


Fig. S1 – Uptake of plasmid DNA is independent of recombination. Chitin-dependent transformation assay with the indicated strains using either linear PCR product or plasmid as tDNA. All data are shown as the mean \pm SD and are the result of at least three independent biological replicates. * = $p < 0.05$, ** = $p < 0.01$, and LOD = limit of detection.

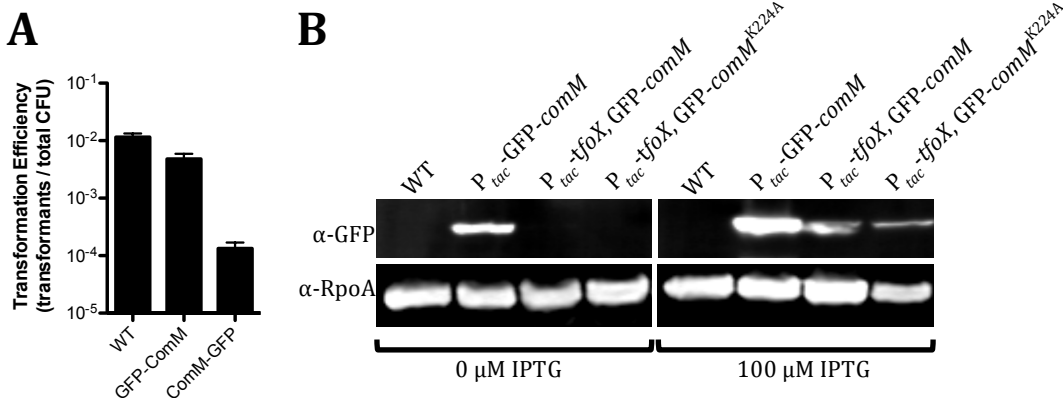


Fig. S2 – N-terminal ComM fusions are functional. (A) Chitin-dependent transformation assay with the indicated strains using linear PCR product as the tDNA. (B) Representative western blot to detect GFP-ComM and RpoA (loading control) in the indicated strains grown in the presence or absence of 100 μ M IPTG. Blot indicates that GFP-*comM* and GFP-*comM*^{K224A} at the native locus are induced when TfoX is ectopically expressed to induce competence. Also, this blot indicates that the P_{tac}-GFP-*comM* construct is leaky and expressed in the absence of inducer. Data in A are shown as the mean \pm SD and are the result of at least three independent biological replicates.

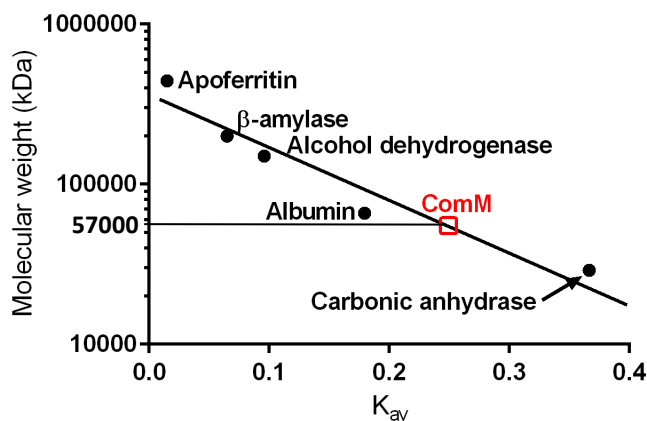


Fig. S3 – *ComM* is monomeric in soluble form. Purified StrepII-*ComM* was analyzed by gel filtration and compared to a set of protein standards to determine size.

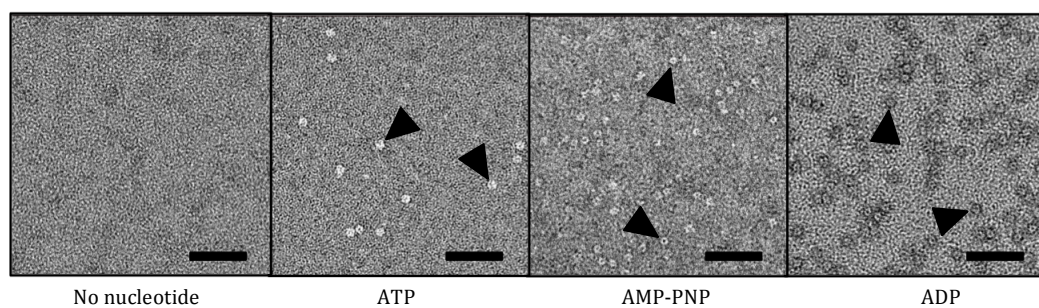


Fig. S4 – *ComM* oligomerizes in the presence of ADP and AMP-PNP. Negative stain EM of purified *ComM* incubated with 5 μM ssDNA and 5 mM of the indicated nucleotide or nucleotide analog. Representative ring-like densities in these samples are indicated by black arrows. Scale bar = 50 nm.

