Structural characterization of NrnC identifies unifying features of diribonucleotidases

Justin D. Lormand[1], Soo-Kyoung Kim[2], George A. Walters-Marrah[1], Bryce A. Brownfield[3], J. Christopher Fromme[3], Wade C. Winkler[2], Jonathan R. Goodson[2], Vincent T. Lee[2], and Holger Sondermann[1,4,5#]

[1] Department of Molecular Medicine, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

[2] Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, USA

[3] Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

[4] Centre for Structural Systems Biology, Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany

[5] Christian-Albrechts-Universität, 24118 Kiel, Germany

[#] To whom correspondence should be addressed:
Email: holger.sondermann@cssb-hamburg.de, phone: +49 (0)40 8998 87680

Running title: Structural basis for diribonucleotidase function of NrnC-type RNases

Abstract

RNA degradation is fundamental for cellular homeostasis. The process is carried out by various classes of endolytic and exolytic enzymes that together degrade an RNA polymer to mono-ribonucleotides. Within the exoribonucleases, nano-RNases play a unique role as they act on the smallest breakdown products and hence catalyze the final steps in the process. We recently showed that oligoribonuclease (Orn) acts as a dedicated diribonucleotidase, defining the ultimate step in RNA degradation that is crucial for cellular fitness . Whether such a specific activity exists in organisms that lack Orn-type exoribonucleases remained unclear. Through quantitative structure-function analyses we show here that NrnC-type RNases share this narrow substrate preference with Orn. Although NrnC employs similar structural features that distinguish these two classes as diribonucleotidases from other exoribonucleases, these key determinants for diribonucleotidase activity are realized through distinct structural scaffolds. The structures together with comparative genomic analyses of the phylogeny of DEDD-type exoribonucleases indicates convergent evolution as the mechanism of how diribonucleotidase activity emerged repeatedly in various organisms. The evolutionary pressure to maintain diribonucleotidase activity further underlines the important role these analogous proteins play for cell growth.

Introduction

Traditionally, nano-RNases – enzymes that act on short, typically 2 to 7 residue-long RNA substrates – have occupied a distinct role in RNA metabolism as they catalyze the final steps in RNA degradation. To date, three main enzyme families have been assigned this function: Oligoribonucleases (Orn) , nano-RNase A and B (NrnA and NrnB) , and nano-RNase C (NrnC). NrnA is comprised of DHH-DHHA1 domains and is suggested to act as a bidirectional exonuclease that cleaves short RNA fragments from 3’ to 5’ and longer substrates from 5’ to 3’ . Orn and NrnC are 3’-5’ exonucleases with a DnaQ fold containing the catalytic DEDD motif , a domain that is common in enzymes that act on nucleic acids. Notably, Orn (and its eukaryotic ortholog REXO2) or NrnC activity is critical for cellular growth, rendering them unique amongst the exoribonucleases known to date .

Despite its classification as a nano-RNase, we recently reported that Orn acts primarily as a diribonucleotidase, assigning it a highly specific and unique function in clearing the diribonucleotide pool as the terminal step in RNA degradation. Substrate-bound structures of bacterial and human orthologs, Orn and REXO2, respectively, revealed the scissile bond of the dinucleotide surrounded by the conserved, catalytic DEDDh motif that is involved in divalent cation coordination and catalysis. More importantly, key determinants for the RNA length preference of Orn and REXO2 include a leucine residue that wedges between the two bases of the diribonucleotide and a phosphate cap – invariable residues that coordinate the 5’-phosphate on the substrate and limit substrate length to two nucleotides . An independent study on REXO2 confirmed our structural analysis and established the human enzyme as a diribonucleotidease in mitochondria, where its activity alters gene expression , a function that relates to the role of diribonucleotides in transcription initiation .

Our previous work on Orn-type RNases was motivated by three main considerations: Orn’s essential role for growth in many bacteria, its role in cyclic dinucleotide signaling, and a lack of understanding how substrate specificity towards short RNA substrates, and in turn lack of activity towards longer RNAs, is achieved . The latter is a basic question as the unique substrate profile is the defining characteristic of this class of enzymes. We demonstrated that Orn has a much higher preference for diribonucleotides compared to 3-7 residue-long RNAs than anticipated previously. This selectivity is due to an active site that is exquisitely suited for diribonucleotides with a 5’ phosphate. The enzyme’s diribonucleotidase activity is required for normal bacterial growth and clearance of c-di-GMP breakdown products. Notably, a knock-out of *Pseudomonas aeruginosa* Orn can not only be complemented by various Orn orthologs, but also by the other three nano-RNases: NrnA, NrnB, and NrnC . However, it is not clear whether this functional complementary correlates with a narrow substrate specificity for dinucleotides.

A recent structural study of NrnC from the soil bacterium *Agrobacterium tumefaciens* identified a homo-octameric assembly as the enzyme's functional unit . The unit can be divided into two stacked rings, each composed of four NrnC monomers, which together form a central channel lined with eight active sites. Mutagenesis of key active-site residues confirmed the requirement of the conserved DEDDy motif for catalysis and identified positively charged residues lining the channel, which are also important for function. Activity assays indicated single-stranded RNA and double-stranded DNA as potential substrates of NrnC, however the structural basis for substrate specificity has not been established.

Here, we ask the fundamental question: What is the substrate specificity of NrnC-type enzymes, an RNase that is essential for the growth of gram-negative bacteria such as *Caulobacter crecentus, Bartonella henselae* and *Brucella abortus* . We present crystal structures of NrnC from *B. henselae* and *Brucella melitensis* in their substrate-bound and apo states. The structures confirm an octameric assembly as a conserved feature of NrnC-type RNases. The substrate-bound states reveal, similar to Orn, a narrow active site that appears optimized for dinucleotides. This preference is reflected also in the enzyme's activity profile. A comparative genomics analysis indicates that Orn and NrnC, despite using a common DnaQ fold, evolved separately as isofunctional enzymes . Considering also the distribution of the structurally unrelated, yet functionally overlapping NrnA- and NrnB-type RNases predominantly in organisms that lack Orn and NrnC underlines the importance to maintain diribonucleotidase activity for cellular function in the bacterial and eukaryotic domains of life.

Results and Discussion

Overall structure of diribonucleotide-bound NrnC. The earlier observation that NrnC expression is able to complement a *P. aeruginosa orn* deletion strain indicates that both enzymes function on dirbonucleotides , consistent with their initial classification as nano-RNases .

However, while we showed that Orn acts as a dedicated diribonucleotidase , the structural basis for NrnC's substrate preference remained not well defined. Specifically, it was not clear how NrnC distinguishes between short RNAs and longer polymers. To answer this question, we determined crystal structures of wild-type *B. henselae* and *B. melitensis* NrnC (NrnC_{Bh} and NrnC_{Bm}, respectively) bound to pGG and, in the case of NrnC_{Bm}, also in the substrate-free state. NrnC_{Bh} forms a homo-octameric assembly comprised of two C4-symmtric rings (Figure 1A) as observed with the previously determined substrate-free, orthologous protein from *A. tumefaciens* (69% sequence identity compared to NrnC_{Bh}, monomer/octamer all-atom rmsd 0.4/1.3 Å; PDB:5ZO3,). The two rings stack with the same face, tail-to-tail, forming a D4-symmetric octamer. The contacts between the subunits within each ring are dominated by a few polar interactions spanning between 441 to 512 Å^2 (Figure 1-figure supplement 1A-B) (determined by PISA). In contrast, pairwise, homotypic interactions between the two rings involve an extensive hydrophobic interface of 1204 Å^2 via an antiparallel packing of the last helix of the NrnC fold. This mode of ring stacking positions the C-terminus of one monomer so that it reaches into the active site of the adjacent monomer in the other ring (Figure 1-figure supplement 1A and 1C).

The diribonucleotide pGG is bound to all eight active sites of the NrnC_{Bh} octameric assembly (Figure 1B). The active sites face the center of the central cavity formed by the NrnC octamer, positioned mid-way of each ring. In this crystallographic state, the residues of the catalytic DEDDy signature motif (D^{25}, E^{27}, D^{84}, D^{155}, Y^{151}) are primed for accepting divalent metal ions for nucleotide hydrolysis, further supported by the observation that side chain of Y^{151} coordinates a water molecule, which likely serves as the attacking nucleophile in the reaction (Figure 1C).

The structure also reveals the molecular basis for substrate coordination. Reminiscent of Orn's active site, the bases of the diribonucleotide are splayed apart by a leucine residue, or L-wedge (L^{31}) (Figure 1C). Continuing with the parallels to Orn, the 5' phosphate of pGG is coordinated by several residues forming a 'phosphate cap', in this case basic residues H^{79}, K^{103}, and H^{205}, the latter being the second to last residue of the protein, contributed from a subunit from the adjacent ring. The specific motifs coordinating the substrate are invariable in NrnC orthologs, in contrast to the exterior surface with overall lower conservation (Figure 1D), and culminate in a length-restricted active site that appears optimized for diribonucleotides. The comparison with the structurally related Klenow fragment bound to a longer RNA substrate supports this notion as residues K^{103} and H^{205} of NrnC's phosphate cap directly block the path for longer substrates (Figure 1-figure supplement 2A and 2C) . Similarly, RNase D, a homologous exoribonuclease that processes longer and stable RNA molecules, presents a more expansive, open active site, although the lack of a substrate-bound structure prevents a more direct comparison of this state (Figure 1-figure supplement 2B and 2C) .

Similar binding poses to pGG at NrnC were observed with pAA and pGC (Figure 1-figure supplement 3A-C), suggesting that most if not all diribonucleotides can be accommodated by NrnC. A co-crystal structure with the di-phosphorylated mononucleotide pAp, a metabolite described as an inhibitor of NrnC , shows the ligand predominantly occupying the 5' site of the active site with the 3' phosphate engaging the catalytic motif and the 5' phosphate being coordinated by the phosphate cap of NrnC (Figure 1-figure supplement 3D).

To establish the functional relevance of the crystallized state of NrnC, we initiated *in crystallo* catalysis by soaking NrnC$_{Bh}$•pGG co-crystals in solutions with divalent cations, magnesium (Mg^{2+}) or manganese (Mn^{2+}). The enzyme became catalytically active with the addition of either cation, with the resulting electron densities showing a broken phosphodiester bond (Figure 1–figure supplement 4). The experiments also confirmed a two-metal mechanism first by interpretation of electron density upon Mg^{2+} soaking (Figure 1–figure supplement 4A). The placement of active site metal atoms was subsequently confirmed using anomalous data collected on Mn^{2+}-soaked crystals (see anomalous difference map, (Figure 1–figure supplement 4B). In these post-catalysis structures, the 5' GMP appears to leave the active site first, as suggested by weaker electron density indicative of lower mononucleotide occupancy at the 5' site compared to the 3' site.

The characteristic active-site motifs of NrnC$_{Bh}$ contribute to diribonucleotidase activity. The structural analysis revealed the molecular basis for substrate binding to NrnC, identifying molecular features that constrain the active site. To assess their relevance for NrnC's catalytic activity, we tested tag-less, purified NrnC$_{Bh}$ and structure-based point mutants thereof in an in vitro activity assay. All mutant proteins retained their quaternary structure and purified as octamers, indistinguishable from wild-type NrnC (Figure 2-figure supplement 1). 5'-^{32}P-radiolabeled pGG was incubated with wild-type or mutant enzymes in the presence of divalent cations at physiological ionic strength. Quenched reactions were resolved via urea-denaturing

PAGE to observe nucleolytic cleavage over time (Figure 2A). With wild-type NrnC$_{Bh}$ the majority of pGG was processed already by the first timepoint at 30 seconds, and complete cleavage of pGG to GMP was achieved by 3 minutes. Nucleolytic activity on pGG was completely inhibited by alanine mutation of the catalytic DEDDy motif residues D^{25} and Y^{151}. Intermediate cleavage kinetics were evident with proteins with disrupted leucine wedge (NrnC$_{Bh}$-L^{31}A) as well as the phosphate cap (NrnC$_{Bh}$-H^{79}A, -K^{103}A, or -H^{205}A). NrnC binding to and activity on pGG were slightly inhibited by pAp (Figures 2B and 2C). In contrast, the diguanylate compound GG, which lacks a 5' phosphate, did not inhibit NrnC's binding to or activity on pGG (Figures 2B and 2C), further elaborating on the importance of the interaction between phosphate cap residues and the 5' phosphate for NrnC function.

We previously proposed a model of cellular fitness in which the loss of *orn* leads to toxic diribonucleotide accumulation detrimental to the cell. The *orn* deletion in *P. aeruginosa* manifests as small colony growth, which is reversible by complementation with *orn* or other nano-RNases expressed from a plasmid (Figure 2D) . Here we used the rescue of the small colony phenotype as a readout of NrnC diribonucleotidase activity in cells, quantified as colony size. Complementation of the deletion strain with wild-type NrnC$_{Bh}$ (with a C-terminal HA tag for detection) restored normal colony size, while NrnC alleles containing mutations within the DEDDy motif failed to complement (Figure 2D). Further, NrnC alleles containing mutations in the L-wedge or the phosphate cap showed reduced complementation effects (L^{31}A, K^{103}A, K^{132}A, or H^{205}A; Figure 2D). NrnC$_{Bh}$-H^{79}A failed to complement the *orn* deletion. Western blot analysis established protein expression for all mutant and wild-type NrnC variants, with the exception of NrnC$_{Bh}$-H^{79}A, which expressed poorly in *P. aeruginosa*, preventing a distinction between failure to rescue because of the mutation or protein levels, or both (Figure 2-figure supplement 2). Together, these data confirm the importance of the motifs identified in the substrate-bound NrnC structures for the enzyme's diribonucleotidase activity *in vitro* and in cells.

Structural comparison of NrnC substrate-bound states reveals an active site optimized for diribonucleotides. To further understand the structural basis of NrnC's substrate preference, we determined structures of *B. melitensis* NrnC (NrnC$_{Bm}$) in several apo and partially-pGG-bound states. The overall structure of pGG-bound NrnC$_{Bm}$ is virtually identical to NrnC$_{Bh}$, namely a homo-octameric assembly with eight active sites pointing toward the central channel (Figures 3A and 3B). The acidic active site residues (D^{24}, E^{26}, D^{83}, D^{154}, and Y^{150}) as well as the L-wedge (L^{30}) and phosphate cap (H^{78}, K^{102}, and H^{204}) are structurally and functionally conserved.

NrnC$_{Bm}$ incubated with pGG crystallized with two molecules per asymmetric unit, only one of which was bound to the diribonucleotide. The resulting octameric assembly contains alternating apo- and pGG-bound subunits per tetrameric ring (Figures 3B). This mixed-state structure allowed us to propose features modulating substrate binding. In substrate-bound monomers (including in the structures of NrnC$_{Bh}$), the DEDDy residue Y^{150} points inward toward the scissile phosphodiester bond, coordinating the attacking water. In contrast, the Y^{150} side chain points away from the active site in substrate-free NrnC (Figures 3A and B). A loop from an adjacent subunit that mediates inter-ring contacts between the monomers and that includes residue D^{113} buttresses the 3' base of the substrate. While the octameric assembly remains in the absence of substrate, this loop moves outward from the active site and D^{113} rotates away from the substrate (Figure 3D). The most drastic conformational change however is attributed to a flexible loop spanning residues ^{130}SKQQQS^{135}. This loop is ordered and positioned in contact with the diribonucleotide in both ortholog structures (Figures 1C, 3B, and 3C). In this state, residue Q^{134} of NrnC$_{Bm}$ (or Q^{135} in NrnC$_{Bh}$) contacts the scissile phosphate via hydrogen bond; K^{131} (or K^{132} in NrnC$_{Bh}$) points towards the 5' phosphate, thus contributing to the phosphate cap. In contrast, in the substrate-free state NrnC$_{Bm}$, captured in the mixed-state structure or a homogeneous apo-state structure, this loop swings away from the active site or is completely disordered, leading to an overall widening of the active site (Figure 3A; Figure 1-figure supplement 2C).

In a second apo-NrnC$_{Bm}$ crystal form, we captured an alternate conformation in which the conserved hydrophobic residue F^{79} moves into the active site (Figure 3-figure supplement 1). The movement of this residue, which is adjacent to the phosphate cap residue H^{78}, is realized through a flip of the peptide backbone and would introduce a clash with the nucleotide. In the same structure the flexible SKQQQS loop is collapsed into the active site, trapping the catalytic residue Y^{150} in an intermediate conformation, which would introduce further clashes with the nucleotide substrate. Simultaneously, a rotamer change of the phosphate cap residue H^{204} pivots its sidechain away from the active site, opening it for access to substrates (Figure 3-figure supplement 1). This conformation may depict a post-hydrolysis state, suggesting a mechanism for product release.

Cryo-electron microscopic analysis supports a narrow substrate preference of NrnC. Efforts to crystallize NrnC bound to longer substrates yielded structures with either empty active sites or with only pGG being resolved in the resulting electron densities. Residual RNase activity over the course of the crystallization or an impact of the longer substrates on crystal packing could contribute to the inability to resolve longer substrates, assuming they bind in the first place. As an alternative to crystallography, we determined NrnC$_{Bh}$ structures bound to pGG, pAGG and pAAAGG by cryo-electron microscopy (cryoEM), a technique that can visualize complexes formed after short equilibration periods and does not rely on proteins packing in a lattice. Considering that NrnC forms an octamer as the biological unit, we processed each data set with C1 and D4 symmetry, with the resulting models consisting of 8 independent or an averaged chain, respectively (resolutions range from 2.72 to 3.27 Å) (Figure 4; Figure 4-figure supplements 1-8). Processing with C1 symmetry preserves the individuality of each monomer allowing the observation of differences e.g., between active sites within the octamer. Applying D4 symmetry during processing averages all 8 monomers, yielding a consensus model; however, regions with conformational differences between individual monomers may contribute to apparent disorder in the electron-density maps. In an additional experiment with the 5-mer pAAAGG as the substrate, $CaCl_2$ was added to the buffer to probe whether the addition of Ca^{2+} ions, which prevent catalysis but can support substrate binding, has a qualitative effect on the binding of longer RNAs.

The structure of pGG-bound NrnC$_{Bh}$ confirmed all active site features described above based on the crystallographic data, namely a narrow active site, a phosphate cap coordinating the 5’ phosphate of the substrate, a L-wedge splaying apart the two bases, and an active-site facing, well-resolved SKQQQS loop (Figure 4A, Figure 4-figure supplement 9A). Refinement with lower and higher symmetry resulted in comparable density maps, indicating a consensus state with 8 nearly identical active sites.

The density maps of NrnC bound to any of the longer substrates resolved invariably only a diribonucleotide at the active site (with poorer density indicating the position of the ribose of the third 5’-substrate residue) (Figure 4B). The remainder of the longer substrates appeared disordered. Using D4-symmetry-averaged data, we also noticed consistently disorder of the SKQQQS loop (Figure 4B). Inspection of symmetry-less (C1) density maps revealed different loop conformations in the individual monomers of the octamers that, when symmetry-averaged, result in the apparent disorder (Figure 4-figure supplement 5-8). Notably, the majority of monomers contain a disengaged loop conformation, leaning away from the active site. Addition of Ca^{2+} with the pAAAGG substrate results in increased ordering of the loop in a disengaged state similar to that observed in the substrate-free crystal structures of NrnC*Bm* (Figures 3E, Figure 4-figure supplements 9B). Together, these results suggest that longer substrates may bind NrnC, but only the first two 3' residues are well coordinated at the active site. Furthermore, RNAs with more than two residues in length increase conformational variability at the active site, likely impacting catalytic activity towards those substrates. In summary, the combined structural data indicate NrnC is optimized for diribonucleotide processing over longer substrates, mirroring our analysis of Orn and REXO2 .

NrnC_{Bh} acts preferentially on diribonucleotides. Although the structural analysis revealed a narrow active site akin to that of Orn, NrnC's substrate-length preference has not been formally proven in this new context. To address this, we conducted kinetic experiments with RNAs of increasing length as substrates, following protocols established for Orn . In the initial assay, RNAs with two or more residues were in 200-fold abundance over NrnC, a condition where NrnC turned over the entire pGG pool within 1 min (Figure 5A). In comparison to NrnC's expedient activity on pGG, increasing the substrate length by only a single residue (pAGG) resulted in a striking decrease of nucleolytic cleavage under otherwise identical conditions. For RNAs with four and more residues, the substrate was incompletely processed and a band indicative of a cleavage product that was one base shorter at the 3' end slowly increased over the course of the experiment (Figures 5A and 5B). Increasing the concentration of NrnC increased activity on the longest substrate tested, an RNA with 7 bases (5'-^{32}P-labeled AAAAAGG, pAAAAAGG), but full conversion to mononucleotides required a ratio of 1:1 NrnC:RNA, indicating a comparatively inefficient, and likely less physiological mechanism (Figure 5B). Furthermore, and similar to the kinetics observed with Orn , a diribonucleotide intermediate was undetectable with enzyme concentrations that were required to observe the successive breakdown of the longer RNA, revealing the rapid turnover of dinucleotides, proceeding at much faster timescales than with any other intermediate that could be readily observed (Figure 5B).

