Nucleic Acids Research, 2025, **53**, gkaf644
https://doi.org/10.1093/nar/gkaf644
RNA and RNA-protein complexes

An endoribonuclease of the YicC-like family delays sporulation via sRNA degradation in *Clostridioides difficile*

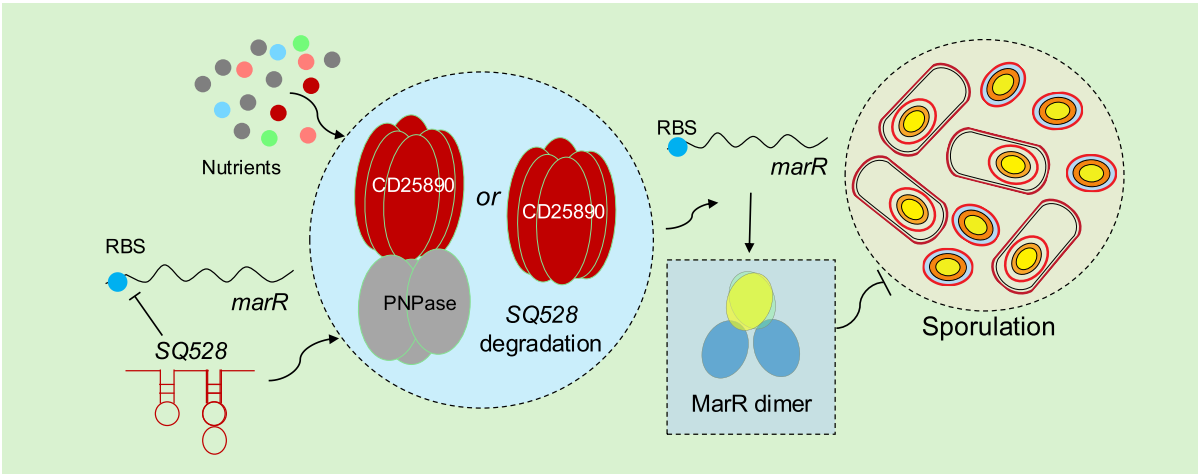
Diogo Martins[1], Bruno Salgueiro[1], Daniel Sobral[2], Marcos Gragera[3], Zach Hensel[1], Adriano O. Henriques[1], Célia V. Romão[1], Mónica Serrano[1,*]

[1]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Avenida da República EAN, 2780-157 Oeiras, Portugal
[2]Genomics and Bioinformatics Unit, Department of Infectious Diseases, National Institute of Health Doutor Ricardo Jorge (INSA), Av. Padre Cruz, 1649-016 Lisboa, Portugal
[3]Centro Nacional de Biotecnología (CNB), CSIC, C. Darwin, 3, Fuencarral-El Pardo, 28049 Madrid, Spain
[*]To whom correspondence should be addressed. Email: serrano@itqb.unl.pt

Abstract

Clostridioides difficile CD25890 is a YicC-like endoribonuclease involved in regulating sporulation initiation, a process critical for the host–host transmission of this anaerobic pathogen. Using comparative transcriptomics we identified a small RNA, *SQ528*, that accumulates at higher levels in a *CD25890* deletion mutant and we show that purified CD25890 cleaves *SQ528* in a metal-dependent manner. Moreover, the overexpression of *SQ528* increases sporulation under certain nutritional conditions phenocopying a *CD25890* deletion mutant. CD25890 is an hexamer in solution and *in vivo*. An N-terminal domain, which self-interacts as assessed by size exclusion chromatography and a two hybrid assay, is essential for oligomerization of CD25890. A C-terminal domain harbours residues H230, E254, and E258, conserved among orthologues, important for catalysis. AlphaFold2 modelling and cryo-EM suggest an elongated barrel-like structure with an internal cavity lined with basic residues that may aid in RNA binding. We show that CD25890 forms a complex with polynucleotide phosphorylase which combines the endoribonuclease activity of the first with the exonucleolytic activity of the latter and leads to the complete degradation of *SQ528*. This study identifies a native substrate for the YicC-family of ribonucleases and advances our understanding of the role of CD25890 in sporulation initiation in *C. difficile*.

Graphical abstract



Introduction

In their natural habitats, bacteria endure recurrent exposure to stressful and shifting environmental conditions and have evolved mechanisms to perceive the surrounding environment and adjust their gene expression patterns to suit the prevailing conditions. At the onset of sporulation, both internal and external cues signal the vegetative cell to halt growth and initiate the formation of a metabolically inactive, resistant spore capable of long-term survival [1]. In most spore-forming bacteria, as in *Bacillus subtilis*, sporulation is triggered by unfavourable growth conditions, including extreme nutrient depletion [2]. Conversely, bacteria such as *Clostridioides difficile* sporulate

Received: November 12, 2024. **Revised:** May 29, 2025. **Editorial Decision:** June 16, 2025. **Accepted:** June 19, 2025

at a low but continuous rate, even under nutrient-rich conditions [1, 3, 4].

Clostridioides difficile is an anaerobic enteric pathogen in which the etiologic agent is the oxygen-resistant spore, while the toxins produced by vegetative cells are the main cause of the disease symptoms [5–8]. In this pathogen, formation of the highly resistant spore is an important factor in persistence in the environment and in the host, causing high rates of disease recurrence [1, 9]. Moreover, *C. difficile* toxins, expressed by sporulating cells, associate in their active forms with spores, making spores toxin-delivery vehicles [10]. Entry into sporulation requires the Spo0A transcriptional regulator which is post-translationally activated by phosphorylation [1, 9, 11]. Once activated, Spo0A induces the expression of genes required for cell-type-specific gene expression during sporulation [1, 12, 13]. We have shown before that *CD25890* is involved in the control of sporulation initiation in response to nutritional signals. *CD25890* represses sporulation by indirectly inhibiting *spo0A* expression [14]. The mechanism by which CD25890 modulates gene expression and influence spore formation remains unknown.

CD25890 is the *C. difficile* member of a novel family of bacterial ribonucleases, the YicC family, widely conserved across bacteria and involved in small RNA (sRNA) degradation [15–18]. To date, characterization of this family has focused on the proteins from the model organisms *Escherichia coli* and *B. subtilis,* called YicC and YloC, respectively. YloC and YicC were shown to be hexameric, metal-dependent enzymes that catalyse cleavage of synthetic single-stranded RNA [16–18]. Both enzymes were shown to form a complex with the polynucleotide phosphorylase (PNPase) [15, 17]. *Escherichia coli* YicC was proposed to work as an adaptor for the degradation by PNPase of *ryhB*, a noncoding RNA involved in the degradation of messenger RNAs (mRNAs) that encode iron-utilizing proteins [15]. So far, however no direct physiological substrate was identified for this family of ribonucleases.

Here, we identify the sRNA *SQ528* as a putative substrate of CD25890. Comparative transcriptomic analysis showed that *SQ528* accumulates at higher levels in the *CD25890* mutant than in the wild-type (WT) strain, even though transcription remains unchanged. Overexpression of *SQ528* from an inducible promoter increases sporulation under certain nutritional conditions. We further show that purified CD25890 assembles into a hexamer with ribonuclease activity against *SQ528*. The N-terminal domain is required for the formation of the hexamer and sufficient for binding nucleic acids, while residues required for catalysis are present in the C-terminal domain of the protein. Structural insights from structural modelling and Cryo-electron microscopy (cryo-EM) suggest that CD25890 adopts an elongated, barrel-like structure with cavity lined with basic residues that may facilitate RNA binding. Additionally, similar to YloC and YicC, we show that CD25890 forms an active complex with PNPase, potentially implicating CD25890 in broader RNA processing mechanisms.

Materials and methods

Bacterial strains, media, and general methods

Bacterial strains and their relevant properties are listed in Supplementary Table S2. All plasmids constructions were conducted in *E. coli* DH5a (Bethesda Research laboratories), BL21 (DE3) (Novagen) was used for protein production, strain HB101 (RP4) was used as the donor in conjugation experiments [19] and strain BTH101 was used in bacterial two hybrid assays [20]. Routine growth of *E. coli* strains was done at 37°C in Luria–Bertani (LB) medium. When needed, ampicillin (100 μg/ml), chloramphenicol (15 μg/ml) or kanamycin (30 μg/ml) was added to the culture medium. All *C. difficile* strains used in this study are derivatives of the WT strain 630Δ*erm* [19] and were routinely grown anaerobically (5% H_2, 15% CO_2, 80% N_2) at 37°C in Brain Heart Infusion (BHI) medium (Oxoid). When necessary, cefoxitin (25 μg/ml) or thiamphenicol (15 μg/ml) was added to the culture medium. For sporulation induction, strains were grown in liquid sporulation medium (SM [21]) and on 70:30 agar medium [22] and incubated at 37°C. When indicated, anhydrotetracycline (aTc) was added to the culture medium to a final concentration of 50 mM. To determine the total number of cells, the cells were serially diluted and plated on BHI with 0.1% taurocholate (TA; Sigma–Aldrich) to ensure efficient spore germination. To determine the number of spores, the cells were heat killed by incubation for 30 min at 70°C prior to plating on BHI with 0.1% TA. Sporulation efficiency (%) was calculated as the total number of heat-resistant spores divided by the total number of viable cells. Statistical analysis was carried out using GraphPad Prism (version 7.0; GraphPad Software Inc.). All plasmids and primers used in this work are listed in Supplementary Tables S3 and S4, respectively.

Construction of His-Tagged CD25890 N- and C-Terminal Domains

DNA fragments encoding the *CD25890* N-terminal and *CD25890* C-terminal were generated by polymerase chain reaction (PCR) from *C. difficile* 630Δ*erm* genomic DNA using primers CD25890_BamHI_Fw/CD25890_NTER_NotI_Rev, CD25890_CTER_BamHI_Fw/CD25890_NotI_Rev, respectively. The resulting DNA fragments, 453 and 429 bp, were cut with BamHI and NotI and cloned between the same sites of pETDuet-1 (Novagen) to produce pDM67 and pDM72.

Construction of His-Tagged *CD25890^{H230A}*, *CD25890^{E254A}*, *CD25890^{E258A}*, *CD25890$^{E254A/E258A}$* and PNPase

Mutations generating single Ala substitutions (H230A, E254A and E258A) were introduced in CD25890 by Quick Change mutagenesis using primers CD25890_Fw_H230A/CD25890_Rev_H230A; CD25890_Fw_E254A/CD25890_Rev_E254A; CD25890_Fw_E258A/CD25890_Rev_E258A, and plasmid pDM35 [14] as template. The double Ala substitution (E258A/E254A) was obtained using primers CD25890_Fw_E254A/CD25890_Rev_E254A and plasmid pDM86 as template. This created plasmids pDM81 (*His6-CD25890^{H230A}*), pDM84 (*His6-CD25890^{E254A}*), pDM86 (*His6-CD25890^{E258A}*), and pDM94 (*His6-CD25890$^{E258A/E254A}$*). *PNPase* gene was generated by PCR from *C.difficile* 630Δ*erm* genomic DNA using primers PNPase_NcoI_Fw/PNPase_BamHI_Rev. The resulting DNA fragment, 2112 bp, was cut with NcoI and BamHI and cloned between the same sites of pACYDUET-1 (Novagen) to produce pDM97.