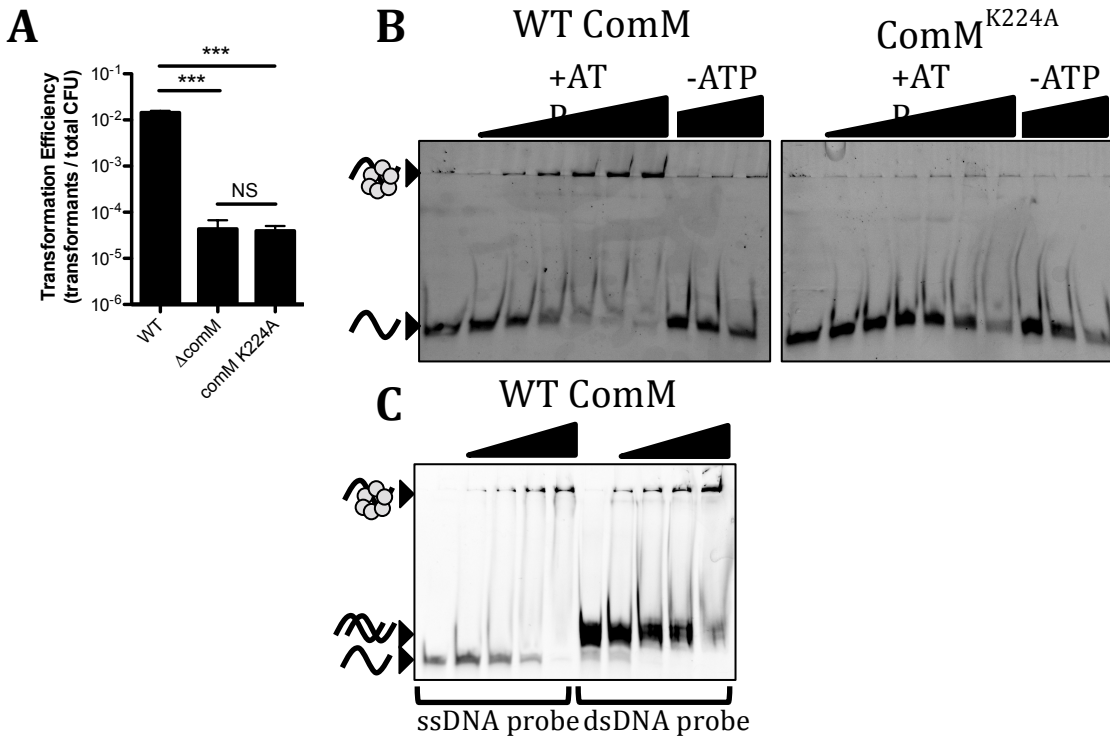


Figure S5 – *ComM* binds ssDNA in the presence of ATP. (A) Chitin-dependent transformation assay of the indicated strains using a linear PCR product as tDNA. Data are shown as the mean \pm SD and are the result of at least three independent biological replicates. *** = $p < 0.001$, NS = not significant. (B) EMSA with purified ComM and ComM^{K224A} and a ssDNA probe. Protein concentrations (of the hexamer) used in the presence of ATP (+ATP) were 0, 5.2, 10.4, 20.8, 41.6, 83.3, 166.6 nM, and in the absence of ATP (-ATP) were 41.6, 83.3, and 166.6 nM. Bound probe is retained in the well due to the large size of the DNA-bound oligomeric complex. (C) EMSA with purified ComM and ssDNA or dsDNA probe. All reactions were performed in the presence of ATP and the protein concentrations (of hexamer) used were 0, 10.4, 20.8, 41.6, and 83.3 nM from left to right. The size of unbound ssDNA and dsDNA are indicated on the right, while bound DNA is shifted to the well. Data in B and C are representative of at least two independent experiments.



Fig. S6 – *ComM* is broadly conserved. (A) Phylogenetic trees of species based on a concatenated alignment of 36 conserved protein sequences. Green text indicates species with an identifiable *ComM* homolog. (B) Phylogenetic tree of *ComM* alleles.

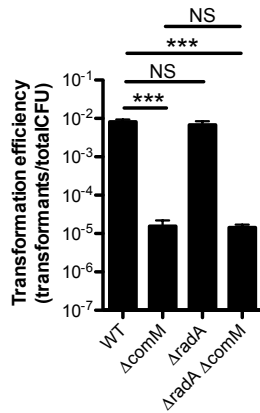


Fig. S7 – *RadA* is not required for natural transformation in *V. cholerae*. All strains contain P_{tac} -*tfoX* mutations and were transformed via chitin-independent transformation assays using a linear PCR product as the tDNA. All data are shown as the mean \pm SD and the result of 6 independent biological replicates. *** = $p < 0.001$, NS = not significant.