Quantification of NrnC’s binding to radiolabeled RNA substrates of different length agreed with the preference of NrnC to cleave diribonucleotides. Affinities of NrnC for radiolabeled substrates (2-7-mer RNA) were determined at physiological ionic strength and in the presence of Ca^{2+} to inhibit any potential residual catalysis . NrnC$_{Bh}$ bound to pGG with a K_d = 17.7 nM. Increasing substrate length by just one additional residue resulted in a nearly 200-fold decrease in affinity (K_d pAGG = 3.49 µM) (Figure 5C). RNA substrates of four, five, or six residues showed similar decreases in affinity, while a 7-mer RNA substrate showed intermediate binding strength with a 32-fold decrease from pGG (K_d pAAAAAGG = 576 nM). As another method to assess substrate preference, competition experiments were performed by incubating NrnC$_{Bh}$ with ^{32}P-pGG with or without unlabeled RNAs of different length. While unlabeled pGG was able to displace radiolabeled pGG quantitatively on NrnC, longer substrates were less potent competitors under otherwise identical conditions (Figure 5-figure supplement 1). Together, the combined structural and biochemical results argue for a strong preference of NrnC towards the shortest species of RNAs, diribonucleotides with a 5’ phosphate.

***NrnC$_{Bh}$* processes DNA under specific experimental conditions.** The *Agrobacterium tumefaciens* NrnC octamer was previously interpreted as a conduit for long, polymeric substrates, in particular single-stranded RNA and single-stranded or double-stranded DNA (dsDNA), based on the octamer’s channel dimensions and positioning of the active sites . DNase activity was proposed to allow NrnC octamers to act on opposite ends of dsDNA to completely unwind and degrade it by passing the strand through the central channel. Here, we asked whether NrnC$_{Bh}$ could act on DNA substrates (Figure 5-figure supplement 2). Under near-physiological ionic strength and in the presence of either Mg^{2+} or Mg^{2+} ions, NrnC$_{Bh}$ failed to degrade a 1.5 kb-long dsDNA fragment (Figure 5-figure supplements 2B and 2C). Degradation of dsDNA was only observed at low ionic strength (~0-60 mM NaCl) and only in the presence of Mn^{2+} ions, conditions that match those used for *Agrobacterium tumefaciens* NrnC. The requirement for Mn^{2+} for activity on DNA substrates mirrors the previously reported phenomenon observed with the 3’-5’ exonuclease EXD2 that acts on both ribo- and deoxyribonucleotides .

To examine whether potential NrnC activity on dsDNA was dependent on the presence of a 5’ phosphate or a 5’ overhang, dsDNA fragments were treated either with restriction enzymes, T4 polynucleotide kinase (PNK), or calf intestinal alkaline phosphatase (CIP) as indicated

(Figure 5-figure supplements 2A and 2D). With any of these modifications and similar to the parent, blunt dsDNA, degradation was only observed in combination of the absence of NaCl and presence of Mn^{2+} ions. These experiments call into question a general DNase activity of NrnC, although such an activity under specific cellular conditions remains plausible.

Phylogenetic analysis indicates repeated evolution of critical diribonucleotidase activity. After having elucidated the mechanisms that signify NrnC-type and Orn-type RNases as dedicated diribonucleotidases, we reinvestigated the evolutionary distribution of these enzymes, relative to their structural homologs RNase D and RNaseT, respectively. We included NrnA and NrnB in this analysis as the only other general RNases unrelated to NrnC and Orn, which also could stand in for Orn in a *P. aeruginosa* deletion strain, as shown previously . We identified homologs of the aforementioned proteins in the full UniprotKB database (version 2020_03), correlated the appearance of homologs in each species, and used the results to identify the spread of each type of nuclease at the order level (Figure 6). Orn was widespread in all Eukaryotic groups, although in bacteria it was infrequently found outside of the Proteobacteria and Actinobacteria phyla. The structurally related RNase T is largely limited to Proteobacteria, with a few exceptions. Of the other major nucleases included in this analysis, NrnA was most frequently found in many bacteria. In contrast, NrnB occurs more sparsely distributed, potentially suggesting a more specialized function in individual organisms. NrnC is primarily found in most Alphaproteobacteria and Cyanobacteria, often overlapping with the occurrence of the homologous RNase D (with respect to the catalytic-domain sequence), although RNase D is present in many more bacterial genomes than NrnC. The three proteins that act as effective dirbonucleotidases (Orn, NrnC, NrnA/NrnB) were largely – although not always – mutually exclusive; most bacterial taxa in this analysis had only one of the three, with the notable exception of Actinobacteria that frequently contained both Orn and NrnC (Figure 6).

We used the identified sequences of DnaQ family ribonucleases, each family curated individually, to create multiple sequence alignments and ultimately a combined phylogenetic tree of representative sequences of Orn, RNase T, NrnC, and RNase D family proteins (Figure 7). Similar to a previous analysis , the DEDDh sequences (Orn and RNase T) segregated from the DEDDy sequences (RNase D and NrnC), highlighting the distinct evolutionary background of NrnC and Orn. Emanating from the ancestorial sequence, the first branch separated RNase T and

Orn from RNase D and NrnC. This ancient split was soon followed by a split between RNase T and Orn. In the other lineage, NrnC and RNase D diverged from each other after a longer effective evolutionary time, presumably following a duplication from the more closely related RNase D (Figure 7A). Due to this relatively recent split, and as it frequently co-occurs with RNase D, NrnC appears to have arisen from a more recent specialization event.

Catalytic, L-wedge, and phosphate cap residues are strictly conserved in many NrnC orthologs (Figures 1D and 7B). RNase D, the closest relative to NrnC, shares the catalytic DEDDy motif with NrnC, but lacks the L-wedge and phosphate cap (Figure 7B), suggesting that these features distinguish NrnC (and Orn) from other RNases and DnaQ-fold enzymes. The phylogenetic analysis also identified a unique group of NrnC-like enzymes in Actinobacteria, which shares most of the characteristic NrnC features, including conservation of the active site, most of the phosphate cap and wedge residues. However, subtle but specific changes in important residues (*i.e.*, a Q-to-R change in the SKQQQS loop and the phosphate caps H^{205}, which is replaced by a tyrosine residue) hint at the possibility of distinct function of the NrnC orthologs in this subgroup (Figure 7). Some Actinobacteria also encode an additional Orn and/or NrnA in the same genome, which could suggest either redundant functions or further specialization of these enzymes in these organisms (Figure 6).

Conclusions

The establishment of NrnC as a dedicated diribonucleotidase led us to compare and contrast the structural features of NrnC and Orn, the other enzyme with such a specific activity . Because NrnC and Orn evolved independently from two different families of RNases, RNase D and RNase T, respectively, their shared substrate-length specificity is particularly noteworthy. On the level of a single subunit, NrnC and Orn contain a conserved catalytic core comprised of a β-sheet and adjacent α-helices, which harbors the residues for divalent cation coordination, the functionally important tyrosine or histidine residue ('DEDDy' or 'DEDDh'), and a wedge residue (L-wedge) that separates the substrate's bases (Figure 8A). In addition, the position of two residues contributing to the respective phosphate caps is conserved, but the identity of the residues varies between the two enzyme groups. For both enzyme families, the active site involves residues from two monomers requiring minimally a dimeric protein (Figure 8B). Thus, the two enzyme families are characterized by a functionally conserved active site optimized to accommodate the shortest RNA substrates, which distinguishes the diribonucleotidases characterized to date from other RNases that act on longer substrates.

Despite these remarkable, function-defining commonalities between NrnC and Orn/Rexo2, their different evolutionary histories have led to distinct implementation of some of these important features. Particularly, NrnC and Orn differ in secondary structure motifs at the periphery of their conserved catalytic core. The enzyme family-specific parts include, for NrnC, additional phosphate cap residues, an N-terminal β-strand and an additional C-terminal α-helix; and for Orn, short helices La, Lb, and Lc forming a lobe that coordinates the 3' base of the dinucleotide substrate (Figures 8A). On the quaternary-structure level, both NrnC and Orn form dimers (Figure 8B). However, the specific dimer arrangements vary between NrnC and Orn. The C-terminal α-helix of NrnC, which is absent in Orn, serves as the major dimerization interface in this family of enzymes, yielding a 2-fold-symmetric dimer. Orn and Orn-related enzymes (such as Rexo2) also form a 2-fold symmetric unit, but via a distinct rotation axis and the interface involves elements of the central core fold (Figure 8B). In contrast to Orn/Rexo2, whose biological unit is the dimeric form, four dimers of NrnC assemble further to the final octameric assembly. Lateral interfaces between dimers of NrnC within the octamer replace the 3' lobe of Orn in coordinating the 3' base of the substrate. Despite these major differences in the architectural composition and unique features of the two enzyme families, the elements that we identified as characteristic for diribonucleotidases align nearly perfectly in space (Figures 8B and 8C).

Of the NrnC-specific motifs, the N-terminal β-strand and near-C-terminal α-helix are conserved in RNase D (Figure 1-figure supplement 2B). NrnC's C-terminal extension including H^{205}, placed in the active site of an adjacent subunit within a NrnC dimer, is involved in 5' base stacking and completes the phosphate cap. Mutations in NrnC's unique phosphate cap retained residual activity, suggesting a role mainly in restricting the length of substrates accommodated at the active site (Figure 2). In contrast, corresponding mutations at the phosphate cap of Orn abolished catalytic activity, suggesting these residues contribute directly to the catalytic mechanism in addition to imposing a substrate-length restriction at the active site . One possibility is that the more ancient Orn evolved more stringent diribonucleotide preference compared to the more recently occurring NrnC-type enzymes. Taken together, the phosphate cap is a unifying feature of Orn and NrnC, though arisen independently in the two enzyme classes. In general, evolution of such specialized active sites that only degrade diribonucleotides would also allow for rapid turnover of this specific RNA pool since competition from longer RNAs would be suppressed. Whether similar motifs have evolved in other enzymes to restrict substrate length remains to be established.

Together, our bioinformatics analysis revealed the unique histories of NrnC and Orn, two nucleases that arose independently to fulfill the crucial role of diribonucleotide degradation. Their activities are at the confluence of RNA metabolism and bacterial second messenger signaling, signifying their importance for cellular homeostasis and regulation. Indeed, functional NrnC or Orn, and consequently the clearance of the cellular dirbonucleotide pool, are necessary for proper growth, being essential in many organisms , which could provide an avenue for targeted antimicrobial intervention. Specifically, as *Bartonella* and *Brucella* species are important pathogens , understanding the implications of NrnC function and failure could offer insight and effective strategies to battle the pernicious impacts of these organisms. Our structure-function studies present blueprints for such endeavors by revealing the specific active-site architectures and activity profiles of NrnC and Orn/Rexo2, their similarities and differences, as well as the general features that distinguish diribonucleotidases from other 3'-5' exoribonulceases.

Material and Methods

Expression constructs and mutagenesis. For protein expression in *E. coli*, codon optimized NrnC genes from *Bartonella henselae* (BH02530) and *Brucella melitensis* (BMEI1828) were synthesized by Geneart (Life Technologies). Genes were cloned into a modified pET28a vector (Novagen) between BamHI and NotI sites using InFusion cloning (Takara Bio). The resulting fusion proteins expressed from these plasmids contained an N-terminal His$_6$-tagged small ubiquitin-like modifier (SUMO) cleavable by recombinant Ulp-1 protease.

For the arabinose-inducible expression and detection of NrnC in *P. aeruginosa*, we used a modified pJN105 vector . The vector pJGHA was constructed by inserting a coding sequence for monomeric superfolder GFP (msfGFP)-HA epitope fusion between the NheI and XbaI sites in pJN105. The plasmid allows insertion of genes of interest between novel NdeI and EcoRI sites, and their expression results in proteins with C-terminal msfGFP-HA. The vector pJHA was constructed by digesting pJGHA with HindIII (New England Biolabs) to remove the msfGFP coding sequence. Following re-ligation of the gel-purified restriction digest, the coding sequence for the HA epitope remained, allowing proteins of interest to be expressed with a C-terminal HA epitope. For expression in *P. aeruginosa*, the codon-optimized NrnC$_{Bh}$ sequences were amplified by PCR and inserted between NdeI and EcoRI sites of the pJHA vector using InFusion cloning.

A QuikChange II site-directed mutagenesis kit (Agilent) was used for the introduction of point mutations in *nrnc*$_{Bh}$ following the manufacturer's instructions. All mutations were verified by DNA sequencing.

Protein expression and purification. *E. coli* BL21 T7 Express cells (New England Biolabs) were transformed with pET28a plasmids encoding His$_6$-SUMO-NrnC$_{Bh}$ or -NrnC$_{Bm}$ and grown in Terrific Broth (TB) supplemented with 50 µg/ml kanamycin at 37°C to an OD$_{600}$ of ~1.0. Induction was carried out at 18°C with 0.5 mM IPTG for 16 hours. Cells were harvested by centrifugation, resuspended in minimal volume of Ni-NTA binding buffer (25 mM Tris-Cl pH 8.5, 500 mM NaCl, 20 mM imidazole), frozen in liquid nitrogen, and stored at -80 °C.

Cell pellets were thawed followed by lysis through sonication and centrifugation. Supernatants were incubated on ice with Ni-NTA resin (Qiagen) equilibrated with Ni-NTA binding buffer for 1 hour with gentle agitation. The NrnC-bound resin was washed three times with 10 column volumes of Ni-NTA binding buffer by gravity flow. Bound NrnC was eluted in 6 column volumes of Ni-NTA elution buffer (25 mM Tris-Cl pH 8.5, 500 mM NaCl, 400 mM imidazole). Eluates were buffer exchanged into gel filtration buffer (25 mM Tris-Cl, 150 mM NaCl, pH 7.5) via a HiPrep 26/10 desalting column (GE Healthcare) and incubated overnight with Ulp-1-His$_6$ to cleave the His$_6$-tagged SUMO moiety from NrnC. Following Ulp-1 cleavage, untagged NrnC protein was recovered in the flow through of a HisTrap Ni-NTA column (GE Healthcare), separated from His$_6$-SUMO, and ULP1-His$_6$. EDTA at a final concentration of 10 mM was added to NrnC before concentration via an Amicon Ultra 10K concentrator (Merck Millipore). Concentrated NrnC was injected onto a HiLoad 16/60 Superdex 200 gel filtration column (GE Healthcare) equilibrated in gel filtration buffer. Fractions containing NrnC were concentrated, frozen in liquid nitrogen, and stored at -80°C.

Crystallization, data collection and structure refinement. NrnC-RNA complexes (pGG and pAA from Biolog Life Science Institute, other nucleotides from Dharmacon) were formed prior to crystallization by mixing 1:2 molar ratio of protein:nucleotide in gel filtration buffer, followed by 30 min incubation at the crystallization temperature. Protein concentrations used in crystallization ranged from 2.5 – 10 mg/ml (NrnC$_{Bh}$) and 1.0 – 8.0 mg/ml (NrnC$_{Bm}$). Crystals were grown via hanging-drop vapor diffusion by mixing equal volumes (0.8 µl) of sample with reservoir solution. NrnC$_{Bh}$ crystals grew at 20°C over a reservoir solution containing 0.1 M succinic acid (pH 6.5), 15 – 20% PEG 3350, and 20% xylitol. NrnC$_{Bm}$ crystals grew at 4°C and 20°C, in reservoir solutions containing 0.1 M Tris-Cl (pH 7), 2.0 - 2.4 M ammonium sulfate or 1.4 M sodium-potassium phosphate, and 20% xylitol. Prior to freezing crystals in liquid nitrogen, they were soaked in reservoir solution with up to 25% xylitol. Data were collected by synchrotron radiation on frozen crystals at Cornell High Energy Synchrotron Source (CHESS) and NE-CAT 24ID-C and 24ID-E beamlines at the Advanced Photon Source (APS) at Argonne National Laboratory. Diffraction data sets were processed using XDS, Pointless and Scala, . The initial structures were solved by molecular replacement using the software package Phenix and MrBUMP in the ccp4 suite with the coordinates of *E. coli* RNase D (PDB:1yt3,) as the search model. Manual model building and refinement were carried out with Coot and Phenix. Illustrations were prepared in Pymol (Version 2.0 Schrödinger, LLC). All software packages were accessed through SBGrid . All data collection and refinement statistics are summarized in Table S1.

Structure determination by cryo-EM. Purified NrnC$_{Bh}$ was diluted in buffer (25 mM Tris-Cl pH 7.5, 150 mM NaCl) to 4, 5, or 6 mg/ml (for incubation with pGG, pAGG, and pAAAGG, respectively). RNA substrate was added at 3-fold molar excess. After 15 min at room temperature, NP40 was added at 0.01% v/v and samples were placed on ice for an additional 15 min. Alternatively, NrnC was diluted to 7.5 mg/ml and incubated with 3-fold excess pAAAGG; after 15 min, 1/5 volume of buffer with CaCl$_2$ and NP40 (25 mM Tris-Cl pH 7.5, 150 mM NaCl, 5 mM CaCl$_2$, 0.05% NP40) was added and incubations proceeded for another 15 min on ice. All cryo-EM samples were prepared with Quantifoil R1.2/1.3 300-mesh grids after glow discharging in a PELCO easiGlow (60 sec glow, 10 mA current, 80% Ar/20% O$_2$) using a FEI Mark IV

Vitrobot (4 °C, 100% humidity, 2.5 sec blot) to plunge grids into liquid nitrogen-cooled ethane immediately after blotting. Data were collected using the Cornell CCMR facility Thermo Fisher Scientific Talos Arctica with a Gatan K3 detector and BioQuantum energy filter operated at 200 kV at a nominal magnification of 63kX (1.24 Å/pixel), 20 eV slit width, and 0.5X binning (super-resolution). Movies were collected with a total dose of 50 e/Å^2, fractionated into either 50 or 75 frames.

Data processing was performed using RELION 3.1 and cryoSPARC . Super-resolution movie exposures were aligned, dose-weighted, and Fourier-cropped to the physical pixel size in RELION using MotionCor2 , and defocus values were estimated using GCTF . Micrographs were then imported into cryoSPARC for manual curation, particle picking and classification. Particles were picked using the cryoSPARC "blob" and template picking and initially extracted with Fourier-cropping to a nominal pixel size of 2.89 Å. This particle stack was cleaned with 2D and 3D classifications in cryoSPARC, then re-extracted in RELION (1.24 Å/pixel) for 3D refinement, CTF refinement, and Bayesian polishing. Further 2D and 3D classification of CTF-refined particles in RELION was performed for the pAGG dataset. For each dataset, no symmetry was imposed during processing until a final refinement was performed imposing D4 symmetry. The crystal structure of NrnC$_{Bh}$ was docked into the reconstructed cryo-EM density maps using the program package Phenix and the models were refined in Coot , ChimeraX , and the real-space refinement module for cryo-EM in Phenix . Illustrations were prepared with ChimeraX and show the density of the sharpened maps. All data collection and refinement statistics are summarized in Table S2.

Complementation analysis in P. aeruginosa. Deletion of *orn* in *P. aeruginosa* UCBPP-PA14 was performed using two-step allelic exchange as described by Hmelo and colleagues . Briefly, deletion alleles were created by overlap extension PCR, and delivered on a pEX18 suicide plasmid to the *P. aeruginosa* host strain by conjugation with *E. coli* donor strain S17.1 leading to the removal of the gene from the genome. To test for complementation, genes were introduced into the *P. aeruginosa orn* deletion strains by using electroporation . Briefly, *P. aeruginosa* cells were grown overnight, centrifuged, then washed with and resuspended in 300 mM sucrose. Expression plasmids based on the pJHA vector were mixed with 100 µl of resuspended cells and electroporated using a Micropulser (Biorad) followed by recovery in 1 ml of lysogeny broth (LB), shaking at 250 rpm for 1 hour at 37 °C. Recovered cells were plated on LB plates containing 60 µg/ml gentamicin. Individual colonies were used for subsequent experiments.