Bacterial two hybrid assay and plasmids construction

The proteins $CD25890^{FL}$, $CD25890^{N}$, $CD25890^{C}$, and PNPase were fused to the T18 and T25 fragments of the CyaA protein. The fragments to be fused to T25 CyaA were generated by PCR from 630Δ*erm* genomic DNA using primers $CD25890^{FL}$_pKT25_Fw_BglII/$CD25890^{FL}$_pKT25_Rev_EcoRI, $CD25890^{FL}$_pKT25_Fw_BglII/$CD25890^{N}$_pKT25_Rev_EcoRI, $CD25890^{C}$_pKT25_Fw_BglII/$CD25890^{FL}$_pKT25_Rev_EcoRI, PnPase_pKT25_Fw_BglII/PnPase_pKT25_Rev_EcoRI. The resulting DNA fragments, 882, 453, 429, and 2112 bp respectively, were cut with BglII and EcoRI and inserted between the same sites of pKT25 [20] to produce pDM68, pDM65, pDM63, and pDM66, respectively. The fragments to be fused to T18 CyaA were generated by PCR from 630Δ*erm* genomic DNA using primers $CD25890^{FL}$_pUT18_Fw_BglII/$CD25890^{FL}$_pUT18_Rev_EcoRI, $CD25890^{FL}$_pUT18_Fw_BglII/$CD25890^{N}$_pUT18_Rev_EcoRI, and $CD25890^{C}$_pUT18_Fw_BglII/$CD25890^{FL}$_pUT18_Rev_EcoRI. The resulting DNA fragments, 882, 453, and 429 bp, respectively, were cut with BglII and EcoRI and inserted between the same sites of pUT18 [20] to produce pDM61, pD62, and pDM64, respectively. The *E. coli* strain BTH101 was co-transformed with combinations of the different T18 and T25 fusion proteins. The activity of β-galactosidase was determined with the substrate *o*-nitrophenyl-β-d-galactopyranoside or in LB agar plates supplemented with 5-bromo-4-cloro-3-indolyl-β-D-galactopyranoside (X-Gal) (EUROMEDEX protocol).

Construction of P_{tetA}-*SQ528*

To construct a P*tetA-SQ528*, *SQ528* was amplified by PCR using primers *SQ528*_Fw_NotI/*SQ528*_Rev_SacI, yielding a 350 bp. The fragments were digested with NotI and SacI and inserted between the same sites of pAM25 [23], yielding pDM91. This plasmid was introduced into *E.coli* HB101 (RP4) and then transferred to strains 630Δ*erm* and Δ*CD25890* by conjugation.

Construction of P_{SQ528}-*gusA*

The promotor region of the *SQ528* was amplified by PCR using primers *SQ528*_Fw_SacI/*SQ528*_Rev_NotI, yielding a 350 bp fragment. The fragment was then digested with NotI and SacI and inserted between the same sites of pRPF185 [24], yielding pDM69. This plasmid was introduced into *E. coli* HB101 (RP4) and then transferred to strains 630Δ*erm* and Δ*CD25890* by conjugation.

Overproduction and purification of proteins

Plasmids were introduced into *E. coli* BL21 (DE3) cells and the resulting strains were grown in autoinduction medium [25]. The cells were then harvested by centrifugation (4000 × *g*, for 10 min, at 4°C) and the sediment resuspended in lysis buffer (20 mM phosphate, pH 7.4, 10 mM Imidazole, 1 mM Phenylmethylsulfonyl Fluoride (PMSF)) containing 1 μg/ml DNase I (Sigma). The suspension was lysed using a French pressure cell (at 18 000 lb/in) and the lysate cleared by centrifugation (15 000 × *g*, 30 min at 4°C). Then, the supernatant was loaded onto a 1 ml Histrap column (Cytiva). The bound protein was eluted using lysis buffer with 300 mM of imidazole, followed by the addition of 50 mM ethylenediaminetetraacetic acid (EDTA) to the elution. This step was performed to ensure the removal of any metal complexes. The protein was dialysed against 50 mM Tris, pH 7.5 with 150 mM of NaCl.

Size exclusion chromatography

Purified proteins were loaded onto a S200 Increase 10/300GL (GE Healthcare), previously equilibrated with a buffer composed of 50 mM Tris–HCl, pH 7.5, 150 mM NaCl. The column was calibrated with the gel filtration molecular markers Dextran Blue and Gel Filtration Standard proteins (Bio-Rad, Hercules, USA; thyroglobulin 669 kDa; ferritin 440 kDa; aldolase 158 kDa; conalbumin 75 kDa; ovalbumin 44 kDa; ribonuclease A 13.7 kDa; rubredoxin 6.6 kDa), which was prepared in the same buffer used for equilibration of the column.

Limited trypsin digestion

The substrate protein, CD25890, was purified as described before, and dialysed in 100 mM Tris–HCl at pH 8.5, overnight at 4°C. The protein concentration was estimated using Bradford Protein Assay, in order to add 100 μg of protein to 2 μg of trypsin. The digestion occurred for 4 h at 37°C. Samples corresponding to 10 μg of substrate were collected from the reaction solution at 5, 15, 30, 45, 60, and 240 h after the beginning of the experiment. The digestion profile was traced by resolving the protein content, at each time, in a 12.5% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE).

Blue native–polyacrylamide gel electrophoresis and immunoblotting

Cell extracts were obtained by withdrawing 20 ml samples from BHI cultures of *C. difficile* 6, 8, 10, and 12 h after inoculation. The cells were collected by centrifugation (4000 × *g*, for 5 min at 4°C), the cell sediment was suspended in 1 ml French press buffer (10 mM Tris, pH 8.0, 10 mM $MgCl_2$, 0.5 mM EDTA, 0.2 mM NaCl, 10% glycerol, 1 mM PMSF). The cells were lysed using a French pressure cell (18 000 lb/in2). Proteins (10 μg) were resolved on Blue native–polyacrylamide gel electrophoresis (BN–PAGE) using gradient gels (5%–15%). Also, 3 μg of His-tagged purified CD25890 was analysed by BN–PAGE. Native Mark Unstained Protein Standard (Thermo Fisher) was used as a molecular weight (MW) marker.

For immunoblot analysis, following separation by BN–PAGE, the proteins were transferred by electrophoresis to Polyvinylidene Difluoride (PVDF) membranes (Millipore Immobilon P). Anti-CD25890 was used at a 1:1000 dilution, and an antirabbit secondary antibody conjugated to horseradish peroxidase (Sigma) was used at a 1:10 000 dilution. The immunoblots were developed with enhanced chemiluminescence reagents (Thermo Scientific). Images were adjusted, cropped, and quantified using ImageJ (http://rsbweb.nih.gov/ij/).

Negative staining and transmission electron microscopy

Purified protein were diluted to 0.2 mg/ml in a buffer solution containing 50 mM Tris–HCl, pH 7.5, 250 mM NaCl, and protein complexes were adsorbed for 2 min to glow-discharged, 400 mesh carbon-coated grids. The grids were washed on 4 drops of distilled H_2O and negatively stained with 2% (w/v) uranyl acetate. Grids were imaged on a Hitachi H-7650

Microscope equipped with an AMT digital camera operated at 120 keV.

Cryo-EM

The protein solution was sent to the CSIC CryoEM Center (Madrid, Spain), for analysis. Samples were prepared on EM grids at various concentrations. Initially, 3 µl of the protein solution at different concentrations (1–0.5 mg/ml, and 1:3, 1:6, 1:10, and 1:50 dilutions) were applied to distinct grids with varying quantifoil hole sizes and materials. The grids were vitrified using a Vitrobot (Thermo Fisher Scientific) under different conditions. These variations were analysed using a Talos Arctica microscope (Thermo Fisher Scientific) under cryogenic conditions to assess particle distribution within the grid holes and check for the presence of ice. Low magnification grid atlases were collected for this purpose. A small dataset was acquired on the Talos Arctica to obtain initial 2D classes. The data were processed using Scipion software [26], a framework which integrates state-of-the-art software for Single Particle Analysis. Movies were motion-corrected with Relion 3.1 [27] CTF estimation was carried out with gctf [28] and particle picking with crYOLO [29]. Finally, 2D-classification, 3D-classification, and nonuniform refinement were performed with cryoSPARC [30, 31].

AlphaFold-Multimer modelling

AlphaFold-Multimer models were generated using ColabFold [32, 33]. Structures were predicted for six repeats of the sequence of the CD25890 protein using the "alphafold2_multimer_v3" model and default selection for additional parameters (without template structures). The highest-ranking model (scored by ipTM, a metric of predicted accuracy at protein–protein interfaces) was analysed out of five structure predictions made using different AlphaFold-Multimer weights [34]. All five predictions were in broad agreement and the predicted structure is largely consistent with structures published to date. However, we consistently obtain predictions with approximate six-fold symmetry; published structures lack this symmetry outside of the hexamerization domain. Structural representations were generated using PyMOL Molecular Graphics System (Schrödinger, LLC).

Assay of β-glucuronidase activity in whole cell extracts

Whole cell extracts were obtained by withdrawing 1 ml samples from SM cultures of *C. difficile* 10 h of after inoculation. The cells were collected by centrifugation (4000 × *g*, for 5 min at 4°C). The activity of β-glucuronidase was determined with the substrate 4-nitrophenyl-β-D-glucuronide. Statistical analysis was carried out using GraphPad Prism (version 7.0; GraphPad Software Inc.).

RNA extraction and quantitative reverse transcriptase-polymerase chain reaction analysis

Sporulating cells were collected from cultures of the 630Δ*erm* strain and the Δ*CD25890* mutant after 10 h of growth in SM. RNA was extracted from two independent cultures using the RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Complementary DNA (cDNA) was synthesized from 4 µg of total RNA using 1 µl of reverse transcription (RT) primer (targeting the *SQ528* or *rpoC*), 4 µl of deoxynucleoside triphosphates (10 mM each), 1 µl of RNase inhibitor, and 10 µl of RT buffer in a final volume of 50 µl (Roche). Samples were heated at 80°C for 5 min, then slowly cooled. cDNA synthesis was carried out at 37°C for 2 h using avian myeloblastosis virus reverse transcriptase (Roche).

Real-time quantitative PCR was performed in a 20 µl reaction containing 20 ng of cDNA, 10 µl of SYBR Green PCR Master Mix (Applied Biosystems), and 400 nM gene-specific primers. Amplification and detection were conducted using the LightCycler 480 system. For each sample, the quantity of cDNA corresponding to the sRNA *SQ528* was normalized to the quantity of cDNA for the *rpoC* gene. The primers used for each marker are listed in Supplementary Table S4. Relative transcript changes were calculated using the 2^−ΔΔCt method, as previously described [35]. Statistical analysis was carried out using GraphPad Prism (version 7.0; GraphPad Software Inc.).

In vitro transcription

For *in vitro* transcription of *SQ528*. The corresponding primers *SQ528*_Fw_T7/*SQ528*_Rev_T7 were used for template generation via high fidelity Phusion DNA polymerase (from Thermo Fisher Scientific). PCR products were purified and concentrated using the NZYGelpure (Nzytech) to prevent the production of side products during *in vitro* transcription. *In vitro* transcription was performed using the HiScribeTM T7 Hight Yield RNA Synthesis Kit (from New England Biolabs) in 20 µl reactions according to the manufacturer's protocol, followed by DNase treatment (2 U TURBO DNase for 30 min at 37°C). RNA products were purified and concentrated using RNeasy® Mini Kit (Qiagen). Purified RNA was eluted in 30 µl of RNase-free water and stored at −80°C. Resulting RNA fragments were separated on a denaturing urea–polyacrylamide gel electrophoresis (urea–PAGE) with 6% polyacrylamide and 8 M urea, followed by ethidium bromide (Carl Roth) staining for 10 min and imaging using an Gel Doc system. Bands of correct size were cut out in small pieces and transferred into 2 ml tubes. For RNA elution, 750 µl RNA elution buffer (0.1 M NaAc, 0.1% sodium dodecyl sulfate, 10 mM EDTA) was added, and the samples were incubated at 4°C and 1000 rpm overnight. Following centrifugation at 5000 *g* at 4°C for 1 min, the supernatants were transferred to new tubes and RNA extraction was performed using a single phenol-chloroform extraction step (phenol/chloroform/isoamyl alcohol). Purified RNA was resuspended in 20 µl RNase-free water and stored at −80°C.