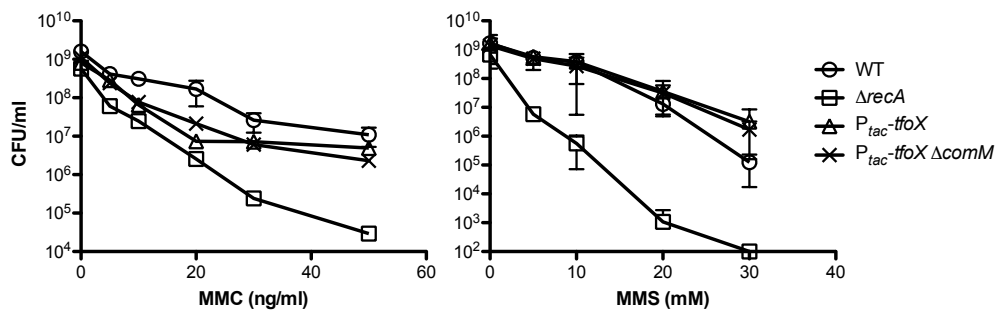


Fig. S8 – *ComM* is not required for DNA repair. Strains were treated with increasing doses of the DNA damaging agent indicated on the X axis and then plated for viability. All data are shown as the mean \pm SD and are the result of at least three independent biological replicates.

SUPPLEMENTARY TABLES

Table S1 – Strains used in this study

Strain Name in Manuscript	Genotype and Antibiotic Resistances	Description	Reference / strain#
WT	Sm ^R	Wildtype <i>V. cholerae</i> O1 El Tor strain used throughout this study	(1) / (SAD030)
Strains Used for Transformation Assays			
$\Delta comM$	$\Delta VC0032::Spec^R$	Replacement of VC0032 with a spectinomycin resistance cassette	This Study (SAD083)
$P_{tac-tfoX}$	VC1153 OE Kan ^R	OE Kan ^R = a fragment of the Tn10 transposon from pDL1093, including the <i>rrnB</i> antiterminator, <i>P_{tac}</i> , <i>LacI</i> , and Kan ^R . This fragment, and its use in overexpression of VC1153 (<i>tfoX</i>) is described in	(2) / (SAD061)
$\Delta comM P_{tac-tfoX}$	VC0032::Spec ^R , VC1153 OE Kan ^R	Replacement of VC0032 with Spec ^R cassette in a VC1153 OE Kan ^R parent strain	This Study (SAD066)
$P_{tac-comM}$	$P_{tac-comM}$ at <i>lacZ</i> locus Spec ^R	VC0032 fused with an IPTG inducible promoter at the <i>lacZ</i> locus	This Study (SAD1065)
$\Delta recA$	$\Delta recA::Spec^R$	Replacement of VC0543 with Spec ^R cassette	This Study (SAD081)
$\Delta pilA$	$\Delta pilA::Spec^R$	Replacement of VC2423 with Spec ^R cassette	This Study (SAD780)
$\Delta dprA$	$\Delta dprA::Spec^R$	Replacement of VC0048 with Spec ^R	This Study (SAD079)
$P_{tac-comM}, \Delta comM$	$P_{tac-comM}$ at <i>lacZ</i> locus Spec ^R , $\Delta VC0032::Amp^R$	VC0032 fused with an IPTG inducible promoter at the <i>lacZ</i> locus in a $\Delta comM$ parent strain	This Study (SAD1066)
$\Delta 29bp$ (Tm ^S)	$\Delta VC1807::TmR^*\Delta 29bp$, $\Delta mutS$ MuGENT edit, LPQEN Kan ^R	$\Delta VC1807::TmR^*\Delta 29bp$, $\Delta mutS$ MuGENT edit, LPQEN Kan ^R ; NOT TmR resistant	This Study (TND0226/SAD1321)
$\Delta 29bp$ (Tm ^S) $\Delta comM$	$\Delta comM::CarbR$, ($\Delta VC1807::TmR^*\Delta 29bp$, $\Delta mutS$ MuGENT edit, LPQEN Kan ^R)	VC0032 deletion in a Tm ^S parent strain	This Study (TND0229/SAD1322)
Point mutant (Tm ^S)	$\Delta mutS$ MuGENT edit, LPQEN::Spec ^R , $\Delta VC1807::TmR^*TI$	$\Delta mutS$ MuGENT edit, LPQEN::Spec ^R , $\Delta VC1807::TmR^*TI$; $\Delta VC1807::TmR^*TI$ is the TmR cassette at VC1807, except it has a transition point mutation that introduces a premature stop codon.	This Study (TND0220/SAD1323)
Point mutant (Tm ^S), $\Delta comM$	$\Delta mutS$ MUGENT edit, $\Delta comM::CarbR$, LPQEN::Kan ^R , $\Delta VC1807::TmR^*TI$	$\Delta mutS$ MUGENT edit, $\Delta comM::CarbR$, LPQEN::Kan ^R , $\Delta VC1807::TmR^*TI$; $\Delta VC1807::TmR^*TI$ is the TmR cassette at VC1807, except it has a transition point mutation that introduces a premature stop codon.	This Study (TND0221/SAD1324)
$comM^{K224A}$	LPQEN::Kan ^R , $comM^{K224A}$	K to A residue substitution disrupts ATP binding. FRT Kan cassette following the LPQEN amino acid sequence in <i>lacZ</i> (VC2338) used for	This Study (SAD1026)