Drip Assay. The indicated *P. aeruginosa* strains harboring expression plasmids were grown overnight with shaking at 37 °C in LB supplemented with 60 µg/ml gentamicin. The cells were adjusted to CFU = 10,000 in fresh LB medium and applied to LB plates supplemented with 60 µg/ml gentamicin and 0.2% arabinose in 20 µl drops. The plates were inverted allowing the culture to drip down the length of the plate, followed by incubation overnight at 37 °C. The plates were imaged using a Chemidoc MP imager (Biorad) with a 0.2 s exposure time.

The colony-measurer Python program (https://github.com/gwmarrah/colony-measurer) was employed to quantify the size of bacterial colonies by pixel measurement. Images were prepared for size quantification by cropping each lane of a drip plate as an individual 8-bit image. ImageJ and colSizeMeasurer.py were used to determine the background pixel intensity and minimum/maximum colony sizes to be measured, respectively. These values were used to refine the parameters in colSizeAnalyzer.py for accurate measurement.

Immunoprecipitation and Western blot. *P. aeruginosa* strains containing plasmid-borne NrnC$_{Bh}$-HA were grown overnight, followed by dilution to an OD_{600} = 0.1 in fresh LB supplemented with 60 µg/ml gentamicin. Cultures were allowed to grow to an OD_{600} = 0.8. Arabinose was added to a final concentration of 0.2% to induce protein expression for 2 hours at 37 °C. Following induction, cultures were normalized by OD, pelleted, and flash frozen in liquid nitrogen. Cells were resuspended in lysis buffer (150 mM NaCl, 25 mM Tris-Cl pH 7.5) followed by sonication. Anti-HA resin (Sigma) was prewashed with lysis buffer. Resin was added to the cleared lysate and incubated with rotation for 1 hour at 4 °C. Following binding, the HA resin was washed with lysis buffer, boiled in SDS loading buffer, and resolved by SDS-PAGE. Western blot transfer to a PVDF membrane proceeded for 90 minutes at constant 0.25 A, followed by overnight blocking with superblock (ThermoFisher) at 4 °C. Rabbit anti-HA primary antibody (Takara Bio) was diluted to 1:100 in TBS-T and incubated with the membrane for 1 hour at 20°C. Following washes with TBS-T, an HRP-conjugated, goat anti-rabbit antibody (GE Life Sciences) was diluted to 1:5000 in TBS-T and incubated with the membrane for 30 minutes at 20 °C. The membrane was washed with TBS-T before treating with SuperSignal West Femto (ThermoFisher) ECL reagent, followed by imaging with a BioRad Chemidoc system.

Size-exclusion chromatography-coupled multiangle light scattering (SEC-MALS). Purified, wild-type or mutant variant NrnC$_{Bh}$ at 2 mg/ml (85 µM) was injected onto a Superdex 200 Increase 10/300 gel filtration column (GE Healthcare) equilibrated with gel filtration buffer (25 mM Tris-Cl, pH 7.5, 150 mM NaCl). Size-exclusion chromatography was coupled to an in-line, static 18-angle light scattering detector (DAWN HELEOS-II, Wyatt Technology) and a refractive index detector (Optilab T-rEX, Wyatt Technology). Data were collected every second. Data analysis was performed with Astra 6.1 (Wyatt Technology) yielding the molar mass and mass distribution (polydispersity) of the sample. Monomeric BSA (Sigma; 5 mg/ml) was used as a control sample and to normalize the light scattering detectors.

Measuring dissociation constant (K_d) and binding specificity by DRaCALA. To measure K_d, the purified protein was serially diluted in binding buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl, and 5 mM $CaCl_2$). Each dilution was mixed with ^{32}P-lableled substrate and spotted onto nitrocellulose. The dried nitrocellulose was exposed to a phosphorimager screens, scanned and analyzed as previously described . The fraction bound was plotted against protein concentration using the program Prism. For competition experiments to determine binding specificity, 100 µM of unlabeled nucleotides were mixed with ^{32}P-labeled pGG; the mixtures were added 200 nM of purified NrnC.

Biochemical assay of RNase activity. The reactions were performed by adding the indicated concentration of enzyme to the indicated concentration of substrate containing a tracer, consisting of 5'-end ^{32}P-labeled substrate with the same length and sequence, in reaction buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl, and 5 mM $MgCl_2$). Reactions are stopped at indicated time points by the addition of 0.2 M EDTA. The samples were mixed with loading buffer (4 M Urea, 20% sucrose, 0.1% SDS, 0.05% bromophenol blue, 0.05% xylene cyanole FF, and 1x TBE), and separated by electrophoresis on 20% polyacrylamide gels.

DNase activity measurement. DNase activity was assessed using an unspecific 1.5 kb PCR fragment, either with blunt ends or after restriction digestion with either KpnI, NdeI, or NotI (New England Biolabs). CIP (New England Biolabs) was used to dephosphorylate overhangs, and PNK (New England Biolabs) was used to phosphorylate blunt PCR products. NaCl, $MgCl_2$, and $MnCl_2$ were added to the concentration indicated in the figures, with concentrations of NrnC and DNA at 1 µM and 20 nM, respectively, except where indicated otherwise in the figure. Reactions were incubated for 30 min or the indicated time at 37 °C. Reactions were stopped by the addition of a stop buffer containing proteinase K (Qiagen) and EDTA (JT Baker) to a final concentration of 0.1 mAU proteinase K and 10 mM EDTA, followed by incubation at 50 °C for 30 minutes. Samples were resolved by electrophoresis in a 1% agarose gels containing GelRed stain (Biotium) and imaged by UV visualization in a GelDocXR system (BioRad).

Identification of RNase homologs. For each group of RNases, a list of seed protein sequences was manually prepared with Uniprot entry names: Orn (ORN_PSEAE, ORN_ECOLI, ORN_HUMAN, ORN_STRGR, ORN_CORDI, ORN_BURMA, ORN_YEAST, ORN_VIBCH), NrnA (NRNA_BACSU, NRNA_MYCPN, NRNA_THET8, NRNA_MYCTU, A0A3R9HUU0_STRSA), NrnB (NRNB_BACSU, A0A5C5X7X8_9BACT), NrnC (A9CG28_AGRFC, G4VUY7_9RHIZ, A1UU18_BARBK), RNase T (RNT_ECOLI, RNT_VIBCH, RNT_BUCAP, RNT_HAEIN, RNT_XYLFA), and RNase D (RND_ECOLI, Q9ZD81_RICPR, RND_HAEIN, I6XF17_MYCTU).

Initially a seed alignment was prepared for each RNase family using T-COFFEE in the Expresso structural alignment mode . A search on the 2020_03 release of UniprotKB was performed with an HMM prepared from these alignments with hmmsearch from HMMER 3.3 . Results were filtererd by two criteria, a hit score above 125 and a template/query length ratio between 0.8 and 1.2. Search hits were used to construct a new multiple sequence alignment with MAFFT in E-INS-i mode . The hmmsearch was repeated with an HMM constructed from this expanded alignment and results were filtered with the same hit score cutoff but a more generous template/query length ratio between 0.6 and 1.5. The resulting hits were considered to be the final list for each RNase family, with any sequences found in multiple categories assigned to the category for which it had a higher hit score.

To determine whether a particular taxa contained homologs for each RNase group, the total number of proteins for each taxon in the NCBI taxon database present in our dataset were counted . The average number of proteins per genome for each taxon was determined for all genomes annotated as a descendent taxon available in the NCBI genome database. Finally, the number of homologs found in each taxonomic category was calculated as a fraction of the total proteins in that taxon, multiplied by the average genome size to get an average presence-per-genome. For visualization purposes, values above 0.5 are considered present and non-zero values below 0.5 are potentially or partially present. The minimal species tree was extracted from the NCBI taxonomy database using ETE3 , followed by visualization of the resulting tree and annotation with iTOL .

All commands and code from this analysis were constructed as a reproducible SnakeMake pipeline available at https://github.com/jgoodson/rnases (commit 66b2664 used in current versions of the figures).

Phylogenetics of DnaQ-family RNases. To construct a phylogenetic tree of the DnaQ-fold RNases, the final sequences identified from the previous analysis for each family were clustered by sequence identity with MMSeqs2 . Targeting a final sequence count of 600, the sequence identity threshold was determined for each family necessary to approximately maintain the original proportion of each family in the final representatives (30% for Orn, 45% for NrnC, 50% for RNase T, and 30% for RNase D). From these cluster representatives, a multiple sequence alignment was constructed using MAFFT in E-INS-i mode using DASH to obtain additional structural homologs . The alignment was trimmed by removing the additional DASH sequences and columns with more than 90% gaps. The most appropriate evolutionary model was determined with IQ-TREE ModelFinder (LG+R8) and a phylogenetic tree was constructed using IQ-TREE 2.1.1 with 10,000 UFBoot replicates, and some modified parameters for expanded tree search (additional UFBoot NNI search, initial SPR search radius 20, 500 initial trees, initial search on 100 best trees, maintenance of the 50 best trees, and 500 iterations without improvement as stopping criteria) . The tree was rooted with midpoint rooting on the long internal branch between RNaseT/Orn and RNaseD/NrnC. Sequence logos were created from monophyletic subgroup alignments using Logomaker .

Data deposition. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 7MPL, 7MPM, 7MPN, 7MPO, 7MPP, 7MPQ, 7MPR, 7MPS, 7MPT, 7MPU, 7MQB/EMD-23941, 7MQD/EMD-23943, 7MQF/EMD-23945, 7MQH/EMD-23947, 7MQC/EMD-23942, 7MQE/EMD-23944, 7MQG/EMD-23946, 7MQI/EMD-23948).

Acknowledgement

This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). 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. Additional crystallographic research was conducted at the Center for High Energy X-ray Sciences (CHEXS), Cornell High Energy Synchrotron Source (CHESS), which is supported by the NSF under award DMR-1829070. The MacCHESS resource is supported by NIGMS award 1-P30-GM124166-01A1 and NYSTAR. This work made use of the Cornell Center for Materials Research Shared Facilities, which are supported through the NSF MRSEC program (DMR-1719875). The work was supported by NIH awards R01AI142400 (to. V.T.L.), R01GM123609 (to H.S.), and R35GM136258 (to J.C.F. and B.A.B.).

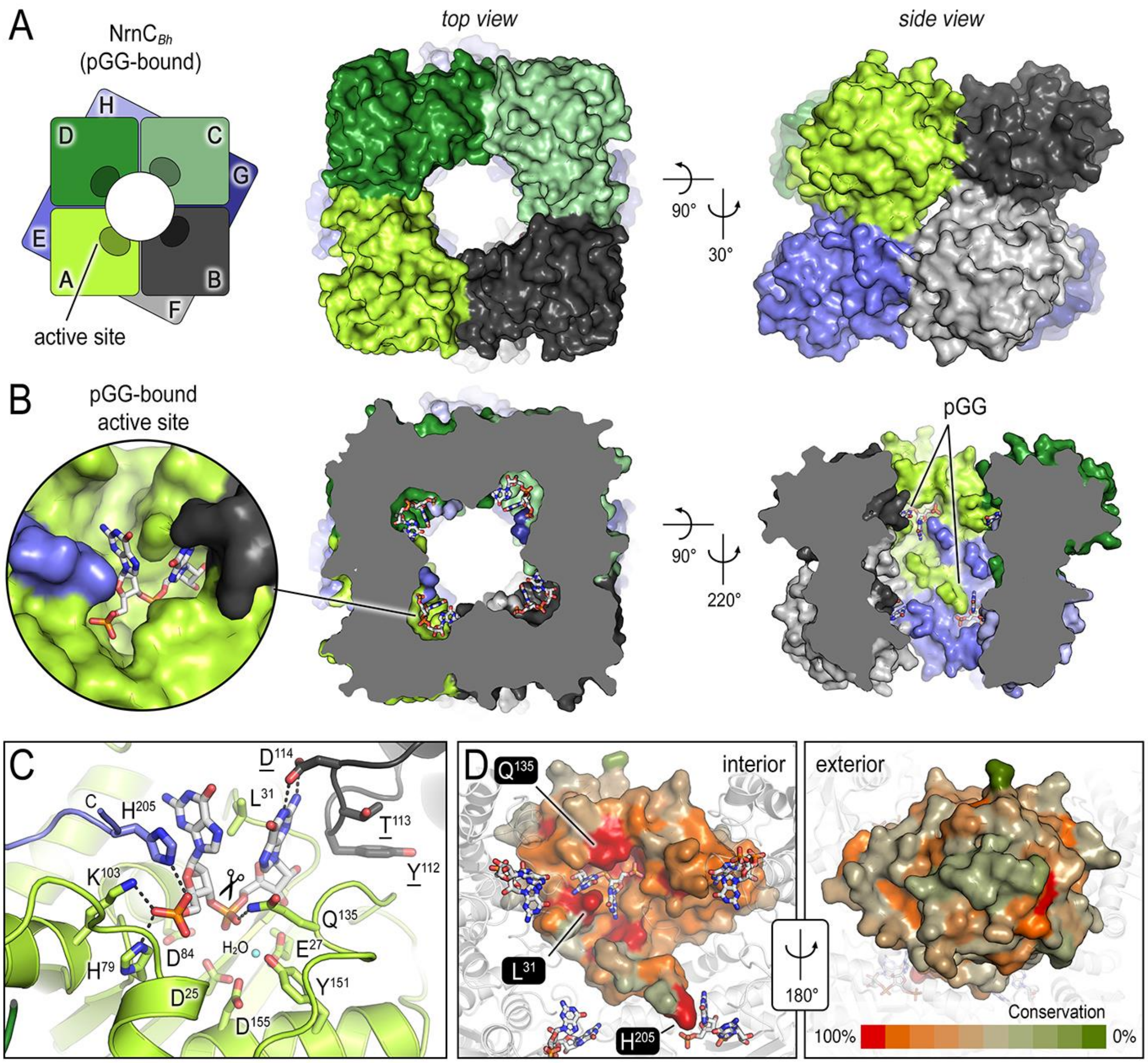


Figure 1. The crystal structures of *B. henselae* NrnC bound to pGG reveals motifs defining substrate specificity. (A) The octameric assembly. NrnC$_{Bh}$ is shown as surface representation in two views. Each monomer is shown in a distinct color. The cartoon illustrates the stacking of the two tetrameric NrnC rings that form the octamer with a central, round opening. (B) Active-site position. Each monomer contributes one active site, here bound to the substrate pGG, facing toward NrnC's central pore. Each active site includes a C-terminal tail of a subunit from an adjacent ring. (C) Substrate coordination. The catalytic DEDDy motif and residues coordinating each moiety of pGG contacts are shown as sticks, with carbon residues colored according to monomer identity. Residue Y^{151} coordinates water molecule near the scissile bond. (D) Conservation mapping on a surface representation of a NrnC monomer. Conservation scores

were calculated based on a multisequence alignments (MUSCLE;) of NrnC homologs identified using a sequence search on the EggNOG resource, version 5.0.0 and the sequence of NrnC$_{Bh}$ as the input. Outliers were identified based on sequence length and non-consensus insertions, resulting in a final collection of 560 sequences of putative NrnC orthologs. The two views, separated by a 180° rotation, show the cavity-facing (interior, left) and outer-facing (exterior, right) surface regions.

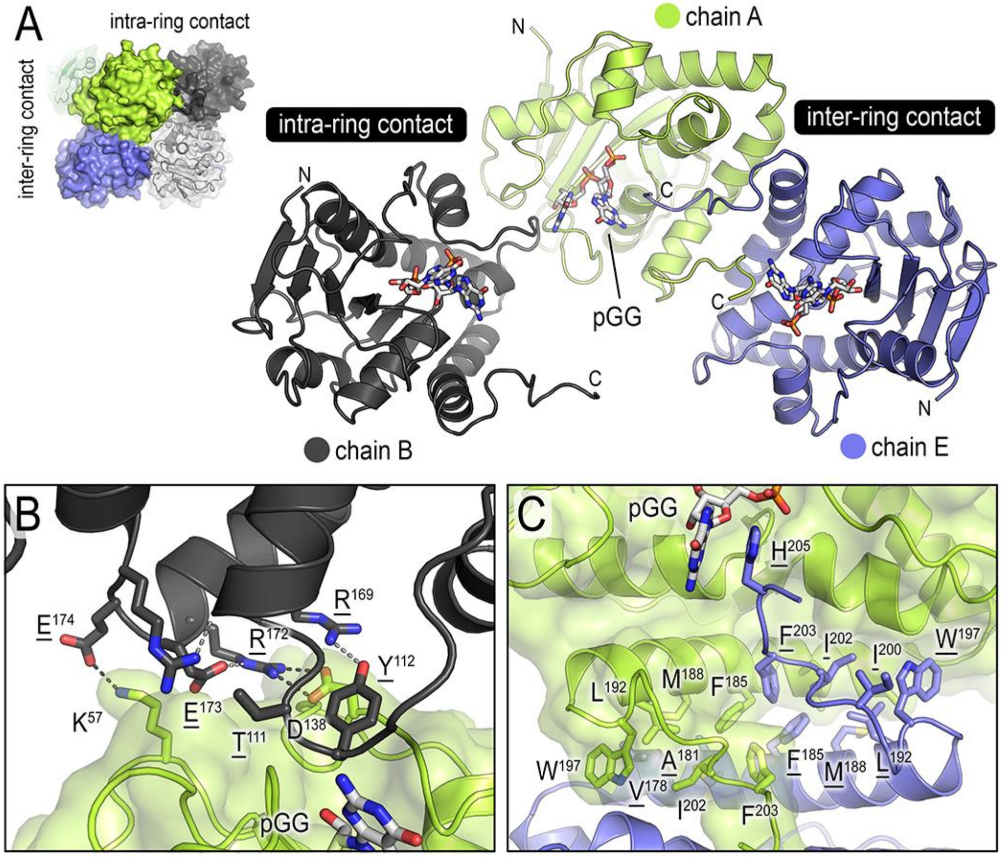


Figure 1-figure supplement 1. Inter- and intra-ring contacts in the NrnC$_{Bh}$ octamer. (A) Overview of three pGG-bound NrnC$_{Bh}$ monomers. The monomers are shown in cartoon representation The chain colored in dark grey forms an intra-ring contact with the central, green-colored chain, whereas the chain colored in blue forms a representative inter-ring contact within the octameric NrnC. (B) Detailed intra-ring interface. (C) Detailed inter-ring interface. Residues contributing direct interactions between monomers are shown as sticks. Representative hydrogen bonds are shown as dashed lines.