Ribonuclease activity assays

Activity assays were performed by incubating 5 µM of RNA either alone or with 0.5 µM of CD25890 proteins, for different times (0, 15, 30, 60, and 90 min). All components were mixed to a final concentration of 1× structure buffer (20 mM HEPES, pH 8, 40 mM NaCl, and 10 mM KCl). Manganese or other bivalent ion (when indicated) were added to 10 mM. Resulting RNA fragments were separated on a denaturing urea–PAGE with 6% polyacrylamide and 8 M urea, followed by ethidium bromide (Carl Roth) staining for 30 min and imaging using an Gel Doc system. The stained acrylamide gels were scanned and the amount of RNA at each time point was quantified using ImageJ *1*.37v [36] and the activity defined as $(1 - [SQ528]t_n/[SQ528]t_0) \times 100$.

Fluorescence endoribonuclease assay

A dual-labelled fluorogenic RNA substrate, 5′-6-FAM-dArUdAdA-6-TAMRA-3′, was synthesized and HPLC-purified (custom-ordered from Metabion). The probe includes a 6-carboxyfluorescein (6-FAM) fluorophore at the 5′ end and a 6-carboxytetramethylrhodamine (TAMRA) quencher at the 3′ end. Endoribonuclease activity was monitored in a total volume of 20 µl containing 50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 10 mM $MnCl_2$. The fluorogenic RNA probe was used at a final concentration of 100 nM. CD25890 was added at the indicated concentrations to initiate the reaction. Fluorescence was monitored in real-time at 37°C using a fluorescence microplate reader (LightCycler 480) with excitation at 485 nm and emission at 528 nm. Cleavage of the probe separates the 6-FAM fluorophore from the TAMRA quencher, resulting in increased fluorescence. Fluorescence intensity was recorded every 30 s.

Fluorescence *in situ* hybridization

Cells were collected from cultures growing in SM for 10 h in the presence of aTc and fixed in Histochoice solution (VWR) for 15 min at room temperature and then for 30 min on ice. The samples were centrifuged three times at 3000 × *g* for 2 min and washed in PBS. The cell pellets were resuspended in 100 µl GTE buffer (50 mM glucose, 20 mM Tris–HCl, pH 7.5, 10 mM EDTA, pH 8.0). Sixteen microliters of a 10 mg/ml mutanolysin solution (prepared in GTE buffer containing 4 mM vanadyl ribonucleoside complex and 10 mg/ml mutanolysin) were added to 48 µl of the cell suspension. The mixture was incubated for 10 min at room temperature. The samples were then centrifuged at 3000 × *g* for 2 min and gently resuspended in 100 µl of GTE. The suspension was immediately applied to 0.1% poly-L-lysine-coated multi-well slides and incubated for 10 min at room temperature. Excess liquid was aspirated, and the cells were allowed to air dry completely. Once the slides were dry, a 100 µl drop of 0.1% Triton X-100 was added to the cells and incubated for 5 min at room temperature. The excess liquid was aspirated, and the cells were allowed to air dry completely. The slides were then washed twice with 100 µl of 2 × SSCT buffer (300 mM sodium chloride, 30 mM sodium citrate, 0.1% Tween-20, pH 7.0) for 5 min each at room temperature in a humid chamber. The slides were subsequently incubated in a solution of 50% formamide and 2 × SSCT for 30 min at 37°C in a humid chamber. The probes (20ng/µl) were added to the hybridization solution (composed of 1.75 g sodium chloride, 0.882 g sodium citrate, 1 g dextran sulfate, and 5.5 ml formamide, brought to a final volume of 10 ml with ddH_2O), which was then incubated at 95°C for 5 min and centrifuged at 16 000 × *g* for 1 min. Slides were placed on a heat block at 95°C, and 25 µl of the hybridization solution was added. The slides were incubated for 4 min, then coverslips were applied and the slides incubated in a humid chamber for 12 h at 42°C. After hybridization, slides were washed in 25% formamide and 2 × SSCT at 37°C for 30 min, followed by three washes in 2 × SSCT at room temperature for 5 min each. The slides were then briefly rinsed five times in PBS. Fluoromont-G™ solution (Invitrogen) was added to each well, and the slides were covered and sealed with clear nail polish. The slides were either visualized immediately or stored in the dark at –20°C. The sequences of the Cy5 labelled probes (FISH_SQ528_Cy5_1 and FISH_SQ528_Cy5_2) are given in Supplementary Table S4.

Fluorescence microscopy and image analysis

Samples were withdrawn from SM cultures at the desired times following inoculation, and the cells collected by centrifugation (10 min, 4000 × *g*, at 4°C). The cells were washed with 1 ml of phosphate-buffered saline (PBS; 137 mM NaCl, 10 mM Phosphate, 2.7 mM KCl, pH 7.4), and resuspended in 0.1 ml of PBS supplemented with the lipophilic styryl membrane dye *N*-(3-triethylammoniumprpyl)-4-(*p*-diethylaminophenylhexatrienyl) pyridinium dibromide (FM4-64, Molecular Probes, Invitrogen; 10 µg ml^{-1}) [37]. Images were acquired using a Leica DM6000B upright microscope equipped with an Andor iXon 885 EMCCD camera and controlled with the MetaMorph V5.8 software. The images were acquired with a 100 × 1.4 numerical aperture (NA) immersion objective and a 1.6× optovar, the fluorescence filter sets TX2 and phase contrast optics.

Cell segmentation was conducted using Cellpose (version 3.1.1.1), a generalist deep learning-based algorithm for cellular segmentation [38]. Depending on batch size, the software was operated via its Python-based graphical user interface or command-line interface. The estimated cell diameter was set to 30 pixels, based on empirical measurements and Cellpose's automatic estimation. Input images were processed in greyscale mode. Quantitative image analysis, including pixel-based measurements and downstream processing, was conducted using Fiji/ImageJ [39]. Statistical analysis was carried out using GraphPad Prism (version 7.0; GraphPad Software Inc.).

RNA-seq data analysis

RNA extraction and RNA-sequencing was performed as described previously [14]. Sequencing reads were aligned with bwa to the genome of the *C. difficile* strain 630Δ*erm* (CP016318.1 and CP016319.1). Tables of read counts per gene were obtained using featureCounts with the Genbank gene annotations, extended with an extra set of sRNA annotations from [40]. Differential expression analysis was performed using the edgeR R package. Namely, count data was normalized using the TMM normalization [41], and moderated bayesian statistics [42] was applied to obtain differentially expressed genes between conditions. For visualization purposes, alignments were used to estimate normalized counts per million (CPM) in 10 bp windows using the bamCoverage tool from Deeptools.

Results

sRNA SQ528 accumulates in a *CD25890* deletion mutant

It is likely that, similar to other members of the YicC-family, CD25890 also participates in sRNA regulation. Previous research has identified sRNAs involved in initiating sporulation that modulate *spo0A* translation [43]. Thus, we investigated whether CD25890 regulates the activity of these or other sRNAs whose functions remain unidentified. In our previous study we compared the transcriptome between Δ*CD25890* mutant and WT cultures after 10 h of growth in SM, a medium that induces sporulation [14]. We utilized this data set to identify small, noncoding RNAs [40, 44] differentially expressed between the Δ*CD25890* mutant and WT. From this new analysis, we identified 11 sRNAs that accumulated at higher levels in the mutant (*SQ528*,

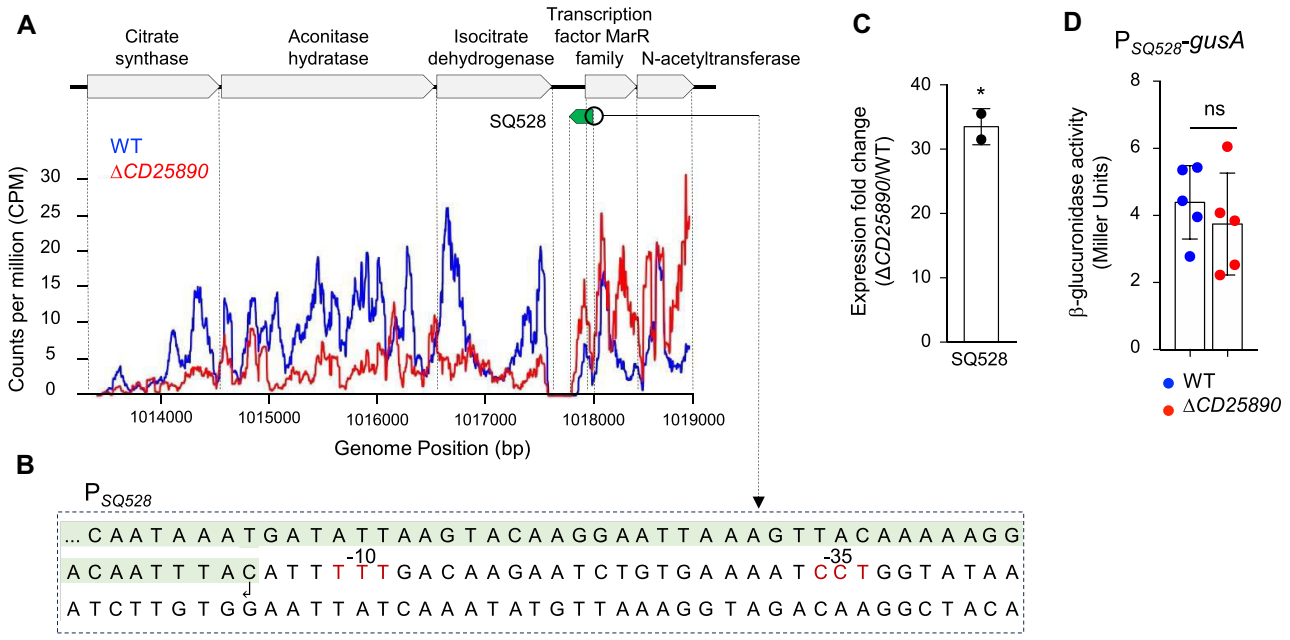


Figure 1. The sRNA *SQ528* is more abundant in a *CD25890* mutant. (**A**) RNA-Seq transcriptomics data on the *SQ528* genomic region in WT and *CD25890* mutant, visualized as normalized CPM per 10 bp windows. The sRNA *SQ528* (in green) is antisense to the 5′ end of the gene *CD08350*, coding for a transcription factor of the MarR family. Upstream of *CD08350* are the three genes coding for the enzymes involved in the oxidative part of the Krebs cycle (citrate synthase, aconitase and isocitrate dehydrogenase). (**B**) *SQ528* promotor region. A putative σ^H promoter on is represented. In red are the –35 and –10 consensus sequence recognized by *C. difficile* σ^H [45] and in green the annotated 5′ terminus of *SQ528*. (**C**) Quantification of the expression of SQ528 by qRT-PCR. Total RNAs were extracted from *C. difficile* 630Δ*erm* and Δ*CD25890* strains grown in SM for 10 h. The graph shows the fold change of the differential expression of *SQ528* between the Δ*CD25890* and WT. Dots correspond to the mean of four technical replicates in two biological replicates. Error bars correspond to standard deviations from two biological replicates. **P* < .05 by by a Student's *t*-test. (**D**) Expression of P_{SQ528}-*gusA* was monitored in the WT and *CD25890* mutant. Strains were grown for 10 h in SM, samples were taken and assayed for β-glucuronidase activity. Endogenous levels of β-glucuronidase activity were determined in a WT strain carrying a plasmid with a promoter less *gusA* and subtracted from the values obtain in the other strains. Error bars correspond to standard deviation from five biological replicates. Differences between the WT and *CD25890* means are nonsignificant (ns) by a Student's *t*-test.

CD630_n00210, CD630_s0591, *SQ2591*, *SQ1001*, *SQ758*, CD630_n00820, CD630_n00760, CD630_n00710, *SQ1366*, and *SQ2284*; Supplementary Table S1).