		selection during co-transformation.	
<i>gfp-comM</i>	<i>gfp-comM</i>	N-terminal GFP-comM at native locus	This Study (SAD924)
<i>comM-gfp</i>	<i>comM-gfp</i>	C-terminal comM-GFP at native locus	This Study (SAD925)
parent (ADP1)	Δ ACIAD1551::P _{tac} -lacZ, Δ mutS::Spec ^R	lacZ introduced into a defunct transposase (neutral gene = ACIAD1551), mutS deleted and replaced with Spec ^R (ACIAD1500)	This Study (TND0137/SAD1325)
Δ comM (ADP1)	Δ ACIAD1551::P _{tac} -lacZ, Δ mutS::Spec ^R , Δ comM	lacZ introduced into a defunct transposase (neutral gene = ACIAD1551), mutS deleted and replaced with Spec ^R (ACIAD1500), comM in-frame mutation (ACIAD0242)	This Study (TND0149/SAD1326)
Δ 180 (ADP1)	Δ mutS::Kan ^R	MutS deleted and replaced with a Kan ^R cassette	This Study (SAD742)
Δ 180 Δ comM (ADP1)	Δ mutS::Kan ^R , Δ comM	MutS deleted and replaced with a Kan ^R cassette, Δ comM in-frame	This Study (TND0144/SAD1327)
Point mutant (ADP1)	ACIAD1551::Spec ^R (point mutant), Δ mutS::Kan ^R	SpecR cassette in ACIAD1551 is inactivated with a point mutation. MutS deleted and replaced with Kan ^R	This Study (TND0150/SAD1328)
Point mutant, Δ comM (ADP1)	ACIAD1551::Spec ^R (point mutant), Δ mutS::Kan ^R , Δ comM	SpecR cassette in ACIAD1551 is inactivated with a point mutation. mutS was deleted and replaced with Kan ^R and comM was deleted in-frame	This Study (TND0164/SAD1329)
pBAD18 Kan	pBAD18 Kan	TG1 E. coli strain used to purify a replicating plasmid with a Kan ^R cassette	This Study (SAD233)
P _{tac} -tfoX, Δ radA	VC1153 OE Kan ^R , <i>radA</i> ::Spec ^R	<i>radA</i> (VC2343) deleted and replaced with Spec ^R in a TfoX overexpressing background.	This Study (TMN0135/SAD1813)
P _{tac} -tfoX, Δ radA, Δ comM	VC1153 OE Kan ^R , <i>radA</i> ::Spec ^R , <i>comM</i> ::Carb ^R	<i>comM</i> (VC0032) deleted and replaced with Carb ^R in a TfoX overexpressing, <i>radA</i> deletion background.	This Study (TMN0136/SAD1814)
Strains used in ComM Induction			
P _{tac} -gfp	P _{tac} -GFPmut3 Spec ^R at the <i>lacZ</i> locus	Replaced lacZ gene in SAD030 with a fragment from the transposon vector pDL1098 which encodes LacI, SpecR and has a P _{tac} promoter. Cloned GFPmut3 downstream of P _{tac} promoter	This Study (SAD559)
P _{tac} -gfp-comM	P _{tac} N-terminal <i>gfp-comM</i> at <i>lacZ</i> Spec ^R , Δ comM Kan	VC0032 deletion strain containing an N-terminally GFP tagged comM at the <i>lacZ</i> locus under the control of P _{tac}	This Study (SAD921)
<i>gfp-comM</i>	N-terminal GFP-comM at the native locus	N-terminally GFP tagged comM was cloned into SAD030 at the native locus	This Study (SAD924)
P _{tac} -tfoX, <i>gfp-comM</i>	VC1153 OE Kan ^R , N-terminally GFP-comM at the native locus	Amplified VC1153 OE Kan from SAD061 and cloned fragment into SAD924	This Study (TMN0140 / SAD1320)