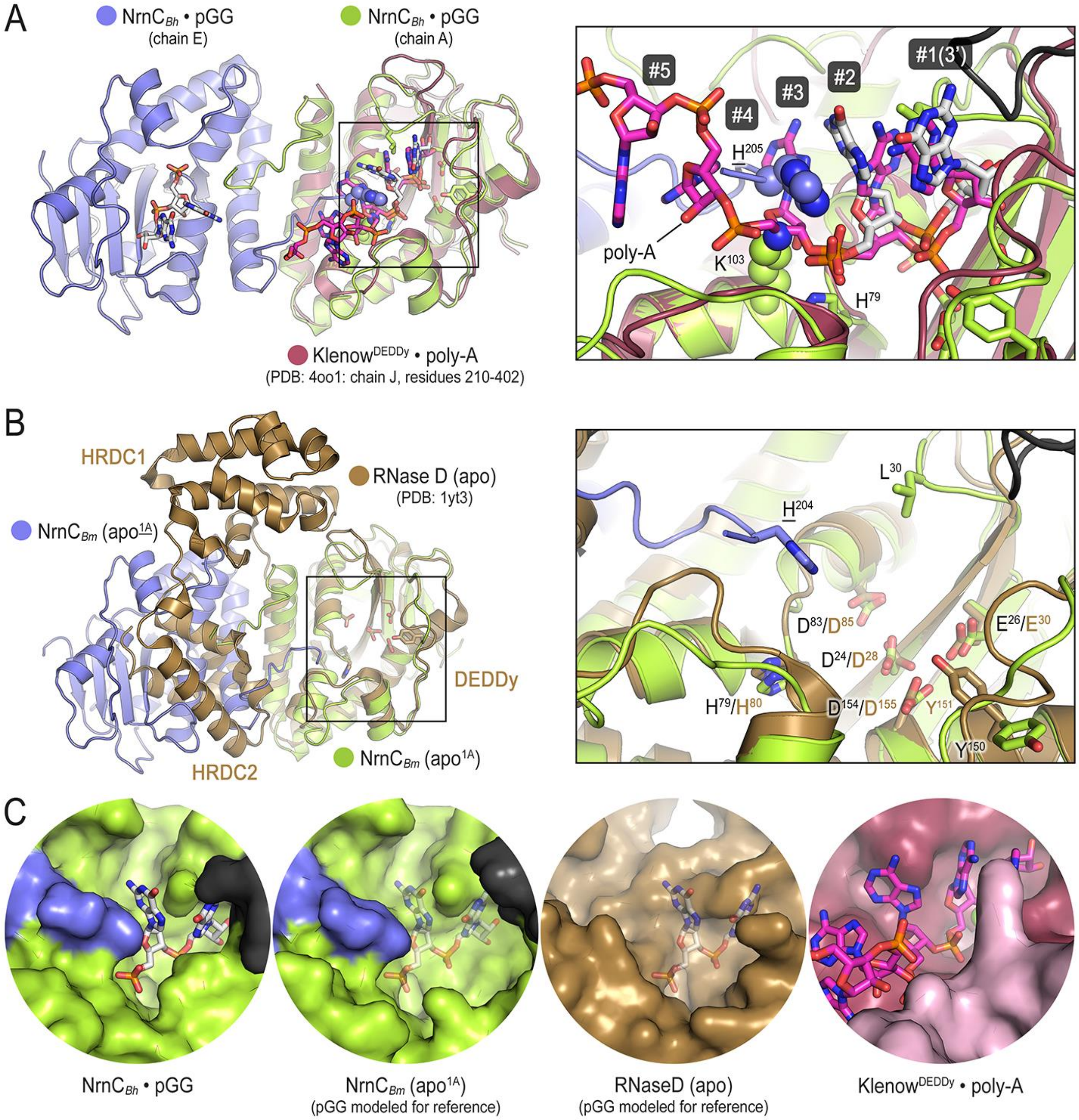


Figure 1-figure supplement 2. Comparison of NrnC to structurally related proteins reveals the constricted nature of NrnC's active site. (A) A pGG-bound NrnC$_{Bh}$ inter-ring dimer (slate and green chains, with pGG carbon atoms shown in white) was superimposed on a structure of Klenow fragment (purple chain) bound to poly-A (pink carbon atoms; PDB 4oo1,). The inset shows a detailed view of the superimposed active sites. Numbering refers to the residues in the

substrate. Phosphate cap residues of NrnC that block the path of poly-A substrate are shown as spheres. The black protein chain in top-right corner stems from an adjacent intra-ring monomer. (B) A substrate-free NrnC$_{Bm}$ inter-ring dimer (slate and green chains) was superimposed on a structure of apo-RNase D , highlighting conservation of the catalytic DEDDy motif and differences in regions around NrnC's L-wedge and phosphate cap (inset). (C) Surface views of the active sites of NrnC, RNaseD, and Klenow$^{\text{DEDDy}}$ accentuate the constraint of the NrnC active site. Translucent pGG represents modeled substrate as opposed to co-crystallized substrate.

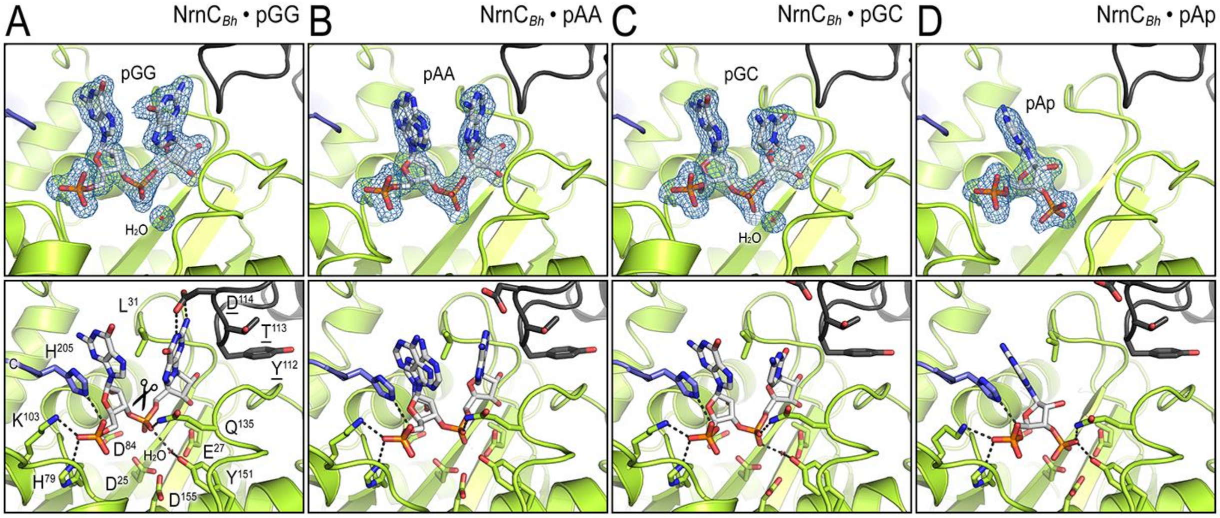


Figure 1-figure supplement 3. **Structural comparison of NrnC$_{Bh}$ bound to various ribonucleotides.** (A) Diribonucleotide pGG-bound NrnC$_{Bh}$, identical to the structure shown in Figure 1. (B) Diribonucleotide pAA-bound NrnC$_{Bh}$. (C) Diribonucleotide pGC-bound NrnC$_{Bh}$. (D) Adenosine-3',5'-bisphosphate (pAp)-bound NrnC$_{Bh}$. The top row of images shows NrnC as a cartoon representation with the nucleotide substrate represented as sticks with carbon atoms colored white. Polder omit maps for each substrate are shown as blue mesh. The bottom row shows detailed views of the active site residues contacting each ribonucleotide.

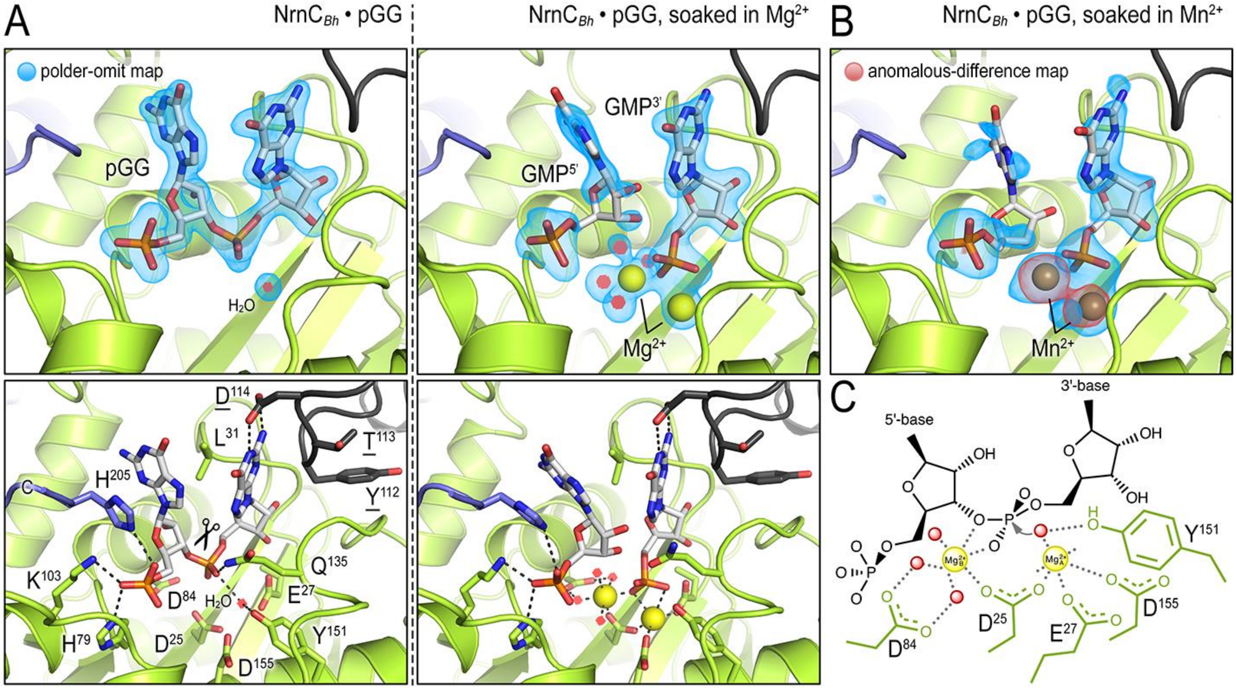


Figure 1-figure supplement 4. *In-crystallo* catalysis indicates a two-metal mechanism of NrnC$_{Bh}$ activity. (A) Active sites of NrnC$_{Bh}$•pGG, before and after soaking crystals in a solution containing Mg^{2+} prior to data collection. (B) Active site of NrnC$_{Bh}$•pGG, after soaking crystals in a solution containing Mn^{2+} prior to data collection. Top panels show polder omit maps highlighting nucleotide and metal density. The red density in (B) represents an anomalous-difference map calculated from data collected at the Mn^{2+} absorption edge. The bottom panels show specific active site contacts between protein, nucleotide, ions, and water molecules. (C) Schematic overview of two-metal coordination at the active site of NrnC.

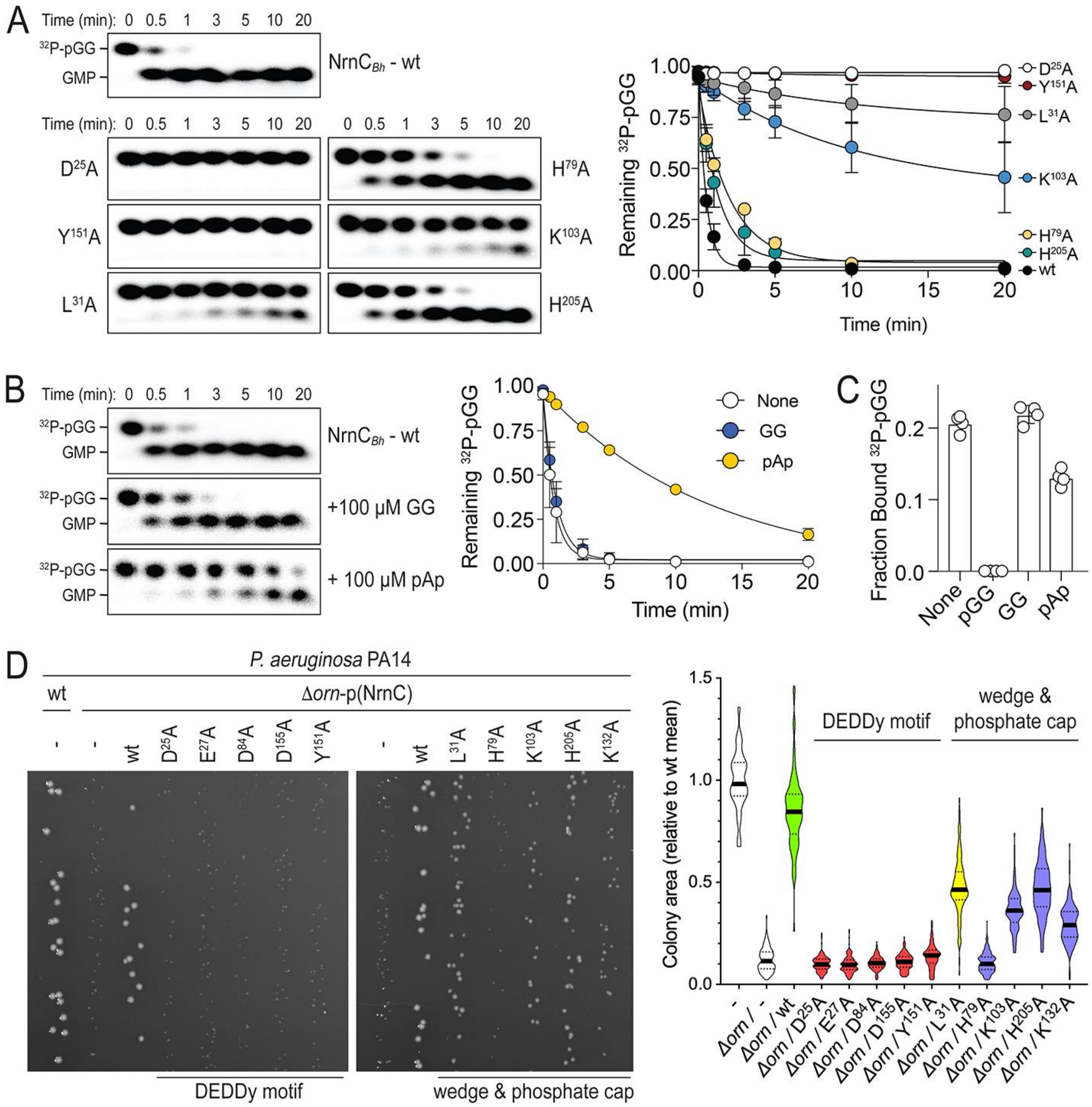


Figure 2. Phosphate cap and L-wedge contribute to NrnC's diribonucleotidase activity. (A) In vitro enzyme activity. Degradation of ^{32}P-pGG (1 μM total) by purified wild-type NrnC$_{Bh}$ or variants with alanine substitutions (5 nM) at the indicated sites was assessed. Samples were stopped at the indicated times (min) and analyzed by denaturing 20% PAGE. Representative gels are shown (left). The graph (right) shows the means and SD of three independent experiments. (B) Effect of a dinucleotide lacking the 5' phosphate (GG) and pAp on NrnC catalysis. pGG processing was assessed as in (A) but in the presence or absence of 100-fold excess (over ^{32}P-pGG) GpG or pAp. Representative gels (left) and quantification from three independent

experiments (right) are shown. Means and SD are plotted. (C) Competition binding studies. Fraction bound of ^{32}P-pGpG to 200 nM purified NrnC$_{Bh}$ in presence of no competitor, 100 µM pGG, 100 µM GpG, or 100 µM pAp is plotted as individual data, means, and SD of four independent experiments. (D) Complementation of the small-colony phenotype of *P. aeruginosa* Δ*orn* by wild-type and mutant NrnC$_{Bh}$. Bacterial cultures were diluted and dripped on LB agar plates. After overnight incubation, representative images of the plates were taken (left). Experiments were performed in triplicate. Quantification of respective colony sizes are shown as violin plots (right).

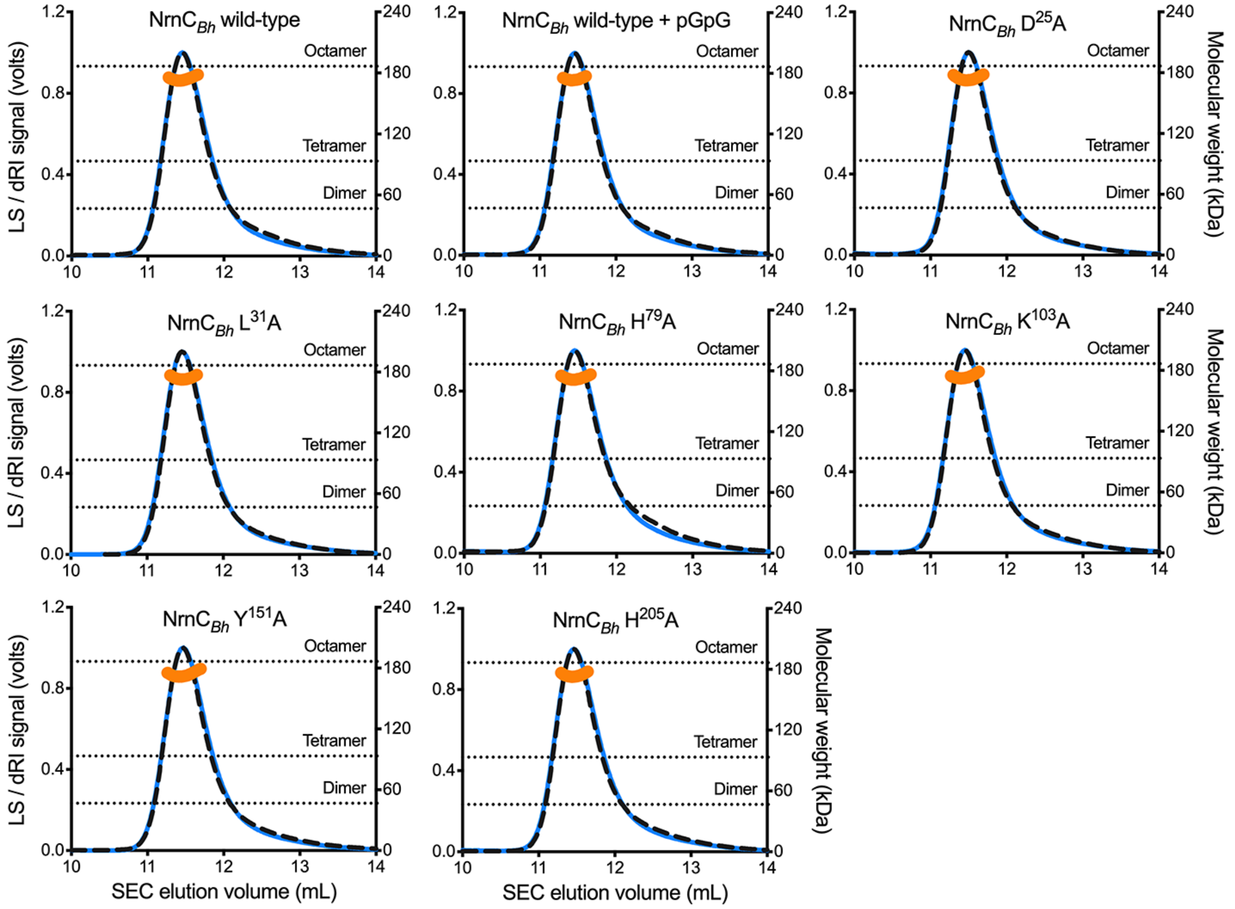


Figure 2-figure supplement 1. **SEC-MALS of NrnC$_{Bh}$ wild-type and mutant variants.** Molecular weight determination indicates that pGG binding does not impact oligomerization, and that purified NrnC$_{Bh}$ point mutants remain octameric in solution. Absolute molecular weights of NrnC are shown as orange data points across elution peaks plotted on the right axis.

Theoretical oligomerization states are shown as dashed horizontal lines. 90°-light scattering: blue solid lines; refractive index signal: black dashed lines; plotted on left axis.

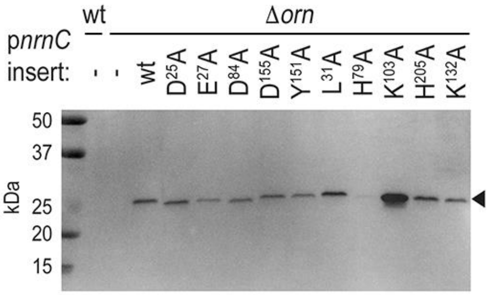


Figure 2-figure supplement 1. Expression of NrnC*Bh* wild-type and mutant variants in *P. aeruginosa* Δ*orn*. Cell lysates were analyzed by Western blotting, detecting the C-terminal HA-tag in recombinantly expressed NrnC*Bh*.