Since the hyper-sporulating phenotype of the Δ*CD25890* mutant is only seen under certain nutritional conditions, we chose *SQ528* (with a fold change of ~6) for further analysis, as it is located in the intergenic region between the operon coding for the enzymes of the oxidative half of the Krebs cycle and a transcriptional factor from the MarR family (Fig. 1A). To validate the RNA-seq results, we performed quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR) using RNA isolated from the Δ*CD25890* mutant and the WT (Fig. 1C). The average transcript levels for *SQ528* was increased (~30-fold) in the Δ*CD25890* mutant relative to the WT in agreement with the RNA-seq data. To determine whether increased *SQ528* levels arise from reduced degradation or increased transcription, we constructed a transcriptional fusion of the promoter region of the sRNA (218 bp; Fig. 1B) to the reporter gene *gusA*, which codes for β-glucuronidase. Activity assays were done in samples collected from cultures after 10 h of growth in SM showed that *SQ528* is transcribed at similar levels in the WT and Δ*CD25890* mutant. Thus, *CD25890* does not significantly affect transcription of *SQ528* (Fig. 1D). Given that our transcriptome data show higher levels of *SQ528* in the absence of CD25890, this suggests that the sRNA *SQ528* is stabilized, rather than more actively transcribed, when CD25890 is absent. We manually searched for a possible promoter located upstream of the annotated *SQ528*, and identified a sequence consistent with the consensus for promoters recognized by SigH (Fig. 1B [45]). SigH is an alternative sigma factor involved in the transition to the post-exponential growth phase in C. *difficile*, which is consistent with our observation of *SQ528* expression after 10 h of growth in SM [45].

Overexpression of the sRNA *SQ528* phenocopies a *CD25890* mutant

Since CD25890 appears to play a direct or indirect role in the degradation of *SQ528*, we reasoned that *SQ528* accumulation could be responsible for the Δ*CD25890* phenotype. If so, overexpressing *SQ528* would trigger an increase in sporulation, as observed in the *CD25890* deletion mutant. The *CD25890* deletion mutant exhibited enhanced sporulation efficiency compared to the WT strain when cultured in SM, but not when grown on 70:30 agar plates. Both media are known to promote sporulation in C. *difficile* [14, 22]. To overexpress *SQ528*, we fused the sRNA to the anhydrotetracycline (aTc) inducible promoter (P_{tetA}) in a multicopy plasmid, and introduced it in the WT and Δ*CD25890* strain. As a control, we used strains carrying a plasmid with only the P_{tetA} promoter (empty plasmid [24]). To confirm *SQ528* overexpression, we performed fluorescence *in situ* hybridization (FISH) using two Cy5-labelled antisense RNA probes specific to *SQ528* in both WT and mutant strains grown in SM for 10 h (Supplementary Fig. S1). FISH imaging and quantification revealed increased fluorescence intensity in the Δ*CD25890* strain compared to the WT, both in strains car-

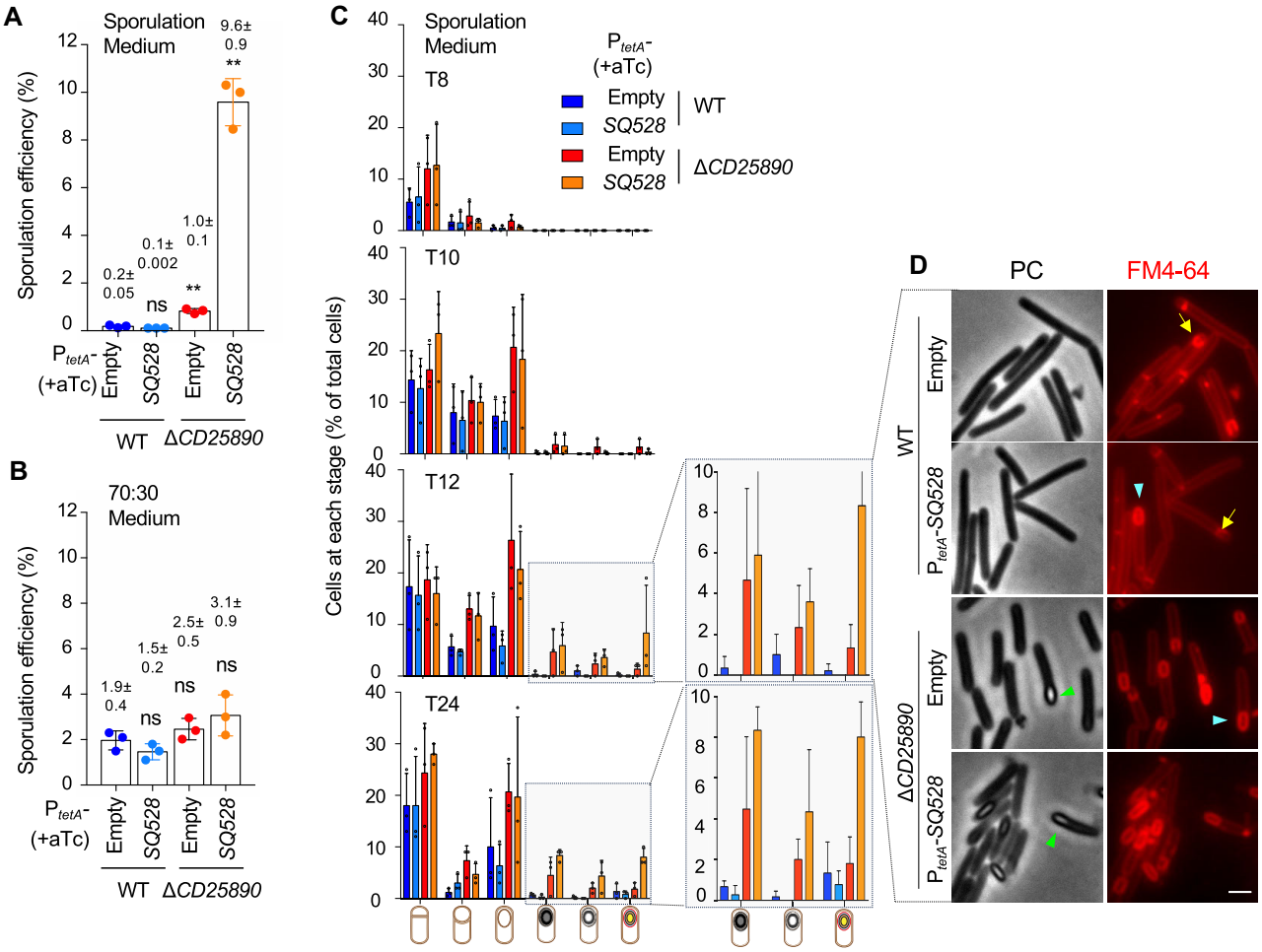


Figure 2. Overexpression of sRNA *SQ528* increases sporulation. Sporulation efficiency of the WT and the *CD25890* mutant carrying P_{tetA}-*SQ528* or P_{tetA}-*empty*. Strains were grown in SM (**A**) or 70:30 medium (**B**) for 24 h and the cultures were plated in BHI medium supplemented with 0.1% TA before (to estimate total CFUs/ml) and after heat treatment at 70°C for 30 min (spores/ml). The results shown are averages and standard deviations for three biological replicates. Asterisks indicate statistical significance determined with a Student's *t*-test (** $P < .01$). (**C**) Samples of an SM culture of the WT strain with P_{tetA}-*SQ528* and Δ*CD25890* mutant strain with P_{tetA}-*SQ528* were collected at 8, 10, 12, and 24 h after inoculation, stained with the membrane dye FM4-64 and examined by phase-contrast or fluorescent microscopy. The stages in the sporulation pathway are represented schematically at the bottom of the panel: asymmetric division of the sporangium, intermediate stage in the process of engulfment of the forespore by the larger mother cell; engulfment completion, isolating the forespore from the surrounding medium, synthesis of the spore surface layers, the spore cortex, and the coat and exosporium. Percentage of cells in the morphological classes represented relative to the total viable cell population is quantified. At minimum of 150 cells were counted at each hour for each strain in each biological replicate. The data represent the mean ± standard deviation of three independent experiments. The panels on the right, showing data for 12 and 24 h after inoculation, have an extended *y*-axis scale compared to the graphs in the panels on the left. (**D**) Phase-contrast and fluorescence images of sporulating cells from the indicated strains (see also Supplementary Fig. S1). The cells were collected after 12 h of growth in SM and stained with the membrane dye FM4-64. Yellow arrows indicate asymmetric division of the sporangium or an intermediate stage in the engulfment of the forespore by the mother cell; blue arrows indicate sporangia in which engulfment is completed, and green arrows highlight sporangia of mature spores. Scale bar, 1 μm.

rying the empty vector, supporting the RNA-seq findings (see above), and in those overexpressing *SQ528*. Furthermore, fluorescence intensity was consistently higher in strains harboring the *SQ528* overexpression plasmid (Supplementary Fig. S1). Strains were grown in both SM and 70:30 plates, each supplemented with 50 nM of aTc, and the percentage of heat-resistant spores was estimated after 24 h of incubation. As reported previously [14], the Δ*CD25890* mutant sporulates nearly five times more than the WT in SM (empty plasmid control in Fig. 2A). Induction of *SQ528* expression in a WT background has no impact on the efficiency of sporulation but it increases the sporulation efficiency by 10-fold in the mutant background (Fig. 2A). This result suggests that, in a WT background, enough CD25890 is present in the cell to deal with the increased expression of *SQ528* from the *tetA* promoter. Although the difference in fluorescence intensity detected by FISH between the *SQ528*-overexpressing strain and the Δ*CD25890* mutant was modest, it appears to be sufficient to account for the enhanced sporulation phenotype observed in the Δ*CD25890* mutant (Supplementary Fig. S1). In contrast, in the mutant, accumulation of *SQ528* leads to an increase in sporulation efficiency. As anticipated [14], on 70:30 plates all the strains sporulated with similar efficiency (Fig. 2B).