$P_{tac-tfoX}$, $gfp-comM^{K224A}$	VC1153 OE Kan^R , N-terminal $gfp-comM^{K224A}$ at the native locus, $Spec^R$	Amplified $comM^{K224A}$ mutation from SAD1026 and cloned fragment into SAD1320	This Study (TMN0148 / SAD1545)
Strains Used for Protein Purification			
ComM	ComM cloned into Amp^R StrepII expression vector	ComM N terminally tagged with 4x StrepII	This Study (pMB486 / SAD1330)
ComM ^{K224A}	ComM ^{K224A} cloned into Amp^R StrepII expression vector	ComM ^{K224A} N terminally tagged with 4x StrepII	This Study (pMB488 / SAD1331)
Rosetta 2 (DE3)		<i>E. coli</i> expression strain	(pMB131/ SAD1332)

Table S2- Primers used in this study-

Primer Name	Primer Sequence (5' to 3')*	Description
Mutant constructs		
ABD855	CATGAATCACTTTGGCATGAGG	$\Delta comM$ F1
ABD856	gtcgacggatccccggaatCATTGCTTCCCTTAGTATTTGATC	$\Delta comM$ R1
ABD857	gaagcagctccagcctacaTAGTACTCTGACCTGCAGAGTTC	$\Delta comM$ F2
ABD858	AAATTCAGAAAAACCACGTC	$\Delta comM$ R2
BBC749	CCGTGAAGCGAGCATGGTcgACCCGTCGCCGAG	$comM^{K224A}$ R1
BBC750	CTCCGGGACGGGTgcgACCATGCTCGCTTCACGG	$comM^{K224A}$ F2
ABD812	AAATGGAGTTTGATCGCATTTGGC	$\Delta recA$ F1
ABD921	gtcgacggatccccggaatCATTACTCTCTCCGGATAGTCACTC	$\Delta recA$ R1
ABD922	gaagcagctccagcctacaTAATCGGCAGGCTGAATGCAAAG	$\Delta recA$ F2
ABD815	TGATCAGCGTTTGGAAATACGTCCG	$\Delta recA$ R2
BBC401	ACCAGCAAAGCTAATAAAATCGAG	$\Delta pilA$ F1
BBC402	gtcgacggatccccggaatGAGCATATGCCTTGCTACACAAG	$\Delta pilA$ R1
BBC403	gaagcagctccagcctacaACTGCAGGTGCAACAATTAATACTAA	$\Delta pilA$ F2
BBC404	CGCCATACTAACCCAATACACTC	$\Delta pilA$ R2
ABD820	CGCTCTTATCTGCTTGATAATGG	$\Delta dprA$ F1
ABD998	gtcgacggatccccggaatCATTAACTGGCATCATCAACC	$\Delta dprA$ R1
ABD999	gaagcagctccagcctacaTAGCTATGATGATGGATATTTTGGATG	$\Delta dprA$ F2
ABD823	TGAAGTACAAGGCCAGTTACTGG	$\Delta dprA$ R2
BBC907	AAAGAGCAGTTGTCTGCTAGAC	$\Delta radA$ F1
BBC908	gtcgacggatccccggaatCAATCCTCGAAGTTGCTCTCAC	$\Delta radA$ R1
BBC909	gaagcagctccagcctacaTAATGGGTAGTTGGTTTTGAAC	$\Delta radA$ F2
BBC910	ATGAAGAAATCTTAGTCCGCAG	$\Delta radA$ R2
ABD824	TTTAGCCCCATTGGCGAAGTGGG	$\Delta mutS$ F1
ABD825	GAGTATCTTTGACGTATTGGATCtcatattatactaCATAATCTTATGTCGCTGCTTATC	$\Delta mutS$ R1
ABD826	GATAAGCAGCGACATAAGATTATGtagtataatatgaGATCCAATACGTCAAAGATACTC	$\Delta mutS$ F2
ABD360	AGATCTTGCTGATGACGCTTTACTC	$\Delta mutS$ R2
BBC717	AAATAGATTTGGTGACTTTACCTCC	VC1807::Ab ^R F1
ABD340	gtcgacggatccccggaatACGTTTCATTAGTCACCTCTATTGTAACTGTTTC	VC1807::Ab ^R R1
ABD341	gaagcagctccagcctacaTAGTCGAAAATAAAAAAAGAGGCTCGCCTC	VC1807::Ab ^R F2
BBC718	CTTTACGCCTGATTGTCTACAC	VC1807::Ab ^R R2