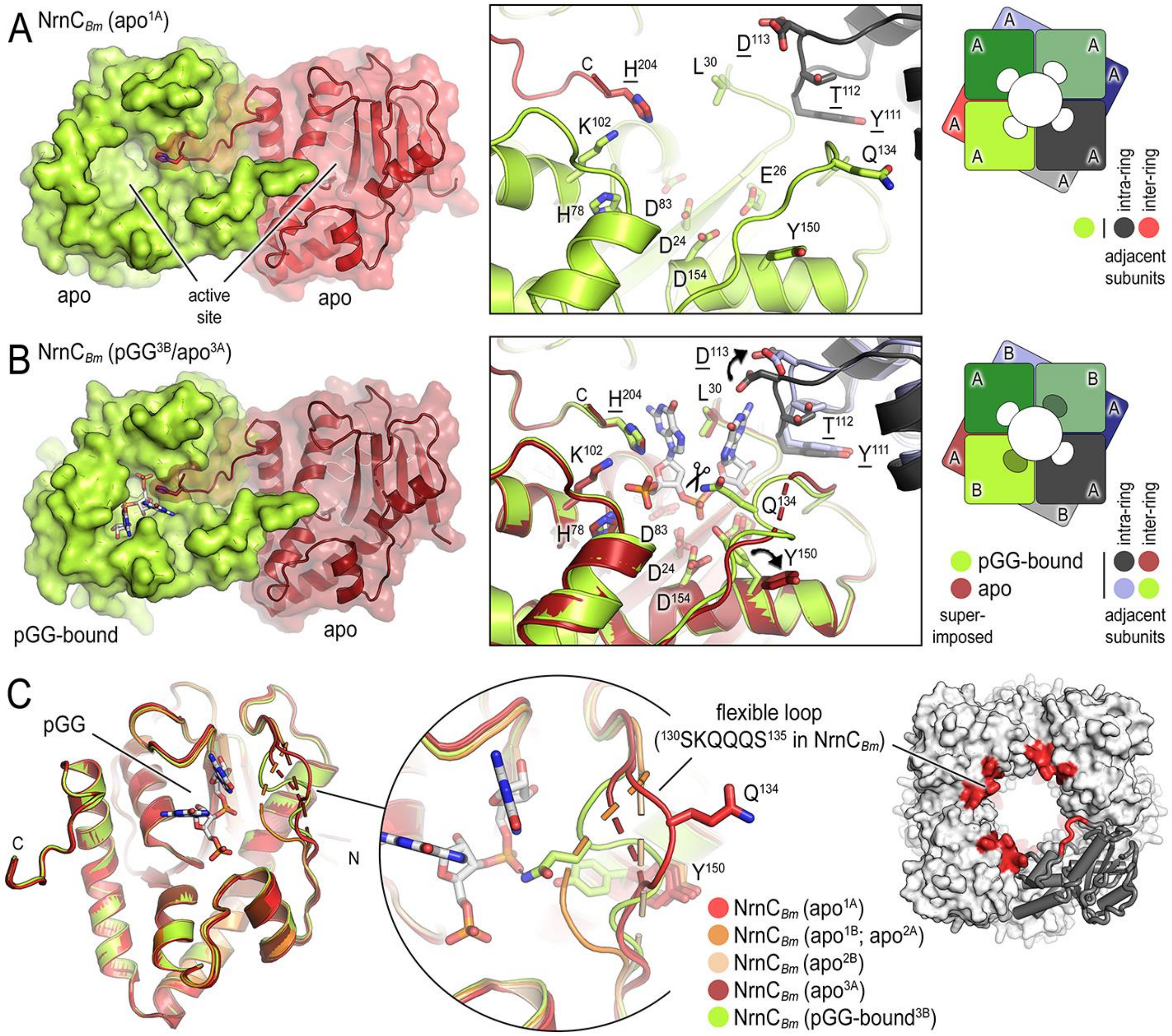


Figure 3. *Brucella melitensis* NrnC crystal structures reveal a flexible loop that constraints the enzyme's active site. (A) Crystal structure of apo-NrnC$_{Bm}$. A crystallographic dimer as part of the octameric assembly is shown as surface presentation (left) and close-up of the active site (middle). The diagram (right) depicts the octamer and the spatial relationship of the monomers shown. (B) Crystal structure of NrnC$_{Bm}$ with alternating substrate-bound and empty active sites. The close-up (middle) shows a superposition of the two monomers in the asymmetric unit, depicting their conformational difference and adjacent monomers, with intra- and inter-ring neighbors colored as shown in the diagram (right). (C) Superposition of four apo-NrnC$_{Bm}$ conformations based on three independent crystal forms, compared to the pGG-bound conformation of the same protein shown in (B). The position of the flexible loop (red) in the NrnC octamer is shown (right panel).

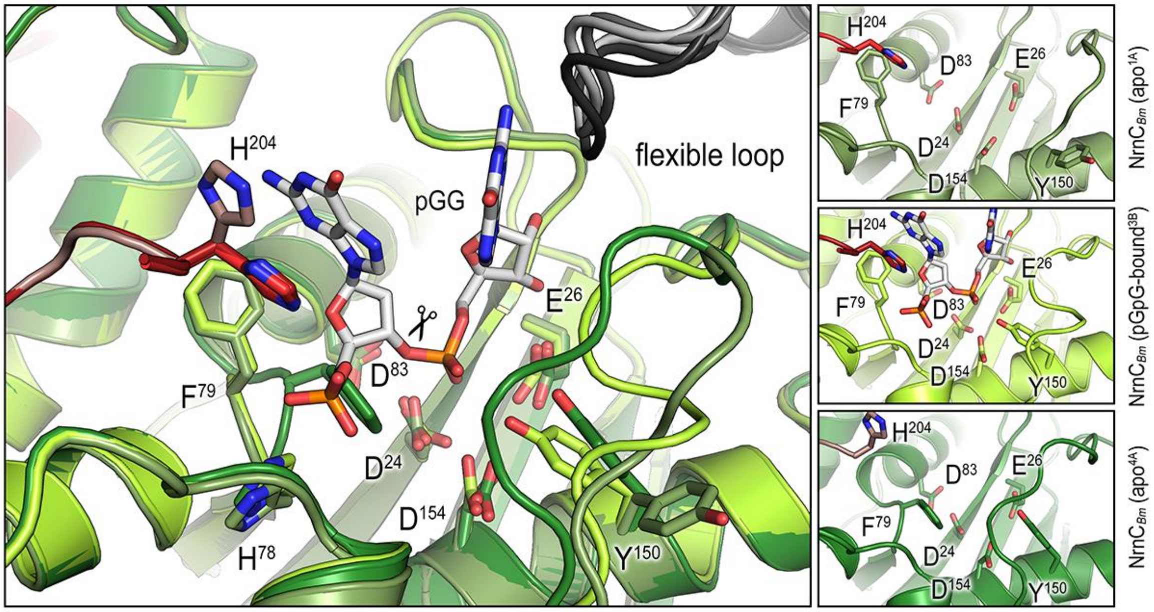


Figure 3-figure supplement 1. Overlay of an alternative crystallographic apo-NrnC$_{Bm}$ state with the apo- and pGG-bound states observed in the crystal structure shown in Figure 3B. The large panel shows three structures superimposed. The smaller panels on the right show each active site isolated with residues of interest labeled and shown as sticks.

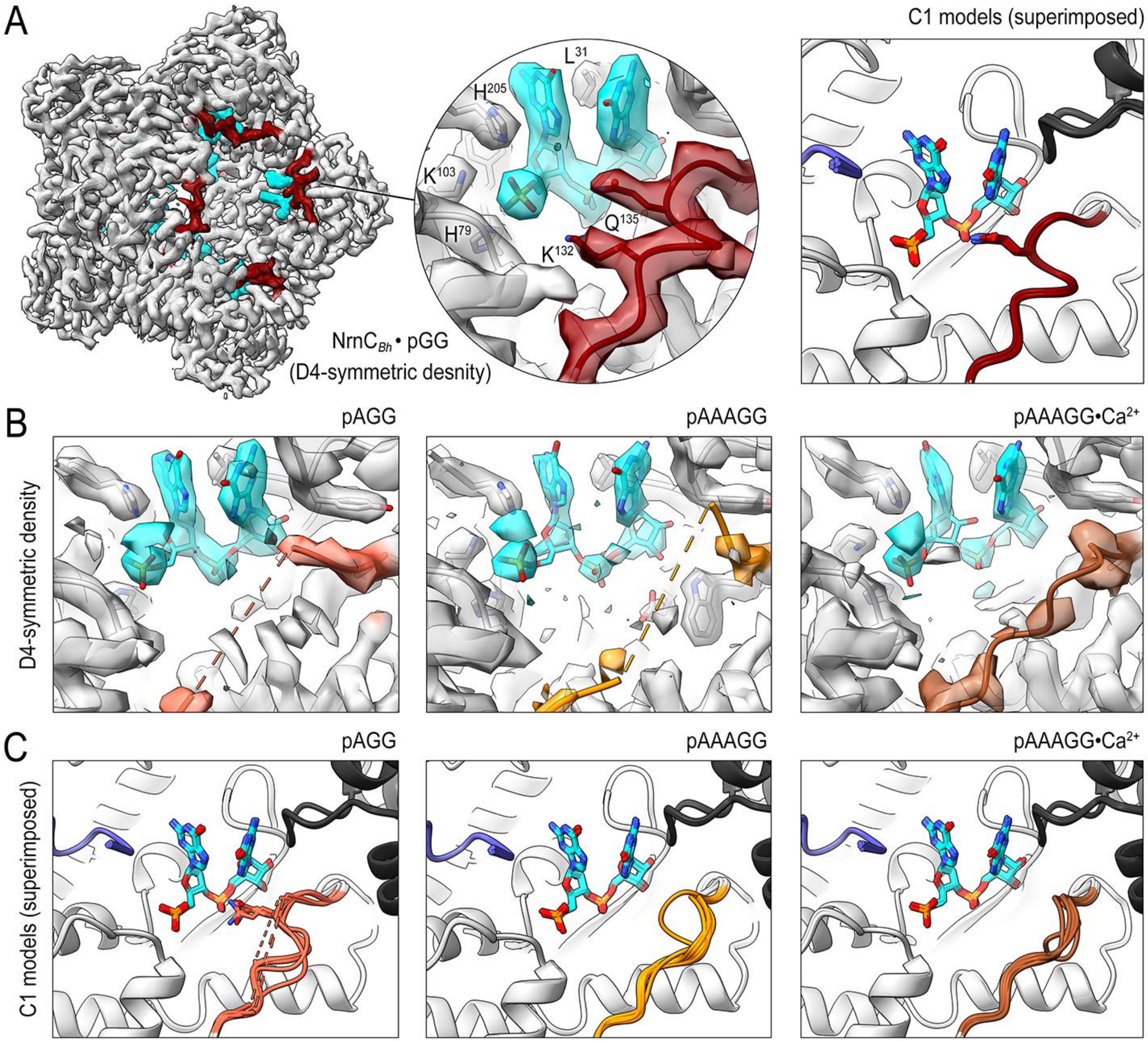


Figure 4. Cryo-EM structures of NrnC$_{Bh}$ with 2-, 3-, 5-mer RNA substrates show substrate length-dependent active site conformations. (A) Electron density map of a NrnC$_{Bh}$ octamer in complex with pGG. D4 symmetry was applied during final map refinement. pGG molecule and density are colored cyan. The SKQQQS-containing loops (residues 130-137) are colored maroon. Superposition of all eight active sites from a reconstruction with C1 symmetry (right panel) shows consensus order in the loop when bound to pGG. (B) Active site images shown for NrnC$_{Bh}$ incubated with 3-mer and 5-mer (with or without Ca^{2+}) RNA substrates. Regions corresponding to those shown in (A) are shown in color. D4-symmetric maps are shown. (C) Superposition of all eight active sites from octamer reconstructions based on respective C1-symmetric maps for each RNA substrate.

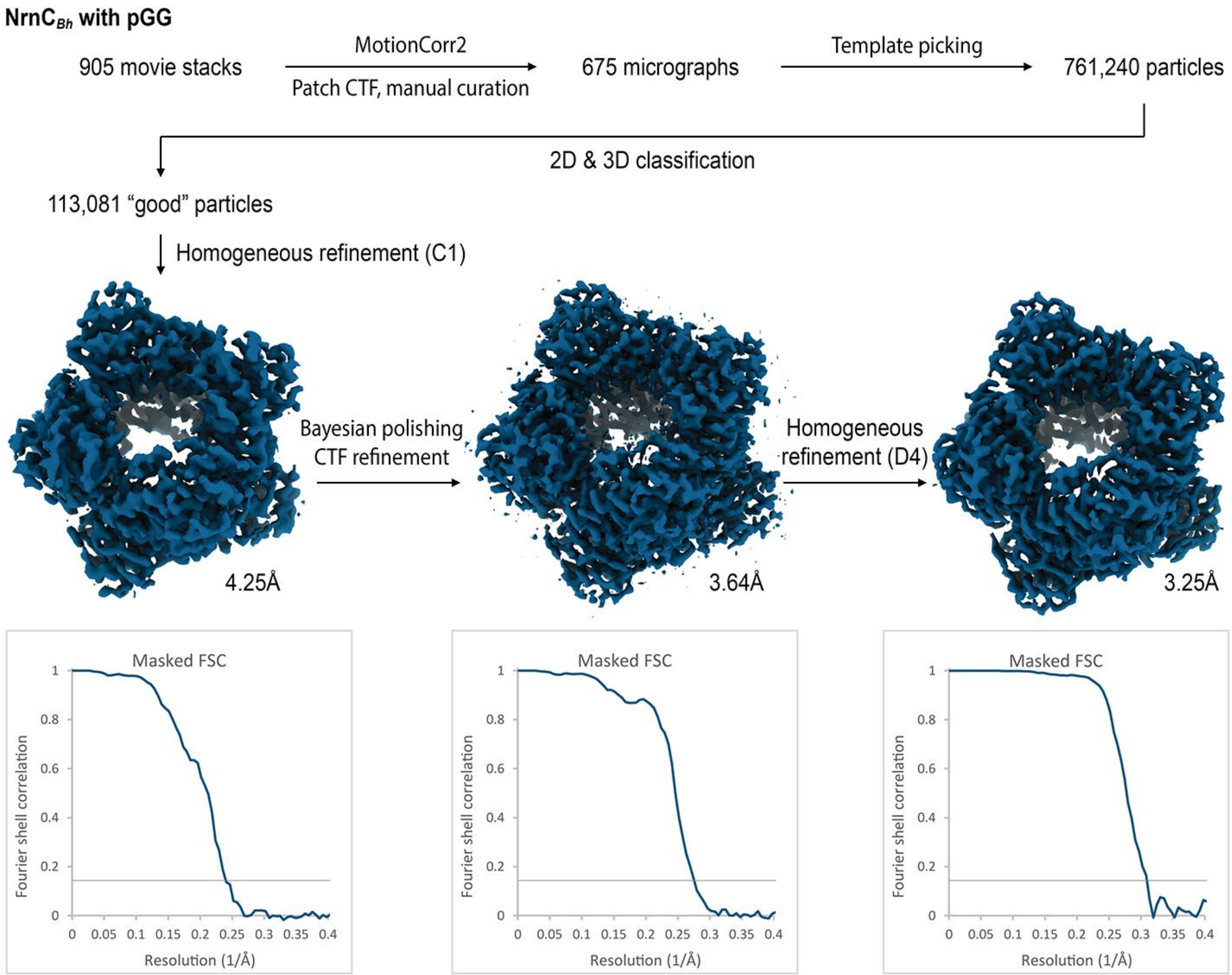


Figure 4-figure supplement 1. Cryo-EM workflow and resolution for NrnC$_{Bh}$•pGG.

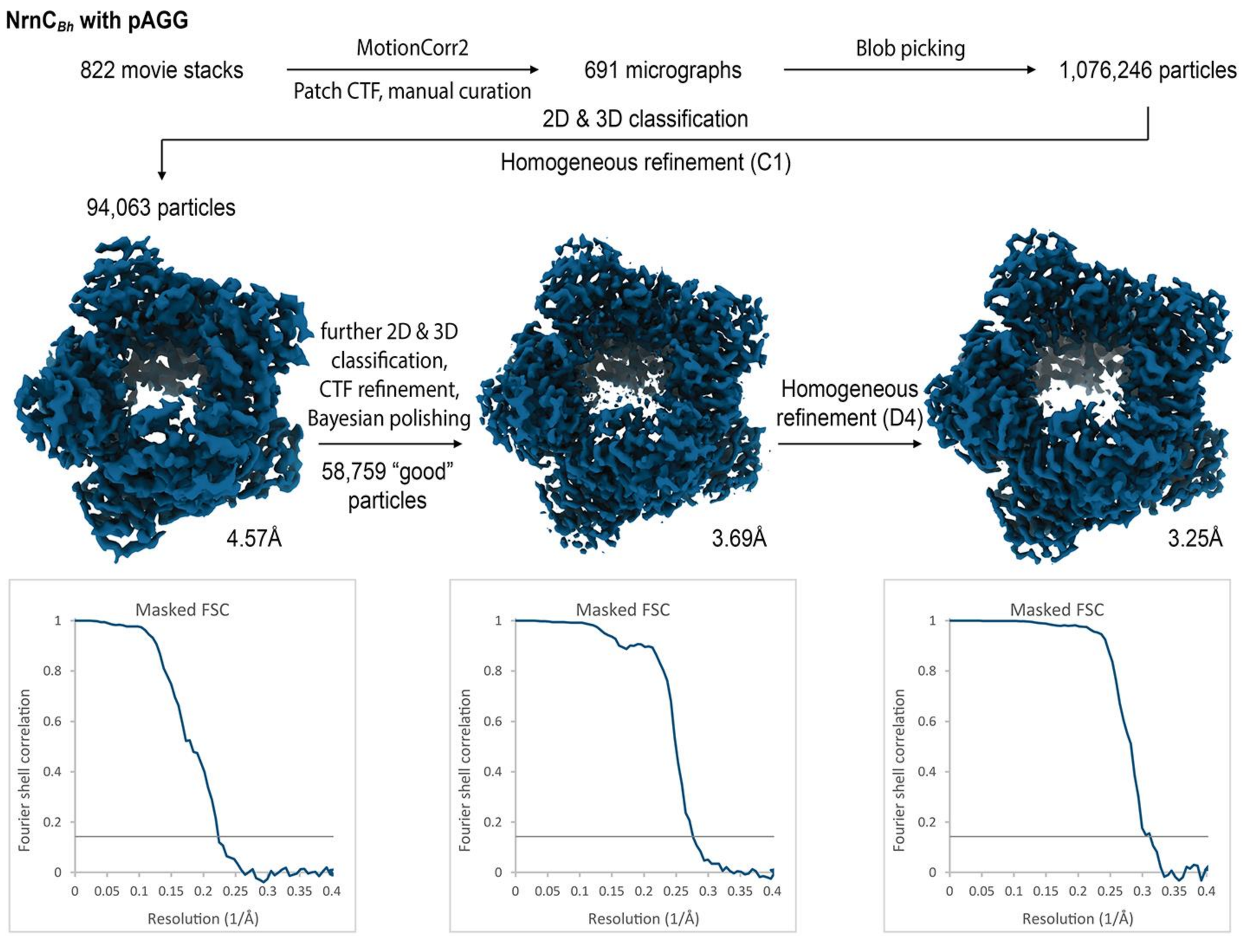


Figure 4-figure supplement 2. Cryo-EM workflow and resolution for NrnC$_{Bh}$•pAGG.

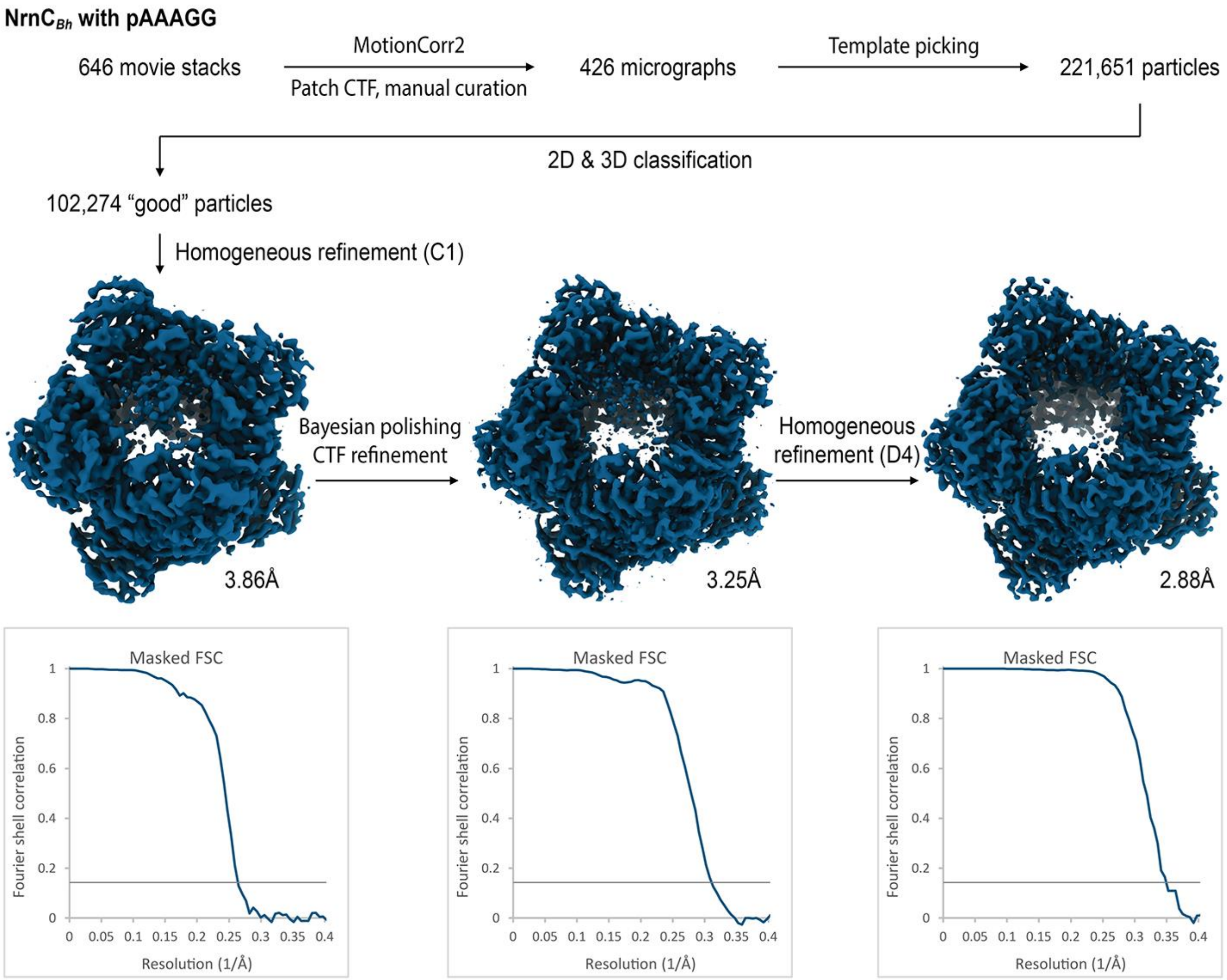


Figure 4-figure supplement 3. Cryo-EM workflow and resolution for NrnC$_{Bh}$•pAAAGG.