We then proceeded to track the morphological stages of sporulation in SM using phase-contrast and fluorescence microscopy (Fig. 2C and D and Supplementary Fig. S2). Both WT and mutant strains were grown in SM, with samples collected

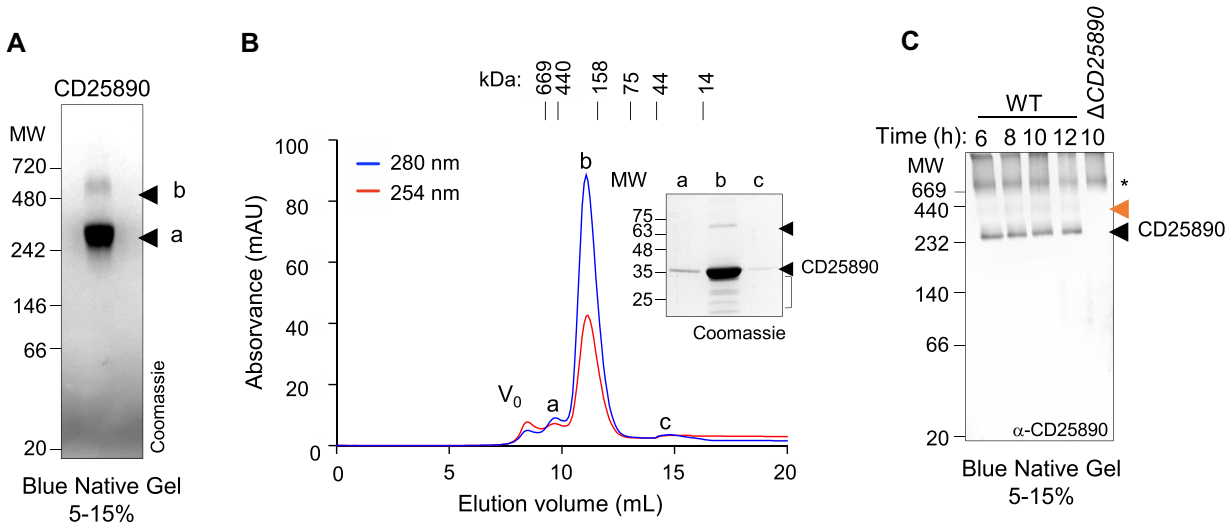


Figure 3. CD25890 is a hexamer in *in vitro* and *in vivo*. (**A**) BN–PAGE of affinity purified His_6-CD25890. The two observed species (a and b) are indicated. MW markers are shown in kDa. (**B**) SEC of affinity purified His_6-CD25890 on a S200 10/300 GL column. On the top is indicated the position of the elution peaks of MW standards (thyroglobulin, 669 kDa; ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 44 kDa; ribonuclease A, 14 kDa). The insert shows SDS–PAGE analysis of the peaks (a, b, and c). His_6-CD25890 is indicated by the black arrow and brackets represent degradation products. (**C**) BN–PAGE with extracts of the WT and the *CD25890* mutant. Strains were grown in SM and samples were collected at 6, 8, 10, and 12 h after inoculation for western blot analysis using anti-CD25890 antibody (*, cross-reactive species). A presumed hexameric species (approximately 232 kDa) and a less abundant species of CD25890, which, based on its apparent size (approximately 440 kDa), may represent a complex with PNPase, are indicated by arrows. The position of MW markers (in kDa) is indicated on the left side of the panels.

at 8, 10, 12, and 24 h post-inoculation. Prior to microscopic analysis, the cells were stained with the membrane dye FM4-64. Over time, we observed a higher percentage of sporulating cells in the *CD25890* mutant compared to the WT, as expected [14]. Consistent with the heat resistance test results (Fig. 2A), overexpression of *SQ528* did not affect sporulation progression in the WT background. However, when *SQ528* was expressed from the P_{tetA} promoter in the *CD25890* mutant background, an increased number of cells completing asymmetric division were observed by hour 10, and phase-bright spores accumulated at higher levels by hour 12 and 24. (Fig. 2C and D and Supplementary Fig. S2). These findings indicate that, while CD25890 inhibits the initiation of sporulation, *SQ528* appears to be also involved in later stages of the process.

CD25890 is a hexamer *in vitro* and *in vivo*

YicC from *E. coli* and YloC from *B. subtilis* were shown to form hexamers in solution, although the oligomeric state of the proteins *in vivo* was not reported [16–18]. To uncover the oligomeric state of CD25890, the full-length protein, His_6-$CD25890^{FL}$, was produced in *E. coli* BL21 (DE3) and purified over a Ni^{2+}-column, as previously described [14]. The final yield of purified His_6-CD25890 (estimated MW 35.4 kDa) was approximately 3 mg/l of culture. Oligomer formation by His_6-CD25890 was tested using BN–PAGE and size exclusion chromatography (SEC). Using BN–PAGE, two species were observed that exhibited MWs compatible with a hexameric or heptameric form (*a*, around 245 kDa, the most represented species) and a higher-molecular-weight species of near 586 kDa (*b*, approximately twice the MW of *a*) (Fig. 3A). When the His_6-CD25890 protein was eluted through a SEC, at an initial concentration of 10 μM, three distinct peaks were observed (*a*, *b*, and *c* in Fig. 3B) all of which contained the protein (insert in Fig. 3B). Comparison of the elution volumes relative to marker proteins filtered through the same column indicated a molecular mass of 35.3 kDa for peak *c*. This peak presumably represents monomeric His_6-CD25890. The calculated mass for the larger peak (*b*) was 216.3 kDa, which is consistent with formation of hexamers. Quantification of the peak area indicated that 89% of total protein was hexameric. Peak *a* has a species with a calculated size of 360.6 kDa which most likely results from a complex with nucleic acids, as shown by the absorbance at 254 nm; this species may correspond to the band of higher MW observed by BN–PAGE (Fig. 3A, *b*). The differences in the sizes between the BN–PAGE and SEC analysis may result from the presence of dodecyl maltoside in the samples separated by BN–PAGE which increases apparent MWs. The re-injection of the major peak (*b*) in a SEC column revealed only one peak corresponding to the same size, showing that the His_6-CD25890 oligomer is stable.

We further investigated the oligomeric state of CD25890 *in vivo* during growth and sporulation in SM. Immunoblot analysis of *C. difficile* extracts resolved by BN–PAGE shows that CD25890 accumulates as a species with a MW compatible with a hexamer (234 kDa), both during growth (6–8 h after inoculation) and when cells start to sporulate (10 h after inoculation; Fig. 3C). A less abundant band just below the 440 kDa marker may correspond to CD25890 in complex with nucleic acids (corresponding to peak *a* in Fig. 3B). Alternatively, it may represent a complex of CD25890 with other proteins, possibly with PNPase (see below). These findings suggest that CD25890 is a hexamer *in vivo* and that the hexamer is the physiologically relevant form of the protein.

The N-terminal domain of CD25890 is responsible for oligomerization

Sequence alignment shows that the YicC-family of proteins have N- and C-terminal conserved domains linked by a less conserved sequence (Fig. 4A and Supplementary Fig. S3). The C-terminal domain includes the DUF1732 domain that is pre-

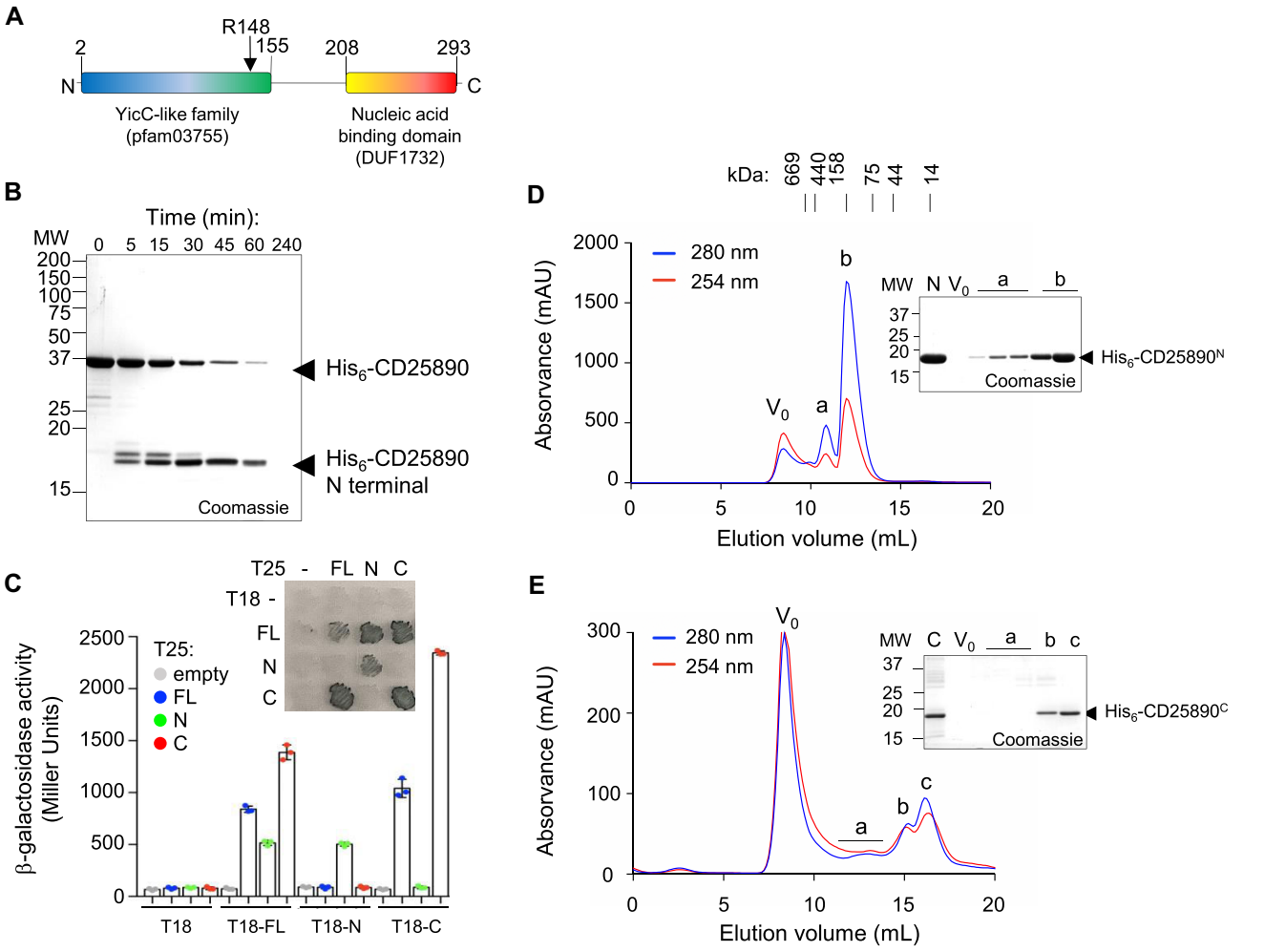


Figure 4. CD25890 is a two-domain protein. (**A**) Schematic representation of the organization predicted for the CD25890 protein. The different domains are labelled: the N-terminal YicC-like family domain and at the C-terminal the DUF1732 domain that is predicted to bind to nucleic acids. (**B**) Limited trypsin digestion of His_6-CD25890 along time. Positions of bands corresponding to His_6-CD25890 and the N-terminal domain, as determined by MALDI-TOF mass spectrometry, are indicated. MW markers are shown in kDa. (**C**) Bacterial Adenylate Cyclase Two-Hybrid (BACTH) analysis of protein–protein interactions. Cultures of transformed BTH101 expressing the indicated fusion proteins were grown in liquid medium until mid log and β-galactosidase activity measured. Results are presented as means ± standard deviation from experiments carried out in biological triplicate. The insert shows colony growth on solid plates for the different combinations of $CD25890^{FL}$, $CD25890^{N}$, and $CD25890^{C}$ proteins, with blue colonies indicating interaction between the different proteins or domains. (**D**, **E**) SEC of affinity purified His_6-$CD25890^{N}$ (residues 1–151) and His_6-$CD25890^{C}$ (residues 152–293) on a S200 increased 10/300 GL column. The insert shows SDS–PAGE analysis of the peaks indicated in the graph. On the top of panel C is indicated the position of the elution peaks of MW standards (thyroglobulin, 669 kDa; ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 44 kDa; ribonuclease A, 14 kDa).

dicted to bind nucleic acids (residues 208–293 [46]). To examine the domain structure of CD25890, we subjected purified His_6-$CD25890^{FL}$ to limited proteolysis with trypsin (Fig. 4B). After 5–15 min of incubation, trypsin cleaved His_6-$CD25890^{FL}$ into 17 and 18 kDa fragments; after longer incubations, (at least 1 h), one stable fragment of approximately 17 kDa accumulated (Fig. 4B). After 4 h of digestion His_6-$CD25890^{FL}$ is completed degraded. We then subjected the 17 kDa His_6-CD25890 fragment to mass spectrometry analysis. Given the size of this fragment, the specificity of trypsin for lysine and arginine residues, and the mass spectrometry sequence information, this fragment corresponds to the N-terminal of the protein, with the C-terminus of this fragment predicted to be located at arginine 148 (Fig. 4A and B).