ABD332	GGCTGAACGTGGTTGTCGAAAATGAC	lacZ F1
ABD263	gtcgacggatccccggaatAACTGATCCAATTTTTTCAGCGCATATTTTGG	lacZ LPQEN::Ab ^R R1
ABD262	gaagcagctccagcctacaTGCCGCAGGAAAACCGCCCCCTaATC	lacZ LPQEN::Ab ^R F2
ABD256	CCCAAATACGGCAACTTGGCG	lacZ R2
ABD269	gaagcagctccagcctacaAATTGTGTAAACGTTTCCACAATTTAAATAG AGG	Spec ^R upstream of lacZ R1
ABD268	gtcgacggatccccggaatGGTGTGGTTCACAGAATCGGTG	Spec ^R upstream of lacZ F2
ABD495	AAAAAATCTTCAATCGCGAGTATCGGGTaGCGGTAGAGATACACA TCGCGAAAAGATGCC	lacZ 820bp linked R1
ABD494	TCTTTCGCGATGTGTATCTCTACCGCtAGCCGATACTCGCGATTGAA GATTTTTTTATCC	lacZ 820bp linked F2
ABD329	GAACATGGGGTGTACGGCAGTGCCATTaAACGATGTGCGGGTTTTG CCAATCTTG	lacZ 245bp linked R1
ABD328	CAAGATTGGCAAACCCGCACATCGTTtAATGGCACTGCCGTACACC CCATGTTC	lacZ 245bp linked F2
BBC1157	GTA AAACTTGAACGTGTTACGAATTGATTCAAAAAGTCTTGCGTC	Tm ^R Δ29bp R1
BBC1158	GACGCAAGACTTTTGAATCAATTCGTAACACGTTCAAGTTTTAC	Tm ^R Δ29bp F2
BBC747	GTCCAACCAACAGCCATTGGTTtTAGGTAATAGCTTTAAACAGGAG C	Tm ^R Point mutation
BBC748	GCTCCTGTTTTAAAGCTATTACCTAaAACCAATGGCTGTTGGTTGGA C	Tm ^R Point mutation
BBC498	GGGTAACGCCAGGGTTTTcCAGTCACGACGTTGTAAAAC	Spec ^R point mutant R1
BBC499	GTTTTACAACGTCGTGACTGGtAAAACCCTGGCGTTACCC	Spec ^R point mutant F2
BBC280	TCCACCACTTCCACctGCGACGTTCTGCGCACTGAGC	GFP-ComM R1
BBC351	GTTCTTCTCCTTTACGCATTACGTTACCTCCTTTTGATCAAAAAGCC TTCAGC	ComM-GFP R1
BBC352	GCaGGTGGAGCAGGTGGAGGACTTGCATCATTTCATAGC	ComM-GFP F2
ABD688	CCACTGTTGCGCAGTTGAATACC	tfoX OE Kan F1
ABD691	ATGATGTCAAACCATGAACCCGG	tfoX OE Kan R2
BBC331	CAATTTACACAGGATCCCGGGAGGAGTAACGTAATGGGACTTGC GATCATTTC	P _{tac} -comM F
BBC365	tgtaggctggagctgcttcCTAGACGTTCTGCGCACTGAG	P _{tac} -comM R
DOG0140	GTTGCTGCATTTGTTTCGATCTG	ΔcomM (ADP1) F1
DOG0141	gtcgacggatccccggaatCATACTATTATTGTTCCATTATGGTGC	ΔcomM (ADP1) R1
DOG0142	gaagcagctccagcctacaTATCGCAGTGAACATAGCTAAAA	ΔcomM (ADP1) F2
DOG0143	ATCAGTGGTTGGGAAGGTG	ΔcomM (ADP1) R2
Inserts for cloning		
MB1225	CGGGATCCATGGGACTTGCATCATTTCATAGCCG	comM ORF F1
MB1214	CGATCGATCTCGAGCTAGACGTTCTGCGCACTGAGC	comM ORF R1
MB1215	TATTGTTTCTCGGCCCTCCGGGGACGGGTGCGACCATGCTCGCTTCACGG CTGTGCGATT	comM ^{K224A} F1
MB1216	AATCGCACAGCCGTGAAGCGAGCATGGTCGCACCCGTCGCCGGAGGGC CGAGAAACAATA	comM ^{K224A} R1
Oligomerization assays/Negative stain EM		
ABD363	CGTTAAATGAAATTAATACGACTCACTATAGGGAGAGAGGTTTGCTGT TTGAGAAGCC	ssDNA substrate for oligomerization
EMSA probe		
BBC742	ATTCCGGGGATCCGTCGACCTGCAGTTCAGAAGCAGCTCCAGCCTACA	EMSA binding probe (ssDNA, dsDNA)
BBC743	TGTAGGCTGGAGCTGCTTCTGAACTGCAGGTGACGGATCCCCGGAAT	EMSA binding probe (dsDNA)