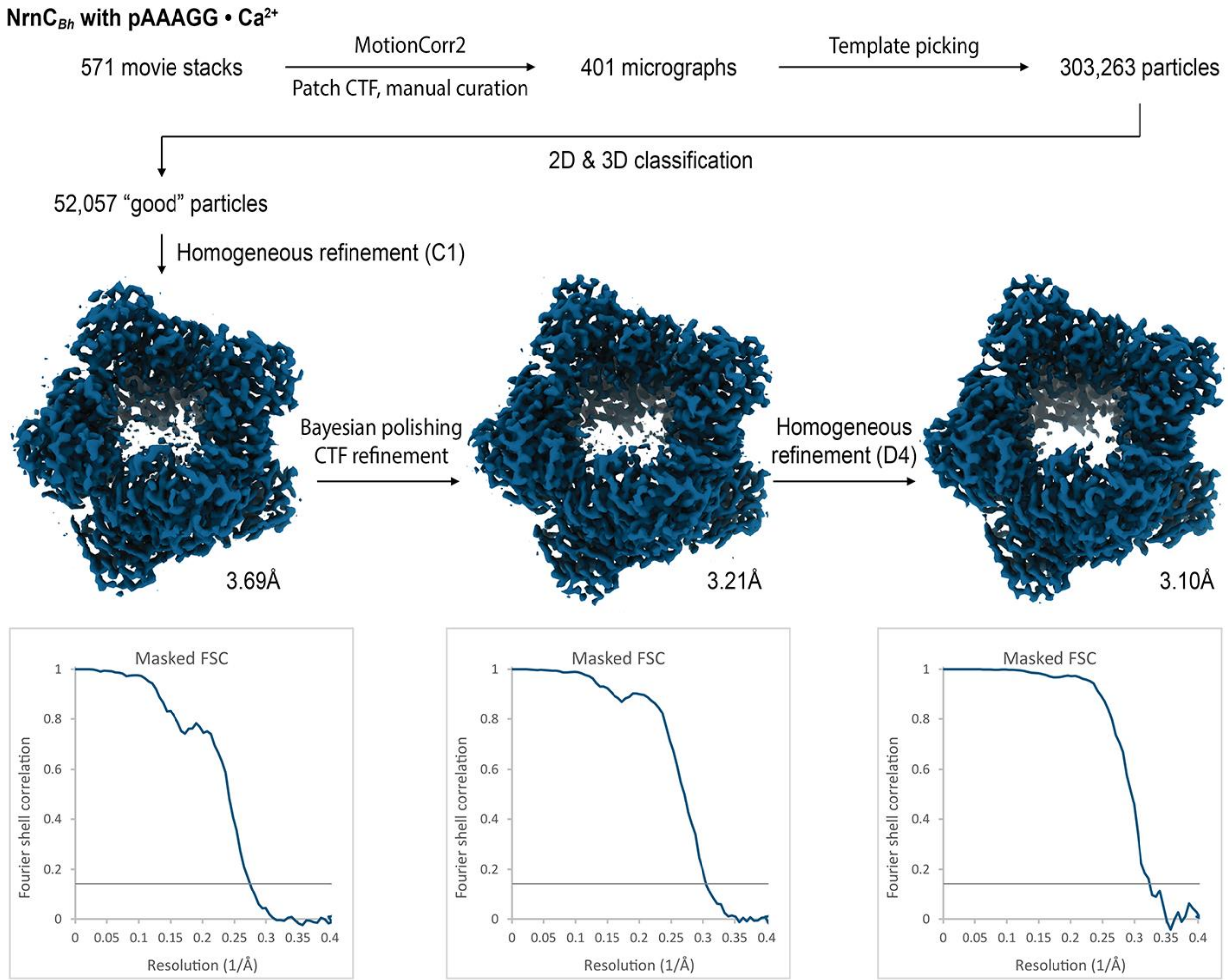
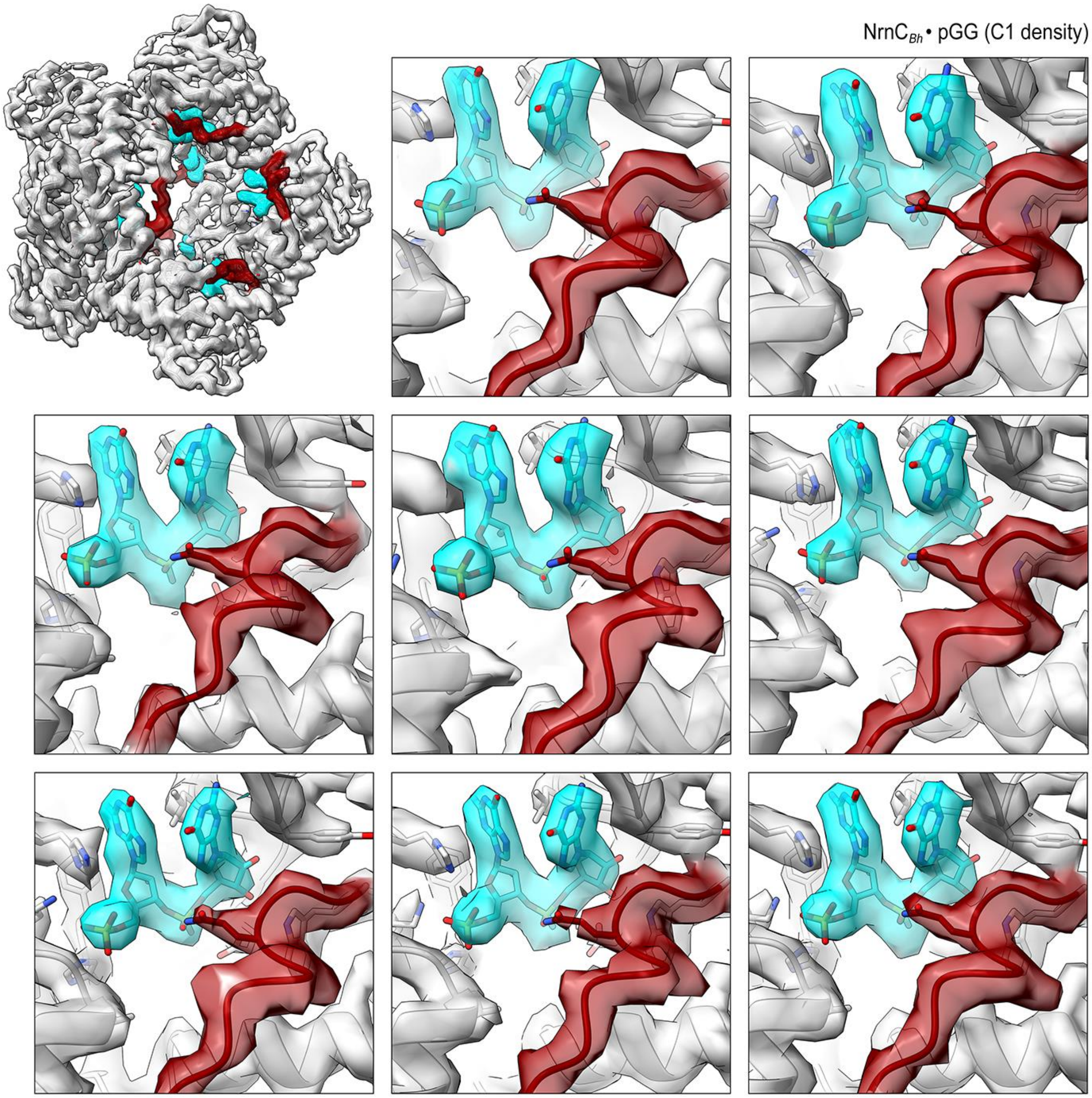


Figure 4-figure supplement 4. Cryo-EM workflow and resolution for NrnC$_{Bh}$•pAAAGG in the presence of Ca^{2+} ions.



833

834**Figure 4-figure supplement 5. Overall and individual active site electron density of a**
835**NrnC$_{Bh}$•pGG octamer after refinement with C1 symmetry.**

836

837

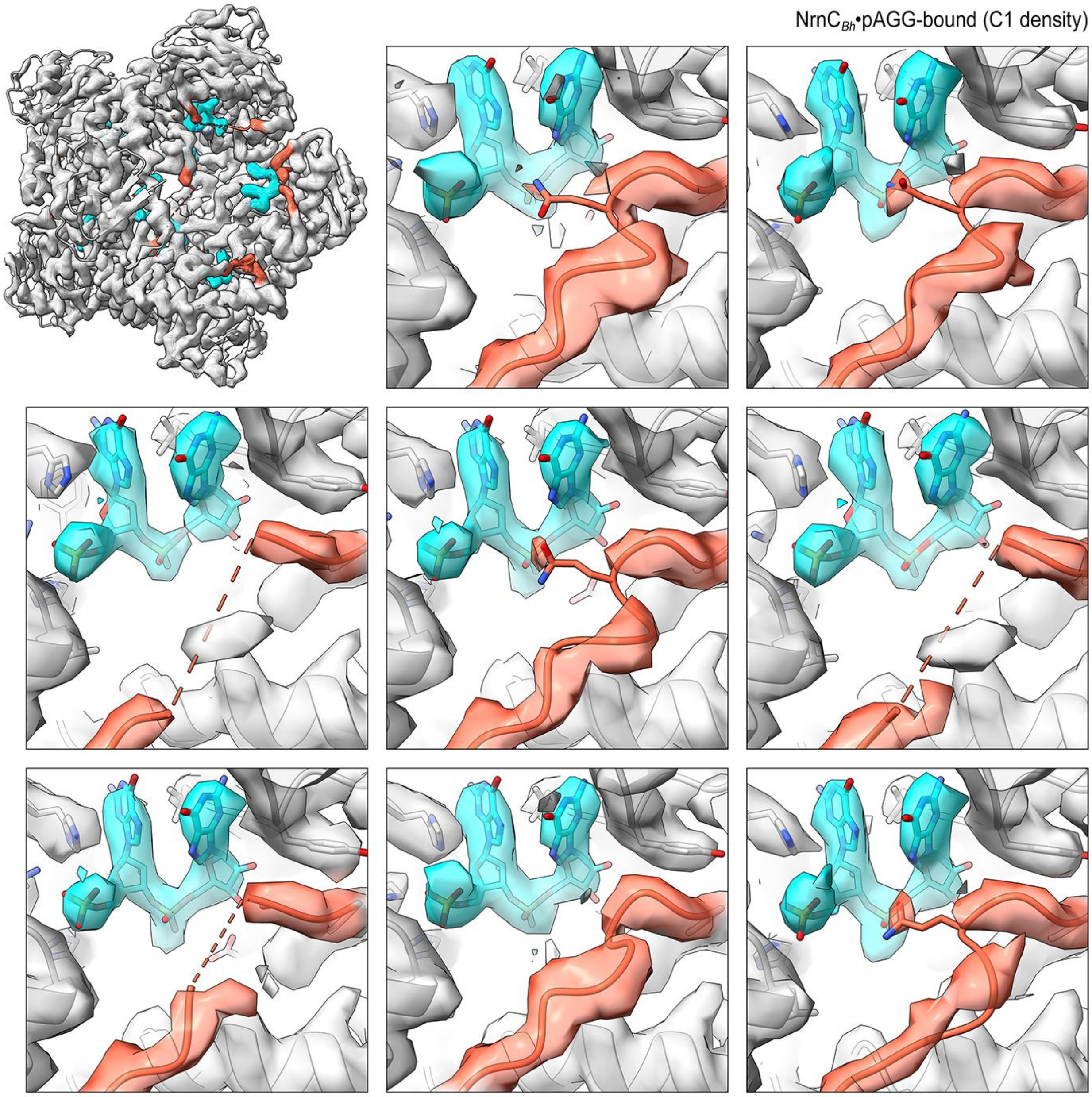
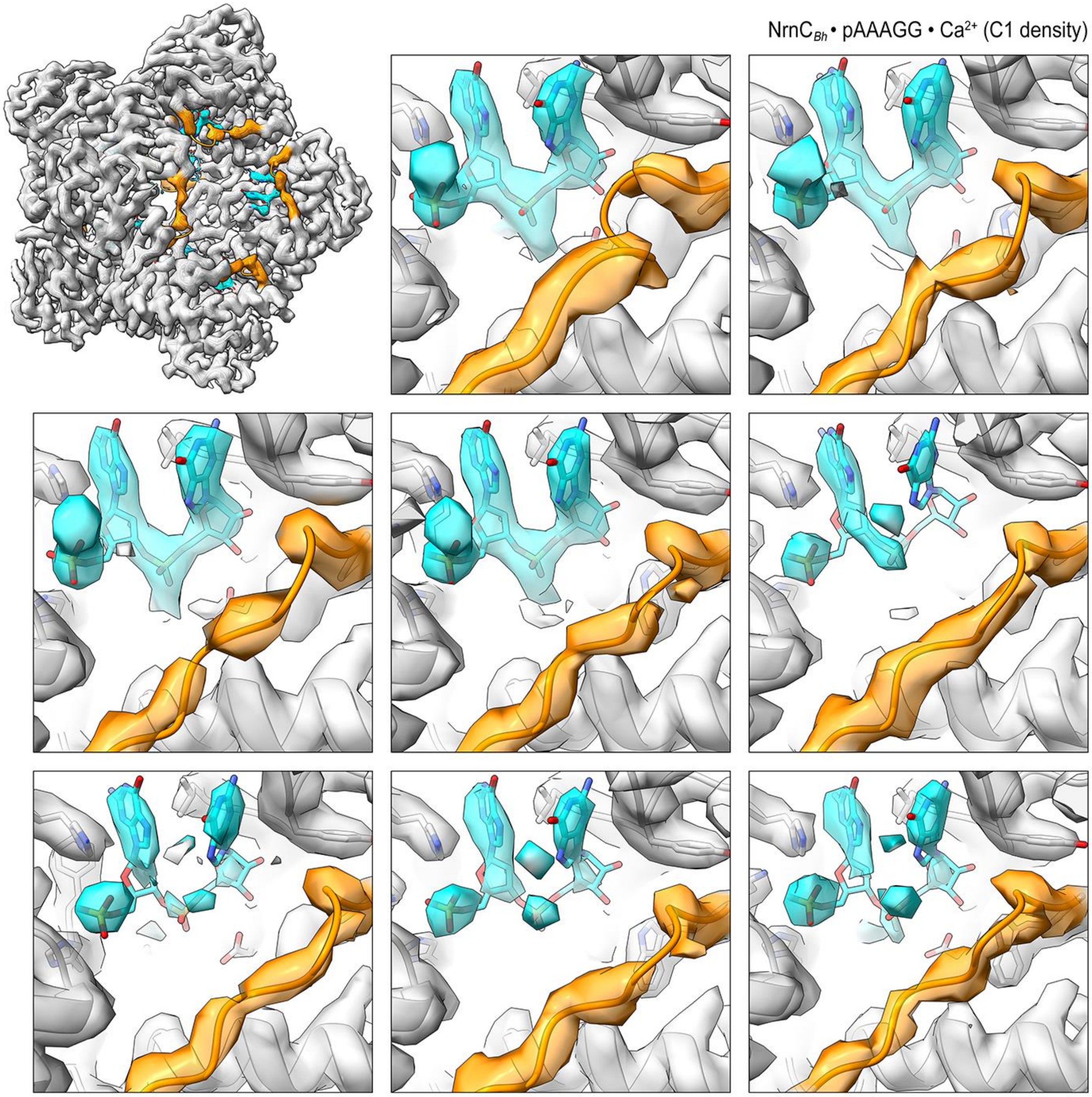


Figure 4-figure supplement 6. Overall and individual active site electron density of a NrnC$_{Bh}$•pAGG octamer after refinement with C1 symmetry.



843

844**Figure 4-figure supplement 7. Overall and individual active site electron density of a NrnC$_{Bh}$•pAAAGG octamer after refinement with C1 symmetry.**

846

847

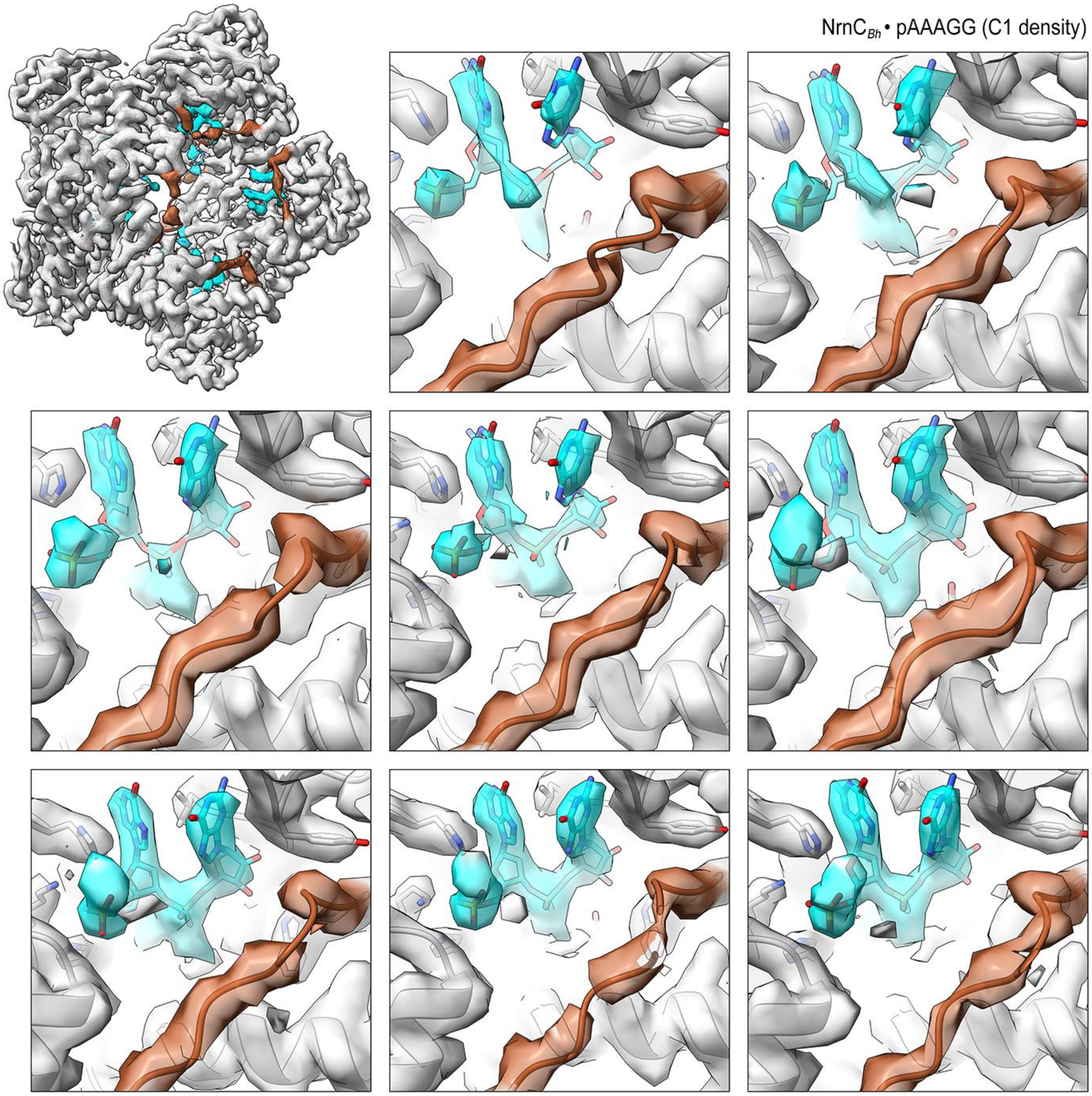


Figure 4-figure supplement 8. Overall and individual active site electron density of a NrnC$_{Bh}$•pAAAGG octamer in the presence of Ca^{2+} ions after refinement with C1 symmetry.