To gain insight into the domain involved in the oligomerization of His_6-$CD25890^{FL}$ we first tested the ability of each domain to self-interact using the BACTH analysis [20]. For this purpose, full-length CD25890 ($CD25890^{FL}$), CD25890 from residue 1–151 ($CD25890^{N}$) and from residue 152–293 ($CD25890^{C}$) were individually fused to the C-terminus of the T25 fragment and to the N-terminus of T18 of the adenylate cyclase (*cyaA*) catalytic domain (Fig. 4C). The different combinations of $CD25890^{FL}$ with $CD25890^{N}$ and $CD25890^{C}$ were tested for interactions. In this system, association between the N- and C-terminal moieties of adenylate cyclase results in production *c*-AMP and β-galactosidase produced from the *lacZ* reporter gene [20]. The interactions between full-length CD25890 and separate domains were quantified by measuring the activity of β-galactosidase. Our BACTH analysis confirmed that $CD2589^{FL}$ interacts with itself (Fig. 4C; 840.55 ± 30.21 Miller units, MU). Moreover, we detected self-interaction of $CD25890^{N}$ (502.92 ± 19.48 MU) and $CD25890^{C}$ (2343.86 ± 19.77 MU), and interaction of each domain with $CD25890^{FL}$ in at least one orientation

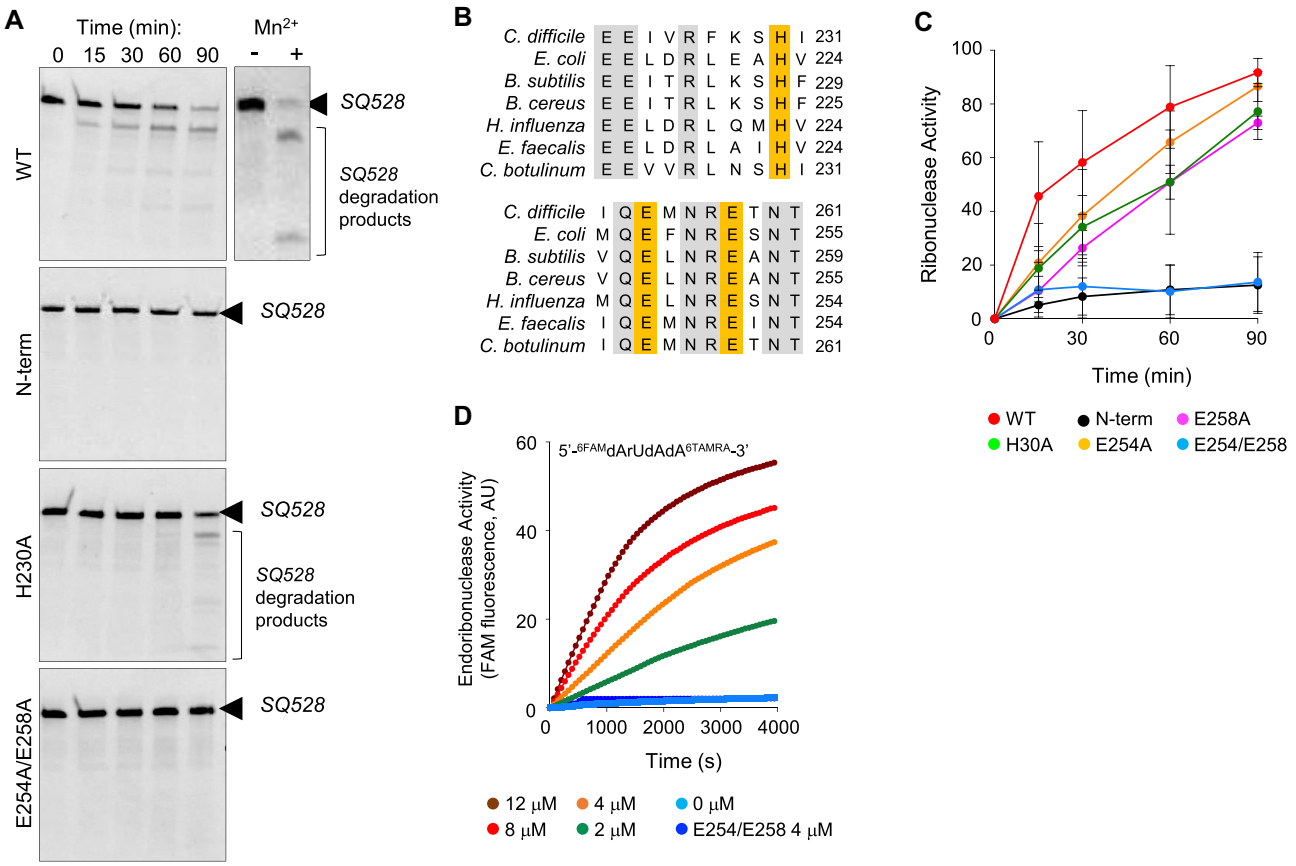


Figure 5. Ribonucleolytic activity of CD25890. (**A**) Time dependent cleavage of *SQ528* sRNA by CD25890^{FL} (WT), CD25890^{H230A}, CD25890$^{E254A/E258A}$, and CD25890^{N}. Resulting RNA fragments were separated on a 8 M urea–6% polyacrylamide gel followed by staining with ethidium bromide for 30 min and imaging using an Gel Doc system. Quantification of these assays (means ± standard deviation, of at least three replicates) are shown in panel **(C)**. (**B**) The alignments show highly conserved residues in the vicinity of H230, E254, and E258 shown before to be important for the activity of YloC from *B. subtilis* and YicC from *E. coli*. (**D**) Time-dependent cleavage of the dual-labelled oligonucleotide 5′-6-FAM- dArUdAdA-6-TAMRA-3′ by different concentrations of CD25890^{FL} or CD25890$^{E254A/E258A}$. In the intact substrate, FAM fluorescence is quenched by TAMRA via Förster Resonance Energy Transfer (FRET). Cleavage at the rU site separates the fluorophores, resulting in increased 6-FAM fluorescence, which serves as a readout of endoribonuclease activity.

(Fig. 4C). An interaction of CD25890^{FL} with CD25890^{N} was only was observed when it was fused to T25 fragment, rather than when fused to the T18 fragment (Fig. 4C). The detection of interactions in only one direction suggests that, in this context, the orientation of the proteins has a significant influence on protein–protein interactions. No β-galactosidase activity was detected when the plasmids containing the different domains of CD25890 fused to T25 or T18 were paired with plasmids expressing T18 or T25 alone (Fig. 4C). Thus, both domains of CD25890 have the ability to self-interact.

To assess the oligomeric state of CD25890^{N} (residues 1–151) and CD25890^{C} (residues 152–293) both domains were fused to a His$_6$-tag, produced in *E. coli* BL21 (DE3) and purified over a Ni^{2+}-column [14]. While a yield of 5 mg/l of culture was obtained for CD25890^{N} (estimated MW 19 kDa), CD25890^{C} (estimated MW 18 kDa) was mainly insoluble with only 0.7 mg/l of culture being recovered in the soluble fraction. SEC of purified His$_6$-CD25890^{N} revealed two main peaks (*a* and *b* in Fig. 4D) both of which contained the protein (insert in Fig. 4D). The larger peak (*b*) corresponds to a species with a calculated size (125 kDa), similar to that of a His$_6$-CD25890^{N} hexamer (Fig. 4D). The smaller peak (*a*) corresponds to a species with a calculated size of 229 kDa, which may correspond to two hexamers (Fig. 4D). SEC of purified His$_6$-CD25890^{C} revealed two main peaks that contain protein (*b* and *c* in Fig. 4E). Peak *c* corresponds to a monomer (19.1 kDa) and peak *b* to a possible dimer (26.3 kDa). No protein was detected in peak *a* and it may correspond to nucleic acids given the higher optical density (OD) at 254 nm observed for this peak. These results suggest that the N-terminal domain alone is able to form a hexamer, and that it may be the domain responsible for the oligomerization of CD25890.

CD25890 shows metal-dependent ribonuclease activity

The *B. subtilis* and *E. coli* members of the YicC-family, YloC, and YicC, have Mn^{2+}-dependent endoribonuclease activity [16–18]. To test whether CD25890 has ribonuclease activity, the His$_6$-CD25890^{FL} protein, purified by Ni^{2+}-affinity chromatography followed by SEC, was incubated with *SQ528*, the sRNA previously identified by RNAseq as a possible CD25890 substrate. The reaction mixture, which included Mn^{2+}, was resolved by electrophoresis in a polyacrylamide gel stained with the nucleic acid dye ethidium bromide. Incubation results in accumulation of a smaller size product over time, concomitant with the disappearance of *SQ528* (Fig. 5A). This intermediate could result from 3′ exonuclease processiv-

ity, or it could result from endonuclease activity. We favour the latter hypothesis since orthologues YloC and YicC were previously shown to

have endonuclease activity [15, 17, 18]. To test this, we measured CD25890 activity as a function of enzyme concentration using a real-time fluorescence-based assay (Fig. 5D). Using a target RNA with an internal cleavage site flanked by a fluorophore and a quencher is a sensitive method for detecting endoribonuclease activity [16, 47]. In this assay, we used a tetranucleotide substrate, 5′-6FAM-dArUdAdA-6TAMRA-3′. Cleavage at the single riboU releases the fluorophore from the FRET quencher, allowing detection of fluorescence in real time. CD25890 was added to the reaction at concentrations ranging from 0 to 12 μM, and the rate of cleavage was measured over 4000 s (Fig. 5D). Activity increased with enzyme concentration but not when a catalytically inactive allele was added to the reaction (see below). These results indicate that CD25890 is indeed an endoribonuclease.

As observed for YloC and YicC, cleavage was not observed when Mn^{2+} was absent from the reaction (Fig. 5A [16–18]). Similar to YloC, CD25890 was active in the presence of Mg^{2+} and Co^{2+}, but not in the presence of Ca^{2+} (Supplementary Fig. S4A, upper panel). In the presence of Zn^{2+}, the sRNA *SQ528* was completely degraded; however, this degradation was also observed in the absence of the protein (Supplementary Fig. S4A, bottom panel), suggesting that it may result from zinc-mediated RNA cleavage, as previously reported [48, 49]. In contrast to YloC, we were able to detect CD25890 activity in the presence of Cu^{2+} [17]. The differences observed may stem from the use of different substrates between our study and previous ones (26-nt oligonucleotide [17]). In any case, our results indicate that CD25890 is a metal-dependent ribonuclease.

The C-terminal domain of CD25890 harbours residues essential for ribonuclease activity

We also tested whether $CD25890^{N}$ alone has activity using *SQ528* as a substrate. We were not able to detect any ribonuclease activity, suggesting that the C-terminal part of the protein contains at least some of the residues directly involved in catalysis (Fig. 5B and C). The highly conserved H230, E254, and E258, residues located in the C-terminal domain were previously shown to be important for YloC and YicC activity [16–18] (Fig. 5B). We produced and purified CD25890 alleles carrying substitutions of the conserved residues H230, E254 and E258 to alanine, and cleavage activity on *SQ528* was assessed as described above. $CD25890^{H230A}$, $CD25890^{E254A}$, and $CD25890^{E258A}$ still show activity although reduced relative to the WT form (Fig. 5B and C, and Supplementary Fig. S4B). In contrast, no activity was detected for $CD25890^{E254A/E258A}$ (Fig. 5B and C). Alanine substitutions of H228 and E256 of YloC, homologous to H230 and E258 of CD25890 (Fig. 5B), were also observed to reduce ribonuclease activity [17].