Helicase substrates		
MB1167	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTCCTCACATAGCGTTC	Poly-dT 5' tail
MB1168	GAACGCTATGTGAGTGACACTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Poly-dT 3' tail
MB820	GAACGCTATGTGAGTGACAC	No-tail bottom primer
MB821	GTGTCCTCACATAGCGTTC	No-tail top primer

*Lowercase letters indicate overlap sequences for SOE PCRs or mutated nucleotides when generating point mutations

SUPPLEMENTARY METHODS

Protein expression and purification

The *comM* open reading frame was PCR-amplified from *V. cholerae* genomic DNA using oligonucleotides MB1225 (CGGGATCCATGGGACTTGCGATCATTCATAGCCG) and MB1214 (CGATCGATCTCGAGCTAGACGTTCTGCGCACTGAGC), digested with *Bam*HI and *Xho*I, and ligated into the same sites in plasmid pMB131 to generate pMB486. This cloning added an N-terminal 4x Strep-tag II to the translated protein. The expression plasmid encoding the ATPase- and helicase-dead *comM-K224A* allele (pMB488) was created site-directed mutagenesis of pMB486 with oligonucleotides MB1215 (TATTGTTTCTCGGCCCTCCGGGGACGGGTGCGACCATGCTCGCTTCACGGCTGTGCGATT) and MB1216 (AATCGCACAGCCGTGAAGCGAGCATGGTCGCACCCGTCGCCGGAGGGCCGAGAAACAATA) and verified by DNA sequencing (ACGT, Inc.). Expression plasmids were transformed into Rosetta 2(DE3) pLysS cells and selected for at 37°C on LB medium supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Fresh transformants were used to inoculate one or more 5-mL LB cultures supplemented with antibiotics and incubated at 30°C for ~6 h with aeration. These starter cultures were then diluted 1:100 in ZYP-5052 autoinduction medium containing 1x trace metals mix (3), 100 µg/mL ampicillin, and 34 µg/mL chloramphenicol and incubated at 22°C with agitation to OD₆₀₀ >3 (15-18 h). Cells were harvested by centrifugation for 10 min at 5,500 x g and 4°C. Cell pellets were weighed and frozen at -80°C prior to lysis or for long-term storage.

Frozen cell pellets were thawed at room temperature by stirring in 4 mL/g cell pellet resuspension buffer (25 mM Na-HEPES (pH 7.5), 5% (v/v) glycerol, 300 mM NaOAc, 5 mM MgOAc, and 0.05% Tween-20) supplemented with 1x protease inhibitor cocktail (Sigma), and 20 µg/mL DNase I. Cells were lysed by six passes through a Cell Cracker operated at >1000 psi. All subsequent steps were performed at 4°C. The soluble fraction was clarified by centrifugation for 30 min at 33,000 x g followed by filtering the supernatant through a 0.22-µm membrane. This mixture was then applied to a Strep-Tactin Sepharose column (IBA) pre-equilibrated in resuspension buffer using an ÄKTA Pure (GE Healthcare Life Sciences). The column was washed with 20 column volumes (CVs) of resuspension buffer, 10 CVs of resuspension buffer supplemented with 5 mM ATP, and 10 CVs of resuspension buffer. Protein was eluted with 15 CVs of resuspension buffer supplemented with 2.5 mM desthiobiotin (IBA). Column fractions were examined on 8% SDS-PAGE gels run at 20 V/cm and stained with Coomassie Brilliant Blue R-250 (BioRad). Peak fractions were pooled, concentrated with Amicon Ultra-4 30K centrifugal filters, and loaded onto a HiPrep 16/60 Sephacryl S-200 HR column (GE Healthcare Life Sciences) pre-equilibrated in gel filtration buffer (25 mM Na-HEPES (pH 7.5), 5% glycerol, 300 mM NaCl, 5 mM MgCl₂, and 0.05% Tween-20). The protein was eluted with 1.5 CVs gel filtration buffer, and fractions were analyzed by SDS-PAGE as above. Peak fractions were pooled, snap-frozen with liquid nitrogen, and stored at -80°C.

The *Saccharomyces cerevisiae* Pif1 helicase was overexpressed in Rosetta cells from plasmid pMB330 as described for ComM above. Pif1 purification was likewise identical, except the protein from the Strep-Tactin column was polished by Ni-affinity chromatography instead of size exclusion. Briefly, the pooled peak fractions were applied to a His60 Ni Superflow (Clontech) gravity column, washed with 10 CVs resuspension buffer supplemented with 25 mM imidazole (pH 8), and eluted with 4.5 CVs of a step gradient of resuspension buffer containing 100 mM, 250, and 500 mM

imidazole (pH 8). Peak fractions were pooled, buffer exchanged into storage buffer (4), snap-frozen with liquid nitrogen, and stored at -80°C.