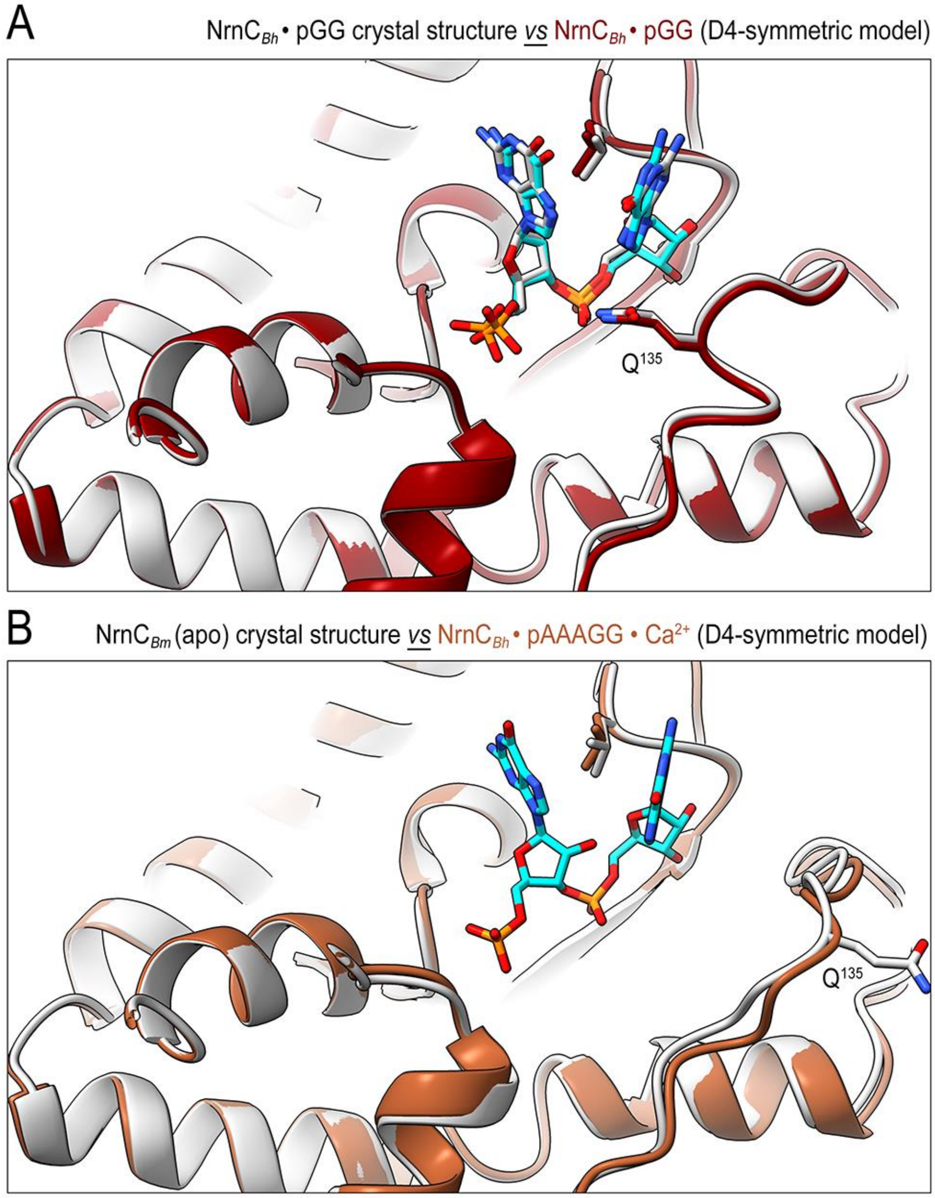


Figure 4-figure supplement 9. The conformation of NrnC bound to substrates with more than two bases resembles the crystallographic apo-state. (A) Comparison of the crystal structure of NrnC$_{Bh}$-pGG with the corresponding cryo-EM structure shows agreement between the solution and crystalline state of the protein with a well-ordered conformation of the loop residues 130-137 engaging the substrate. (B) Comparison of the crystal structure of apo-NrnC$_{Bm}$ with the cryo-EM structure of NrnC$_{Bm}$-pAAAGG. The superposition indicates that longer substrates may bind the active site but only the first full residues appear ordered, resulting in a conformation of NrnC similar to the inactive state observed in the apo-state crystal structures.

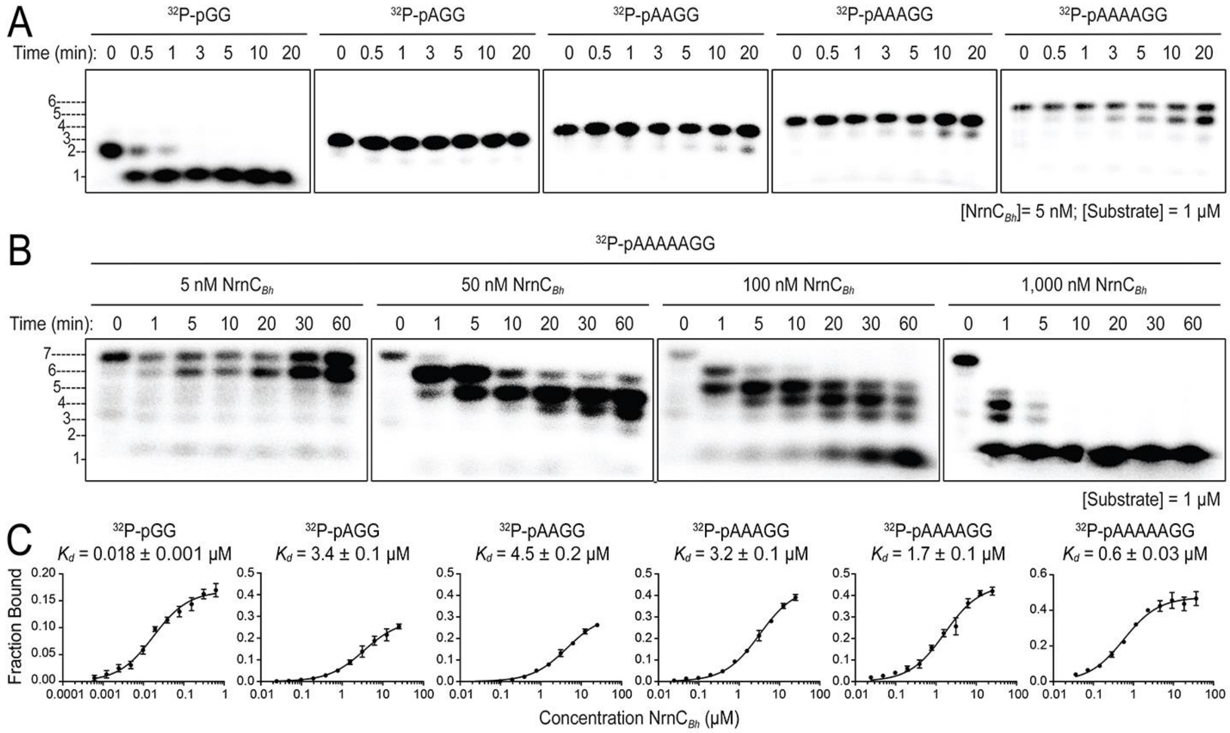


Figure 5. NrnC shows a strong preference for substrates with two residues in length. (A) and (B). RNase assays. Experiments are similar to those in Figure 2 but were performed with radiolabeled substrates from 2 to 7 residues in length. Representative gels of at least two independent experiments are shown. In (B) enzyme concentration was varied from 5-1,000 nM (1:200 to 1:1 enzyme:substrate ratio). Substrate length-dependent binding studies. (C) Affinity of NrnC for RNA with different lengths. Fraction bound of radiolabeled substrates of increasing length was assessed at different NrnC concentrations and is plotted as means and SD from three independent experiments.

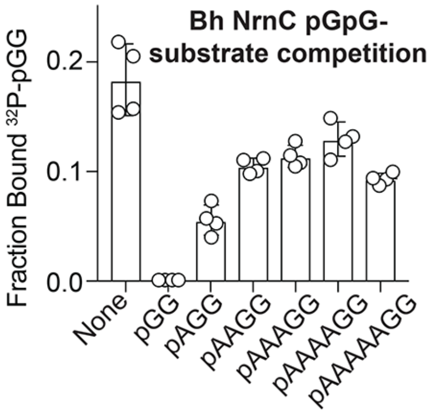


Figure 5-figure supplement 1. Competition binding studies. Fraction bound of ^{32}P-pGG to 200 nM purified NrnC$_{Bh}$ in presence of no competitor or 100 µM unlabeled RNA as indicated. Individual data, means, and SD of four independent experiments are plotted.

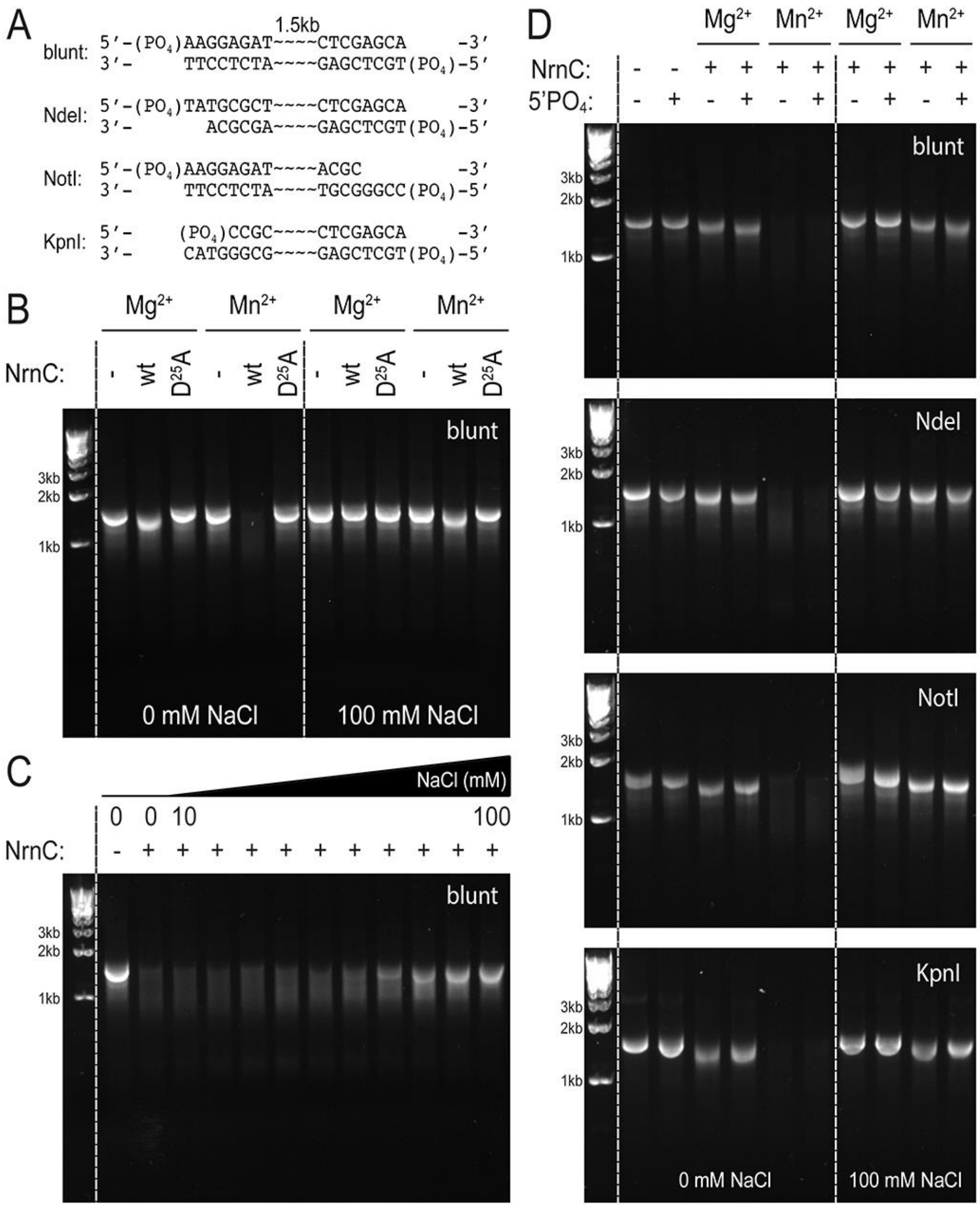


Figure 5-figure supplement 2. NrnC*Bh* degrades long DNA fragments under distinct conditions. (A) DNA fragments tested in this assay. (B) DNase activity of wild-type NrnC*Bh* and a catalytically inactive mutant variant on blunt dsDNA in the presence of either Mg^{2+} or Mn^{2+}, and in the absence or presence of NaCl. (C) NaCl titration on blunt dsDNA using wild-type NrnC*Bh*. (D) NrnC activity on various dsDNA substrates with or without a 5'-PO_4 and in the presence of Mg^{2+} of Mn^{2+}. Representative agarose gels are shown from at least two independent experiments.

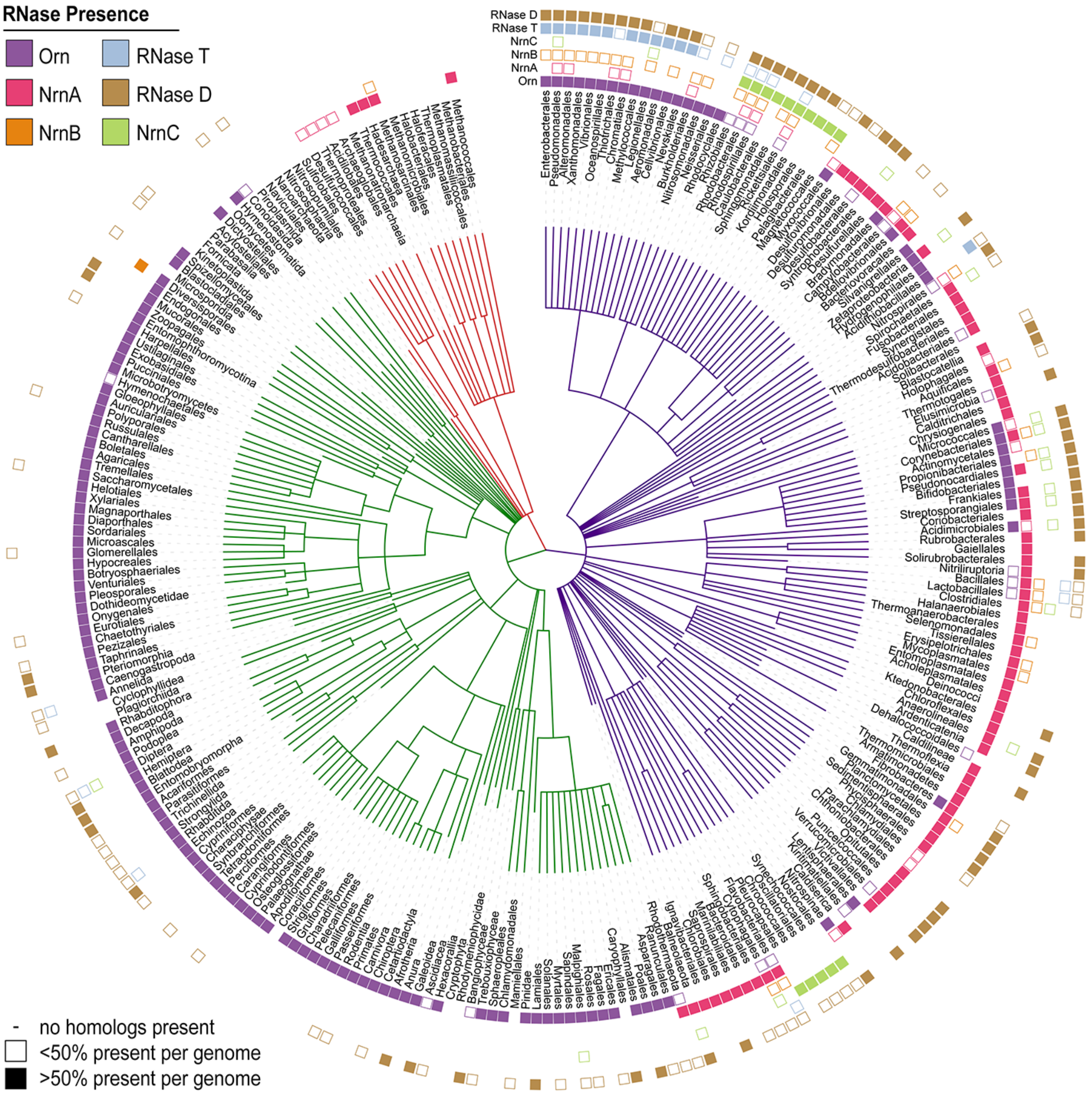


Figure 6. Presence of RNase homologs across sequenced organism classes. Shown is a “Tree of Life” with all taxonomic groups at the class level with at least one substantially-complete proteome available in the dataset. The tree is based on the structure of the NCBI Taxonomy database, with bacterial taxa shown with purple lines, eukaryotic taxa shown with green lines, and archaeal taxa shown with red lines. The presence of each RNase homolog as a proportion of the total proteins in that taxonomic group is shown as either a filled square (>50% presence of a

homolog per genome) or an empty square (<50% presence of a homolog per genome). Lack of a square indicates no homologs for that family were present in genomes of that class.

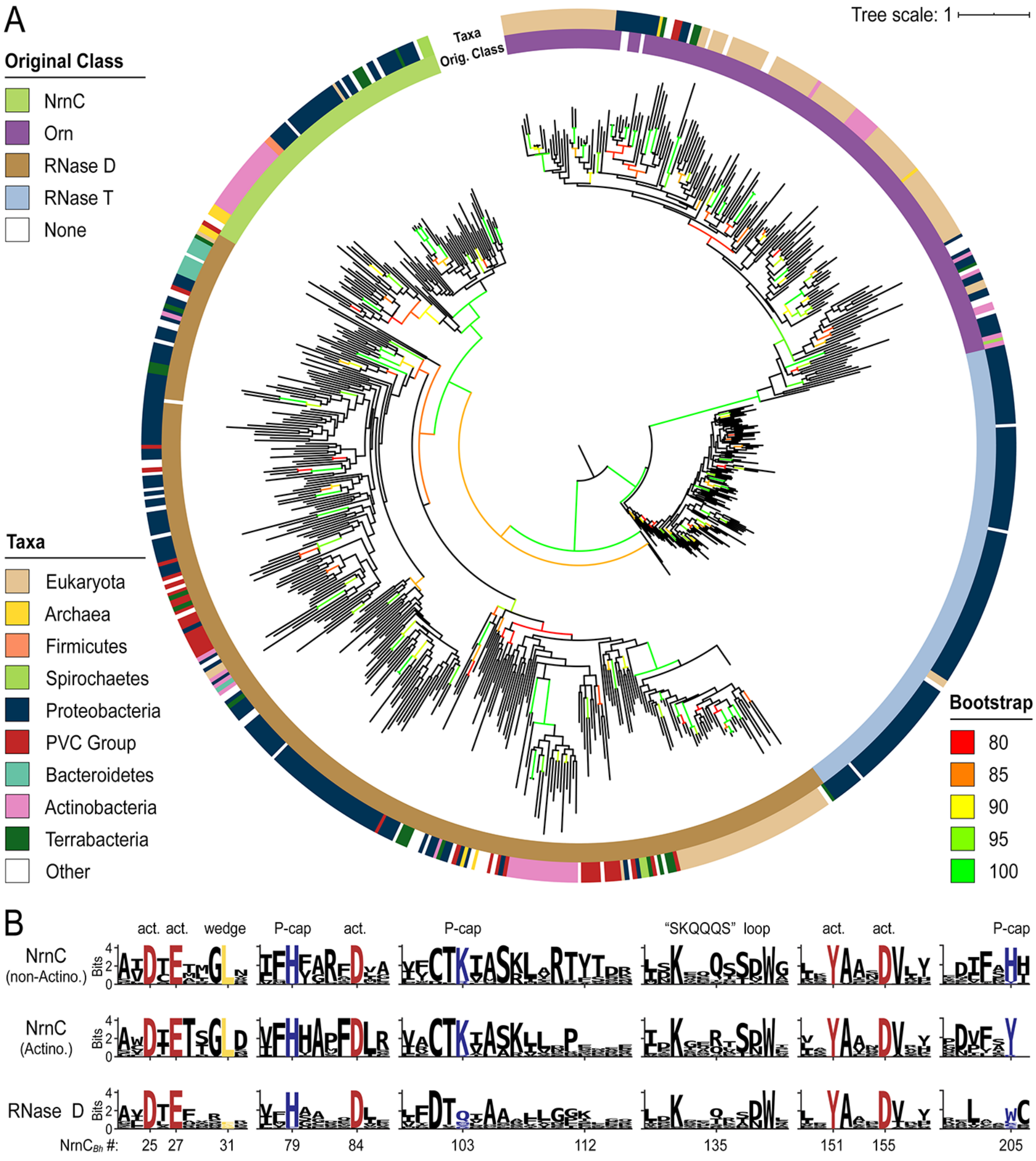


Figure 7. Phylogenetic tree of four DnaQ-fold RNase families. (A) Phylogenetic tree of 669 representatives of the RNase T, RNase D, Orn, and NrnC families of RNase proteins. The inner

ring represents the original classification of each sequence by HMM analysis. The outer ring represents the high-level taxonomic classification of the organism the protein is found in. The color of the branch represents the UFBoot bootstrap value, where black branches are <80%, red is 80%, orange is 85%, yellow is 90%, light green is 95%, and bright green is 100%. Bootstrap values >90% indicate high-confidence splits. (B) Sequence logos of RNase D and NrnC subgroups**.** Sequence logos showing the relative entropy (information content) at selected positions in RNase D as well as the Actinobacterial and non-Actinobacterial subsets of NrnC. Sequence numbering is relative to *Bartonella birtlessi* NrnC (G4VUY7). Active site residues are shown in red, phosphate cap residues in dark blue, and the L-wedge in yellow.