The different forms of CD25890 may have reduced activity due to impaired substrate binding, impaired self-interaction to form the hexamer or because the residues changed to alanine play direct roles in catalysis. SEC was used to test the oligomeric state of the different variants of His_6-CD25890. The elution profiles for all the variants were similar to that of the WT with approximately 89% of the protein eluting as hexamers (Fig. 6A). Given these results, loss of activity is unlikely to be due to impaired hexamer formation. We used electrophoretic mobility shift assay (EMSA) assays to test the ability of the different variants of CD25890 to bind *SQ528*. The formation of the complexes was visualized in a 6% native polyacrylamide gel electrophoresis (PAGE) stained with ethidium bromide (Fig. 6B and Supplementary Fig. S4C). Remarkably, although all forms of CD25890 interacted with *SQ528*, the inactive variants $CD25890^{E254A/E258A}$ and $CD25890^{N}$ bound more strongly to *SQ528* than did full-length CD25890 (Fig. 6B). These results suggest that the active complex exhibits a dynamic behaviour, while the inactive enzyme stabilizes the complex by forming a stronger interaction with the RNA substrate.

Cryo-EM single-particle analysis shows a dynamic CD25890 hexamer

Due to the considerable size of the hexamers, it becomes feasible to visualize individual CD25890 particles using electron microscopy. Negative staining of the WT protein isolated from the SEC column was performed (peak *b*, Fig. 3B) and revealed individual particles. Thus, the sample was submitted to CryoEM analysis; the electron micrographs (Fig. 7A) reveal a monodisperse population of well-separated single particles. The 2D particle classification of the single-particle analysis revealed top and bottom projections of a clearly defined donut-shaped, hexameric structure (Fig. 7B). Side projections were almost absent, suggesting a tendency to adopt a preferential orientation within the ice layer.

In our pursuit of structural insights into the CD25890 hexameric protein, we employed AlphaFold-Multimer to generate a structural model (Fig. 7C and D, and Supplementary Fig. S5). In its monomeric form, CD25890 is predicted to be a lengthy bent molecule with a prominent long helix, reminiscent of YicC [15, 18]. The N-terminal domain, is formed by four antiparallel β-strands followed by α-helices, with a long α-helix, from residues 126 to 195 imparting linear characteristics to the molecule, extending to approximately alanine 208 before coiling back to form a coiled-coil structure (Fig. 7D and and Supplementary Fig. S5B). The C-terminal domain adopts a three-helix bundle configuration in an antiparallel conformation, which together with part of the long helix from the N-terminal domain forms a four-helix bundle. The catalytic residues identified thus far (H230, E254, and E258) are located within this domain. The resulting hexameric model has a barrel-like structure with an internal cavity rich in basic residues, which were predicted in *E. coli* YicC to aid in RNA binding (Fig. 7E and Supplementary Fig. S3) [18]. Notably, the N-terminal domain exhibits a ring-shaped appearance when viewed from the top, aligning well with the 2D averages (Fig. 7B and C). However, the lack of side views and the flexibility beyond the hexamerization domain (see C-terminal densities moving around in the 2D averages) prevented from obtaining a reconstruction at higher resolution.

CD25890 forms an active complex with PNPase

Recent research has suggested that *E. coli* YicC may act as an adaptor, directing specific RNAs for degradation by PNPase, and *B. subtilis* YloC was shown to interact with PNPase [15, 17]. To test for an interaction between CD25890 and *C. difficile* PNPase (CD13180), we utilized the BACTH system (as described above; Fig. 8A). Interactions were quantified by assessing β-galactosidase activity. Our BACTH anal-

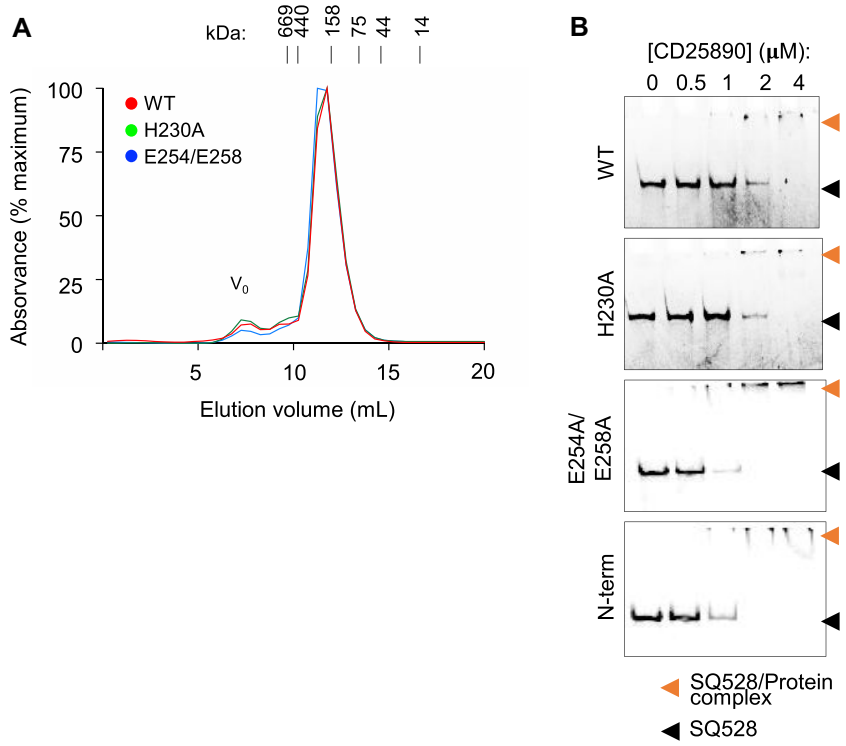


Figure 6. CD25890 binds to *SQ528*. (**A**) SEC of affinity purified His_6-CD25890 His_6-$CD25890^{H230A}$ and His_6-$CD25890^{E254A/E258A}$ on a S200 increase 10/300 GL column. The protein present in the major peak, corresponding to the hexameric form was used in the endoribonucleolytic activity assays. On the top is indicated the position of the elution peaks of MW standards (thyroglobulin, 669 kDa; ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 44 kDa; ribonuclease A, 14 kDa). (**B**) RNA-binding ability of CD25890 WT and different variants to the sRNA *SQ528*. *SQ528* was incubated for 10 min at 37°C with increasing amounts of each recombinant protein (from 0.25 to 8 μM). RNA–protein complexes were resolved in a 6% nondenaturing polyacrylamide gel, and stained with ethidium bromide.

ysis confirmed the interaction between $CD25890^{FL}$ and PNPase (Fig. 8A; 843.04 ± 16.66 MU). No β-galactosidase activity was observed when the plasmid containing PNPase fused to T25 was paired with a plasmid expressing T18 alone or with $CD25890^{C}$. Although light blue colonies were observed with the combination of $CD25890^{N}$ and PNPase, suggesting that the two proteins interacted, no activity was detected when the cells were grown in liquid medium (Fig. 8A). These results suggest that the full-length form of CD25890 is necessary for the interaction.

As an independent test, we produced *C. difficile* PNPase in *E. coli* as a His_6-tagged protein, purified it by Ni^{2+}-affinity chromatography and used SEC to assess its oligomeric state. In SEC, PNPase eluted as a single peak corresponding to a calculated size of 234 kDa (Fig. 8B). This is consistent with the predicted size of a trimer (233.4 kDa) and in line with studies in other species, such as *E. coli*, that also revealed a PNPase trimer in solution [50, 51]. When purified $CD25890^{FL}$ and PNPase were mixed and applied to a SEC column, the proteins co-eluted in a single peak with a calculated size of 445 kDa (Fig. 8B), consistent with the predicted size of 445.8 kDa for the complex. SDS–PAGE confirmed the presence of CD25890, PNPase or both proteins in the peak fractions (Fig. 8B). Thus, the CD25890 hexamer forms a complex with a PNPase trimer. Alongside PNPase, a smaller-sized band is eluted (Fig. 8B), potentially representing a degradation product that continues to form a complex with PNPase and co-purifies with it. An inactive variant, $CD25890^{E254A/E258A}$, is also able to form a complex with PNPase (Fig. 8B). Therefore formation of the complex is independent of CD25890 activity.

We evaluated the catalytic activity of the complex using *SQ528* as a substrate. Reaction mixtures, including Mn^{2+} in all conditions, were subjected to PAGE and visualized with ethidium bromide staining. In contrast to CD25890, which exhibited the accumulation of smaller degradation intermediates over time, PNPase facilitated complete degradation of *SQ528* without the formation of detectable intermediates (Fig. 8C). Notably, when the complex was incubated with *SQ528*, degradation was completed at the earliest time point (Fig. 8C). In an attempt to better establish the enhancement of *SQ528* digestion by the complex, we evaluated catalytic activity at lower protein concentrations (0.3 μM; Fig. 8D and Supplementary Fig. S6). Under these conditions, CD25890 activity was substantially reduced, with only ∼50% of the substrate degraded. PNPase alone, although slower, still achieved complete degradation of *SQ528* after 90 min. This activity was significantly enhanced when the active complex was present, resulting in total degradation of *SQ528* within the first 15 min. In contrast, the catalytically inactive variant $CD25890^{E254A/E258A}$, when in complex with PNPase, severely inhibited PNPase activity. These findings indicate that, although both proteins can independently degrade *SQ528in vitro*, the complex is significantly more efficient. Moreover, PNPase activity appears to depend on CD25890 activity within the complex, reminiscent of dominant-negative effects observed in some enzyme assemblies [52]. This suggests that physical interaction may modulate PNPase activity—possibly by blocking substrate access, altering its conformation, or enabling the sequential action of the two enzymes during RNA degradation.

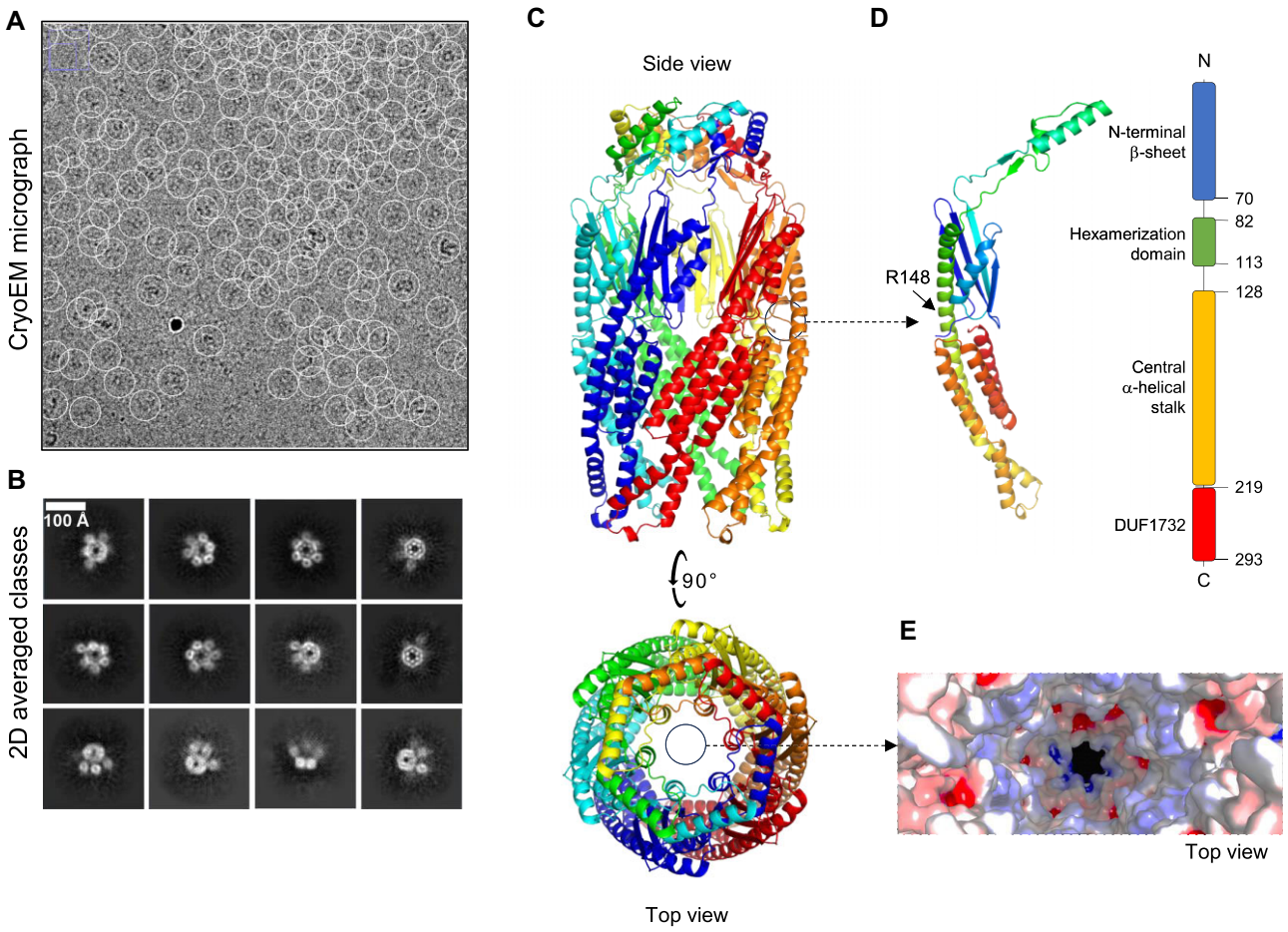


Figure 7. Structural model of the CD25890 hexamer. (**A**) A representative micrograph of the CD25890 specimen by cryo-EM. (**B**) Gallery of representative 2D class averages from a 2D-classification of CD25890 particles. (**C**) An AlphaFold-Multimer model of side view and top view of the hexameric form of *C. difficile* CD25890. The model highlights the different monomers with different colors. (**D**) An AlphaFold-Multimer model of the monomer form of *C. difficile* CD25890. (**E**) View of the internal part of the CD25890 hexamer revealing positive electrostatic surface potential (blue).