Blue native PAGE

Oligomerization of ComM protein was assayed using Blue Native PAGE electrophoresis. 2.5 μ M purified ComM was incubated for 30 min at room temperature in reaction buffer [10 mM Tris-HCl pH7.5, 20mM KCl, 1mM DTT, 10% Glycerol] with 5 mM ATP and/or 5 nM ssDNA (oligo ABD363) where indicated. 1 μ L 20x sample buffer [5% Coomassie G-250, 0.5 M aminocaproic acid pH 7] was added to each reaction and samples were run on 4-16% Native PAGE gels [gel buffer = 0.5 M aminocaproic acid pH 7.0, 0.05 M Bis-Tris pH 7.0]. The cathode buffer was composed of 50 mM Tricine, 15 mM Bis-Tris pH 7.0, 0.02% Coomassie G-250, while the anode buffer was composed of 50 mM Bis-Tris pH 7.0. Samples were run at 150 V for 30 min, then 200 V for 45 min.

Negative stain electron microscopy

The nominal magnification for the images is 60,000x, which is equivalent to 1.8 Å per pixel at the final image. Initial image processing, particle boxing, and CTF determination were performed using EMAN2 (5). A phase-flipped particle dataset was then imported into Relion (6) for 2D classification. Classes showing noisy images were discarded at this stage. As we observed clear six-fold symmetry from the classes, the subsequent processing imposed C6 symmetry. The remaining “good” classes were used to generate the initial models using e2initialmodel.py. The 3D classification was carried out using the initial model that was low-pass filtered to 40 Å to eliminate the possible effect from the model bias. Three 3D classes were obtained; the highest population (46%) of the classes was subjected to further structure refinement in Relion. Approximately 32,958 particles were used to generate the final 3D reconstruction. The reported resolution is ~13.8 Å using gold-standard Fourier shell correlation at a 0.143 cutoff; however, it is an over-estimated value because of the use of negative stain. The structure is rendered using UCSF Chimera (7).

Helicase Assays

Helicase assay substrates were made by 5'-end labelling oligonucleotides MB1167 (TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTCACATAGCGTTC) and MB821 (GTGTCACATAGCGTTC) with T4 polynucleotide kinase (T4 PNK; NEB) and γ [³²P]-ATP. Labelled oligonucleotides were separated from free label using illustra ProbeQuant G-50 micro columns (GE Healthcare) following the manufacturer's instructions. Oligonucleotides were annealed by incubation with an equimolar amount of complementary or partially complementary oligonucleotides overnight at 37°C in annealing buffer (20 mM Tris-HCl [pH 8], 4% glycerol, 0.1 mM EDTA, 40 μ g/mL BSA, 10 mM DTT, and 10 mM MgOAc) (8).

The DNA fork was made annealing oligonucleotides MB1167 with MB1168 (GAACGCTATGTGAGTGACACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT), the blunt dsDNA was made by annealing MB821 to MB820 (GAACGCTATGTGAGTGACAC), the 5' Tail substrate was made by annealing MB1167 and MB820, and the 3' Tail substrate was made by annealing MB821 and MB1168. DNA unwinding was assessed by incubating the indicated concentrations of helicase with 5 mM ATP and 0.1 nM radiolabelled fork in resuspension buffer. Reactions were incubated at 37°C for 30 min and stopped with the addition of 1x Stop-Load dye (5% glycerol, 20 mM EDTA, 0.05% SDS, and 0.25% bromophenol blue) supplemented with 400 μ g/mL SDS-Proteinase K followed by a

10-min incubation at 37°C. Unwound DNA was separated on 8% 19:1 acrylamide:bis-acrylamide gels in TBE buffer at 10 V/cm. Gels were dried under vacuum and imaged using a Typhoon 9210 Variable Mode Imager. DNA binding was quantified using ImageQuant 5.2 software.

SI REFERENCES

1. Miller, V.L., DiRita, V.J. and Mekalanos, J.J. (1989) Identification of *toxS*, a regulatory gene whose product enhances *toxR*-mediated activation of the cholera toxin promoter. *J Bacteriol*, **171**, 1288-1293.
2. Dalia, A.B., Lazinski, D.W. and Camilli, A. (2014) Identification of a membrane-bound transcriptional regulator that links chitin and natural competence in *Vibrio cholerae*. *MBio*, **5**, e01028-01013.
3. Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41**, 207-234.
4. Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C. and Zakian, V.A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, **497**, 458-462.
5. Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *Journal of structural biology*, **157**, 38-46.
6. Scheres, S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology*, **180**, 519-530.
7. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
8. Kanter, D.M. and Kaplan, D.L. (2011) Sld2 binds to origin single-stranded DNA and stimulates DNA annealing. *Nucleic Acids Res*, **39**, 2580-2592.