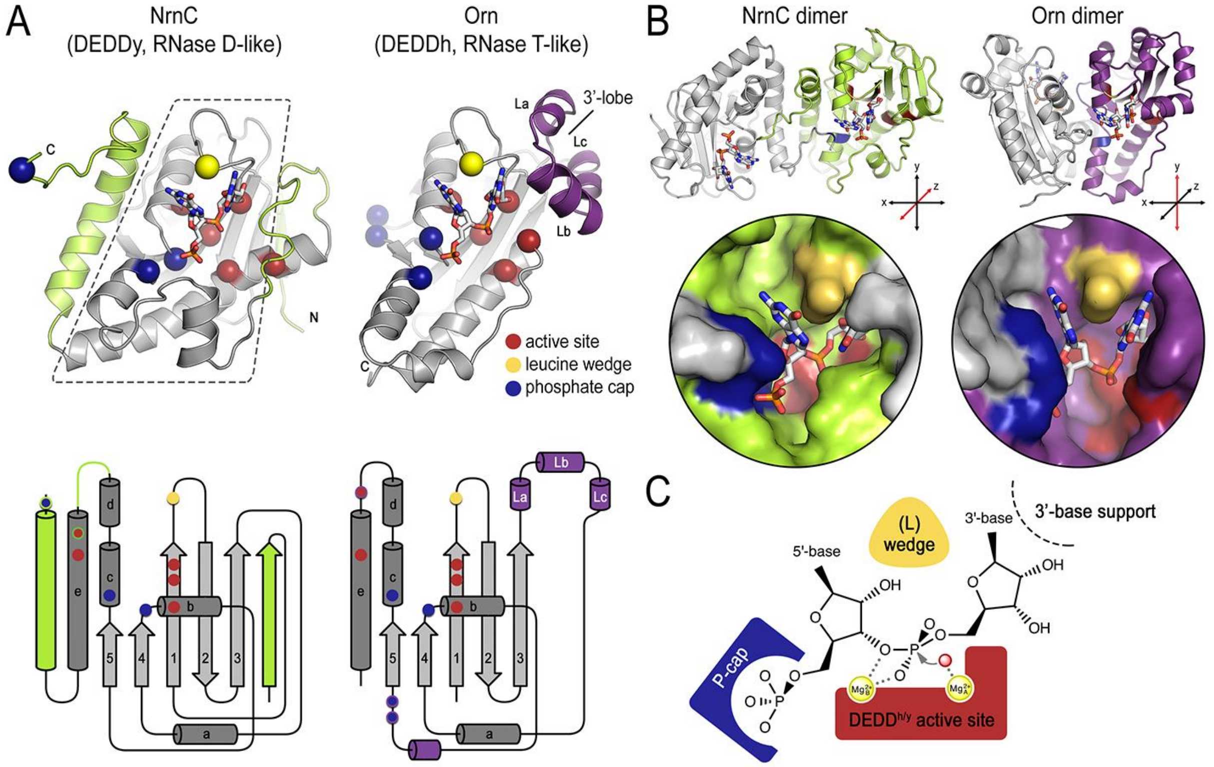


Figure 8. Structural comparison NrnC and Orn. (A) Fold topology. pGG-bound NrnC and Orn monomers are shown in a similar orientation as cartoons (top) or schematic topology diagrams (bottom). Conserved catalytic core elements are colored in grey. NrnC and Orn-specific features are colored in green and purple, respectively. Other color codes mark the positions of the DEDDy/h motif (red spheres), L-wedge (yellow sphere), and phosphate cap

residues (dark blue spheres). (B) Comparison of dimer units of NrnC and Orn (top) with close-ups of the composite active sites of the enzymes (bottom). A NrnC monomer is colored green and an Orn monomer is colored purple, with adjacent monomers in the biological assemblies colored in light grey. Specific residues are colored as in (A). Coordinate systems indicate the 2-fold symmetry axis of the enzyme dimers, with the colored monomers shown in a similar orientation. (C) Structurally and functionally conserved features common among NrnC- and Orn-type diribonucleotidases.

References

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A., & Adams, P. D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol*, *74*(Pt 6), 531-544. https://doi.org/10.1107/s2059798318006551

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., & Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res*, *34*(Web Server issue), W604-608. https://doi.org/10.1093/nar/gkl092

Chin, K. H., Yang, C. Y., Chou, C. C., Wang, A. H., & Chou, S. H. (2006). The crystal structure of XC847 from Xanthomonas campestris: a 3'-5' oligoribonuclease of DnaQ fold family with a novel opposingly shifted helix. *Proteins*, *65*(4), 1036-1040. https://doi.org/10.1002/prot.21148

Choi, K. H., Kumar, A., & Schweizer, H. P. (2006). A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods*, *64*(3), 391-397. https://doi.org/10.1016/j.mimet.2005.06.001

Christen, B., Abeliuk, E., Collier, J. M., Kalogeraki, V. S., Passarelli, B., Coller, J. A., Fero, M. J., McAdams, H. H., & Shapiro, L. (2011). The essential genome of a bacterium. *Mol Syst Biol*, *7*, 528. https://doi.org/10.1038/msb.2011.58

Cohen, D., Mechold, U., Nevenzal, H., Yarmiyhu, Y., Randall, T. E., Bay, D. C., Rich, J. D., Parsek, M. R., Kaever, V., Harrison, J. J., & Banin, E. (2015). Oligoribonuclease is a central feature of cyclic diguanylate signaling in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A*, *112*(36), 11359-11364. https://doi.org/10.1073/pnas.1421450112

Datta, A. K., & Niyogi, K. (1975). A novel oligoribonuclease of Escherichia coli. II. Mechanism of action. *J Biol Chem*, *250*(18), 7313-7319.

Dehio, C. (2005). Bartonella-host-cell interactions and vascular tumour formation. *Nat Rev Microbiol*, *3*(8), 621-631. https://doi.org/10.1038/nrmicro1209

Druzhinin, S. Y., Tran, N. T., Skalenko, K. S., Goldman, S. R., Knoblauch, J. G., Dove, S. L., & Nickels, B. E. (2015). A Conserved Pattern of Primer-Dependent Transcription Initiation in Escherichia coli and Vibrio cholerae Revealed by 5' RNA-seq. *PLoS Genet*, *11*(7), e1005348. https://doi.org/10.1371/journal.pgen.1005348

Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Comput Biol*, *7*(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, *32*(5), 1792-1797. https://doi.org/10.1093/nar/gkh340

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 4), 486-501. https://doi.org/10.1107/s0907444910007493

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr*, *62*(Pt 1), 72-82. https://doi.org/10.1107/s0907444905036693

Fang, M., Zeisberg, W. M., Condon, C., Ogryzko, V., Danchin, A., & Mechold, U. (2009). Degradation of nanoRNA is performed by multiple redundant RNases in Bacillus subtilis. *Nucleic Acids Res*, *37*(15), 5114-5125. https://doi.org/10.1093/nar/gkp527

Federhen, S. (2012). The NCBI Taxonomy database. *Nucleic Acids Res*, *40*(Database issue), D136-143. https://doi.org/10.1093/nar/gkr1178

Galperin, M. Y., & Koonin, E. V. (2012). Divergence and convergence in enzyme evolution. *J Biol Chem*, *287*(1), 21-28. https://doi.org/10.1074/jbc.R111.241976

Ghosh, S., & Deutscher, M. P. (1999). Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc Natl Acad Sci U S A*, *96*(8), 4372-4377. https://doi.org/10.1073/pnas.96.8.4372

Goldman, S. R., Sharp, J. S., Vvedenskaya, I. O., Livny, J., Dove, S. L., & Nickels, B. E. (2011). NanoRNAs prime transcription initiation in vivo. *Mol Cell*, *42*(6), 817-825. https://doi.org/10.1016/j.molcel.2011.06.005

Greenfield, R. A., Drevets, D. A., Machado, L. J., Voskuhl, G. W., Cornea, P., & Bronze, M. S. (2002). Bacterial pathogens as biological weapons and agents of bioterrorism. *Am J Med Sci*, *323*(6), 299-315. https://doi.org/10.1097/00000441-200206000-00003

Hmelo, L. R., Borlee, B. R., Almblad, H., Love, M. E., Randall, T. E., Tseng, B. S., Lin, C., Irie, Y., Storek, K. M., Yang, J. J., Siehnel, R. J., Howell, P. L., Singh, P. K., Tolker-Nielsen, T., Parsek, M. R., Schweizer, H. P., & Harrison, J. J. (2015). Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. *Nat Protoc*, *10*(11), 1820-1841. https://doi.org/10.1038/nprot.2015.115

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol*, *35*(2), 518-522. https://doi.org/10.1093/molbev/msx281

Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol*, *33*(6), 1635-1638. https://doi.org/10.1093/molbev/msw046

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., von Mering, C., & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*, *47*(D1), D309-d314. https://doi.org/10.1093/nar/gky1085

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr*, *66*(Pt 2), 133-144. https://doi.org/10.1107/s0907444909047374

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*, *14*(6), 587-589. https://doi.org/10.1038/nmeth.4285

Kamp, H. D., Patimalla-Dipali, B., Lazinski, D. W., Wallace-Gadsden, F., & Camilli, A. (2013). Gene fitness landscapes of Vibrio cholerae at important stages of its life cycle. *PLoS Pathog*, *9*(12), e1003800. https://doi.org/10.1371/journal.ppat.1003800

Katoh, K., Kuma, K., Toh, H., & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*, *33*(2), 511-518. https://doi.org/10.1093/nar/gki198

Keegan, R. M., McNicholas, S. J., Thomas, J. M. H., Simpkin, A. J., Simkovic, F., Uski, V., Ballard, C. C., Winn, M. D., Wilson, K. S., & Rigden, D. J. (2018). Recent developments in MrBUMP: better search-model preparation, graphical interaction with search models, and solution improvement and assessment. *Acta Crystallogr D Struct Biol*, *74*(Pt 3), 167-182. https://doi.org/10.1107/s2059798318003455

Kim, S. K., Lormand, J. D., Weiss, C. A., Eger, K. A., Turdiev, H., Turdiev, A., Winkler, W. C., Sondermann, H., & Lee, V. T. (2019). A dedicated diribonucleotidase resolves a key bottleneck for the terminal step of RNA degradation. *Elife*, *8*. https://doi.org/10.7554/eLife.46313

Köster, J., & Rahmann, S. (2018). Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics*, *34*(20), 3600. https://doi.org/10.1093/bioinformatics/bty350

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, *372*(3), 774-797. https://doi.org/10.1016/j.jmb.2007.05.022

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*, *47*(W1), W256-w259. https://doi.org/10.1093/nar/gkz239

Liao, H., Liu, M., & Guo, X. (2018). The special existences: nanoRNA and nanoRNase. *Microbiol Res*, *207*, 134-139. https://doi.org/10.1016/j.micres.2017.11.014

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J., & Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*, *75*(Pt 10), 861-877. https://doi.org/10.1107/s2059798319011471

Liu, M. F., Cescau, S., Mechold, U., Wang, J., Cohen, D., Danchin, A., Boulouis, H. J., & Biville, F. (2012). Identification of a novel nanoRNase in Bartonella. *Microbiology (Reading)*, *158*(Pt 4), 886-895. https://doi.org/10.1099/mic.0.054619-0

Mechold, U., Fang, G., Ngo, S., Ogryzko, V., & Danchin, A. (2007). YtqI from Bacillus subtilis has both oligoribonuclease and pAp-phosphatase activity. *Nucleic Acids Res*, *35*(13), 4552-4561. https://doi.org/10.1093/nar/gkm462

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol*, *37*(5), 1530-1534. https://doi.org/10.1093/molbev/msaa015

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P. C., Timony, M. A., Ottaviano, M., & Sliz, P. (2013). Collaboration gets the most out of software. *Elife*, *2*, e01456. https://doi.org/10.7554/eLife.01456

Newman, J. R., & Fuqua, C. (1999). Broad-host-range expression vectors that carry the L-arabinose-inducible Escherichia coli araBAD promoter and the araC regulator. *Gene*, *227*(2), 197-203. https://doi.org/10.1016/s0378-1119(98)00601-5

Nicholls, T. J., Spåhr, H., Jiang, S., Siira, S. J., Koolmeister, C., Sharma, S., Kauppila, J. H. K., Jiang, M., Kaever, V., Rackham, O., Chabes, A., Falkenberg, M., Filipovska, A., Larsson, N. G., & Gustafsson, C. M. (2019). Dinucleotide Degradation by REXO2 Maintains Promoter Specificity in Mammalian Mitochondria. *Mol Cell*, *76*(5), 784-796.e786. https://doi.org/10.1016/j.molcel.2019.09.010

Niyogi, S. K., & Datta, A. K. (1975). A novel oligoribonuclease of Escherichia coli. I. Isolation and properties. *J Biol Chem*, *250*(18), 7307-7312.

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol*, *302*(1), 205-217. https://doi.org/10.1006/jmbi.2000.4042

Omelchenko, M. V., Galperin, M. Y., Wolf, Y. I., & Koonin, E. V. (2010). Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. *Biol Direct*, *5*, 31. https://doi.org/10.1186/1745-6150-5-31

Orr, M. W., Donaldson, G. P., Severin, G. B., Wang, J., Sintim, H. O., Waters, C. M., & Lee, V. T. (2015). Oligoribonuclease is the primary degradative enzyme for pGpG in Pseudomonas aeruginosa that is required for cyclic-di-GMP turnover. *Proc Natl Acad Sci U S A*, *112*(36), E5048-5057. https://doi.org/10.1073/pnas.1507245112

Orr, M. W., Weiss, C. A., Severin, G. B., Turdiev, H., Kim, S. K., Turdiev, A., Liu, K., Tu, B. P., Waters, C. M., Winkler, W. C., & Lee, V. T. (2018). A Subset of Exoribonucleases Serve as Degradative Enzymes for pGpG in c-di-GMP Signaling. *J Bacteriol*, *200*(24). https://doi.org/10.1128/jb.00300-18

Palace, S. G., Proulx, M. K., Lu, S., Baker, R. E., & Goguen, J. D. (2014). Genome-wide mutant fitness profiling identifies nutritional requirements for optimal growth of Yersinia pestis in deep tissue. *mBio*, *5*(4). https://doi.org/10.1128/mBio.01385-14

Pappas, G., Papadimitriou, P., Akritidis, N., Christou, L., & Tsianos, E. V. (2006). The new global map of human brucellosis. *Lancet Infect Dis*, *6*(2), 91-99. https://doi.org/10.1016/s1473-3099(06)70382-6

Park, J., Lee, S. Y., Jeong, H., Kang, M. G., Van Haute, L., Minczuk, M., Seo, J. K., Jun, Y., Myung, K., Rhee, H. W., & Lee, C. (2019). The structure of human EXD2 reveals a chimeric 3' to 5' exonuclease domain that discriminates substrates via metal coordination. *Nucleic Acids Res*, *47*(13), 7078-7093. https://doi.org/10.1093/nar/gkz454

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, *30*(1), 70-82. https://doi.org/10.1002/pro.3943

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*, *14*(3), 290-296. https://doi.org/10.1038/nmeth.4169

Roelofs, K. G., Wang, J., Sintim, H. O., & Lee, V. T. (2011). Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc Natl Acad Sci U S A*, *108*(37), 15528-15533. https://doi.org/10.1073/pnas.1018949108

Rosta, E., Yang, W., & Hummer, G. (2014). Calcium inhibition of ribonuclease H1 two-metal ion catalysis. *J Am Chem Soc*, *136*(8), 3137-3144. https://doi.org/10.1021/ja411408x

Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., & Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res*, *47*(W1), W5-w10. https://doi.org/10.1093/nar/gkz342

Schmier, B. J., Nelersa, C. M., & Malhotra, A. (2017). Structural Basis for the Bidirectional Activity of Bacillus nanoRNase NrnA. *Sci Rep*, *7*(1), 11085. https://doi.org/10.1038/s41598-017-09403-x

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, *9*(7), 671-675. https://doi.org/10.1038/nmeth.2089

Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol*, *35*(11), 1026-1028. https://doi.org/10.1038/nbt.3988

Sternon, J. F., Godessart, P., Gonçalves de Freitas, R., Van der Henst, M., Poncin, K., Francis, N., Willemart, K., Christen, M., Christen, B., Letesson, J. J., & De Bolle, X. (2018). Transposon Sequencing of Brucella abortus Uncovers Essential Genes for Growth In Vitro and Inside Macrophages. *Infect Immun*, *86*(8). https://doi.org/10.1128/iai.00312-18

Tareen, A., & Kinney, J. B. (2020). Logomaker: beautiful sequence logos in Python. *Bioinformatics*, *36*(7), 2272-2274. https://doi.org/10.1093/bioinformatics/btz921

Vvedenskaya, I. O., Sharp, J. S., Goldman, S. R., Kanabar, P. N., Livny, J., Dove, S. L., & Nickels, B. E. (2012). Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs. *Genes Dev*, *26*(13), 1498-1507. https://doi.org/10.1101/gad.192732.112

Wasmuth, E. V., Januszyk, K., & Lima, C. D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. *Nature*, *511*(7510), 435-439. https://doi.org/10.1038/nature13406

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*, *67*(Pt 4), 235-242. https://doi.org/10.1107/s0907444910045749

Yu, D., & Deutscher, M. P. (1995). Oligoribonuclease is distinct from the other known exoribonucleases of Escherichia coli. *J Bacteriol*, *177*(14), 4137-4139. https://doi.org/10.1128/jb.177.14.4137-4139.1995

Yuan, Z., Gao, F., Yin, K., & Gu, L. (2018). NrnC, an RNase D-Like Protein From Agrobacterium, Is a Novel Octameric Nuclease That Specifically Degrades dsDNA but Leaves dsRNA Intact. *Front Microbiol*, *9*, 3230. https://doi.org/10.3389/fmicb.2018.03230

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J Struct Biol*, *193*(1), 1-12. https://doi.org/10.1016/j.jsb.2015.11.003

Zhang, X., Zhu, L., & Deutscher, M. P. (1998). Oligoribonuclease is encoded by a highly conserved gene in the 3'-5' exonuclease superfamily. *J Bacteriol*, *180*(10), 2779-2781. https://doi.org/10.1128/jb.180.10.2779-2781.1998

Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods*, *14*(4), 331-332. https://doi.org/10.1038/nmeth.4193

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E., & Scheres, S. H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife*, *7*. https://doi.org/10.7554/eLife.42166

Zuo, Y., & Deutscher, M. P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res*, *29*(5), 1017-1026. https://doi.org/10.1093/nar/29.5.1017

Zuo, Y., Wang, Y., & Malhotra, A. (2005). Crystal structure of Escherichia coli RNase D, an exoribonuclease involved in structured RNA processing. *Structure*, *13*(7), 973-984. https://doi.org/10.1016/j.str.2005.04.015