Discussion

The YicC-family proteins have gained significant attention across various species, including in model organisms such as *E. coli* and *B. subtilis*. Despite the inability to predict function from their primary sequence, these proteins have been demonstrated to possess endoribonuclease activity and to play a role in sRNA silencing [15, 17, 53]. Here, we present functional and structural characterization of the CD25890 protein of *C. difficile,* which shares 29% sequence identity with YicC from *E. coli* and 40% identity with *B. subtilis* YloC (Supplementary Fig. S3). We have shown previously that CD25890 plays a role in regulating sporulation initiation in *C. difficile* [14]. *C. difficile* is an anaerobic enteric pathogen, where the oxygen-resistant spore acts as both the infectious agent and toxin-delivery vehicle. Toxins produced by actively growing cells are primarily responsible for the disease symptoms [1, 4, 7, 10]. In this pathogen, the formation of highly resistant spores is crucial for persistent infection, leading to high rates of disease recurrence and facilitating transmission [9]. Thus, insight into the mechanisms by which CD25890 impacts gene expression to regulate spore formation sheds light on a regulatory network that is a potential target to disrupt *C. difficile* persistence and virulence.

Using RNAseq we identified an sRNA, *SQ528,* that accumulates to higher levels in a Δ*CD25890* mutant, compared to the WT. Even though RNAseq lacks the ability to distinguish between direct from indirect RNase targets, *SQ528* is likely to be a direct, physiological target for CD25890 because *SQ528* accumulation in the Δ*CD25890* mutant does not result from increased transcription, and also because purified CD25890 cleaves *SQ528* (Fig. 1; see also below). Little is known about the substrate specificity of enzymes of the YicC family. Previous studies have shown that YicC from *E. coli* cleaves at different positions within an RNA hairpin. However, it remains unclear whether enzymes of this family exhibit sequence specificity or simply recognize a secondary structure of a certain length [18]. In the future, identifying additional native sRNA substrates will enable comparative analyses of RNA sequence and structure to identify shared or conserved sequence motifs or secondary structures.

SQ528 shows complementary base pairing with the mRNA encoding a transcription factor identified before as MarR from *C. difficile* (*CD08350*) [54]. Initially characterized in *E. coli,* MarR family members constitute a versatile group of transcriptional regulators responsive to environmental cues and governing the expression of genes involved in various cellular processes [54–58]. These processes include, for example, metabolic pathways, stress responses, virulence mechanisms, and the detoxification of phenolic compounds, antibiotics, and household detergents. [59–62]. MarR of *C. difficile*

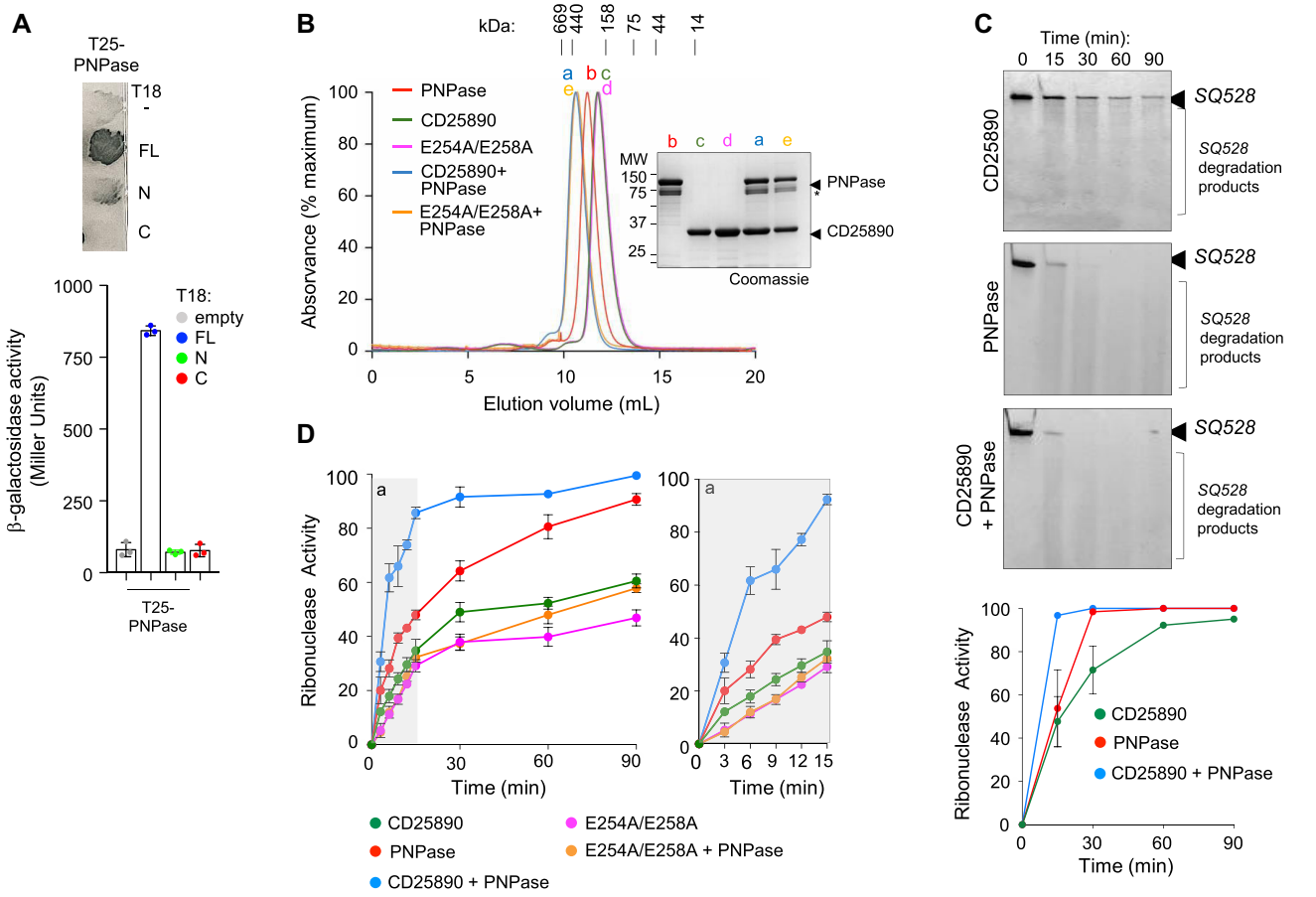


Figure 8. CD25890 and PNPase form an active complex. (**A**) BACTH analysis of protein–protein interactions. The top panel shows colony growth on solid plates for the different combinations of CD25890^{FL}, CD25890^{N}, and CD25890^{C} proteins with PNPase, with blue colonies indicating interaction between the different proteins or domains. Cultures of transformed BTH101 expressing the indicated fusion proteins were grown in liquid medium until mid log and β-galactosidase activity measured (bottom panel). Results are presented as means ± standard deviation from experiments carried out in biological triplicate. (**B**) SEC of affinity purified His$_6$-PNPase, His$_6$-CD25890^{FL}, His$_6$-CD25890$^{E254A/E258A}$, His6-PNPase + His$_6$-CD25890^{FL}, and PNPase + His$_6$-CD25890$^{E254A/E258A}$ on a S200 10/300 GL column. On the top is indicated the position of the elution peaks of MW standards (thyroglobulin, 669 kDa; ferritin, 440 kDa; aldolase, 158 kDa; conalbumin, 75 kDa; ovalbumin, 44 kDa; ribonuclease A, 14 kDa). In the insert is shown the SDS–PAGE analysis of the peaks indicated in the graph. The asterisk marks a degradation product of PNPase, as confirmed by mass spectrometry. (**C**) Time dependent cleavage of *SQ528* sRNA by CD25890^{FL} (WT), PNPase and the complex CD25890/PNPase. Resulting RNA fragments were separated on a 8 M urea–6% polyacrylamide gel followed by staining with ethidium bromide for 30 min and imaging using an Gel Doc system. Quantification of these assays (means ± standard deviation, of at least three replicates) are shown in the graph below. (**D**) Quantification of the time dependent cleavage of *SQ528* sRNA by CD25890^{FL} (WT), CD25890$^{E254A/E258A}$, PNPase and the complexes CD25890/PNPase (0.3 μM of protein; see also Supplementary Fig. S6). Compared to the panel on the left, the panel on the right feature an extended *x*-axis ranging from 0 to 15 min.

shares structural similarity with the MarR proteins from *E. coli* and *Staphylococcus aureus*, indicating that it might also function as a DNA-binding protein and participate in related processes [54].

The genes encoding MarR proteins are often divergent relative to the genes or operons they regulate [55, 60, 61]. In *C. difficile*, MarR is situated downstream of the operon that codes for enzymes involved in the oxidative half of the Krebs cycle, including citrate synthase, aconitase hydratase, and isocitrate dehydrogenase (Fig. 1A). In *Streptomyces coelicolor*, a member of the MarR family known as TamR has been demonstrated to regulate, among other genes, those encoding aconitase and isocitrate dehydrogenase [60]. TamR may respond to metabolites such as citrate to modulate metabolic flux through the Krebs cycle [60, 61]. Given our finding that the increased sporulation phenotype of a Δ*CD25890* mutant is contingent upon the culture medium, we speculate that CD25890 may be part of a regulatory network that integrates metabolic signals with the Krebs cycle to modulate the rate of sporulation initiation in *C. difficile*. In any event, the overexpression of *SQ528*, while enhancing the sporulation efficiency of the *CD25890* mutant, appears to predominantly increase the number of sporulating cells at a late stage in development following the completion of prespore engulfment by the mother cell, including cells with mature phase-bright spores (Fig. 2). Since the primary effect of CD25890 is on the number of cells initiating sporulation [14], this suggests the potential involvement of additional CD25890 substrates in sporulation initiation. In line with this inference, several sRNAs are expressed in *C. difficile* [19, 21], and the *SpoX* and *SpoY* sRNAs have previously been shown to modulate *spo0A* translation, which is essential for sporulation initiation [19]. However, these sRNAs were not identified as potential CD25890 substrates by RNA-seq.

Using a combination of SEC, bacterial two hybrid assays, AlphaFold-Multimer modelling and direct visualization by

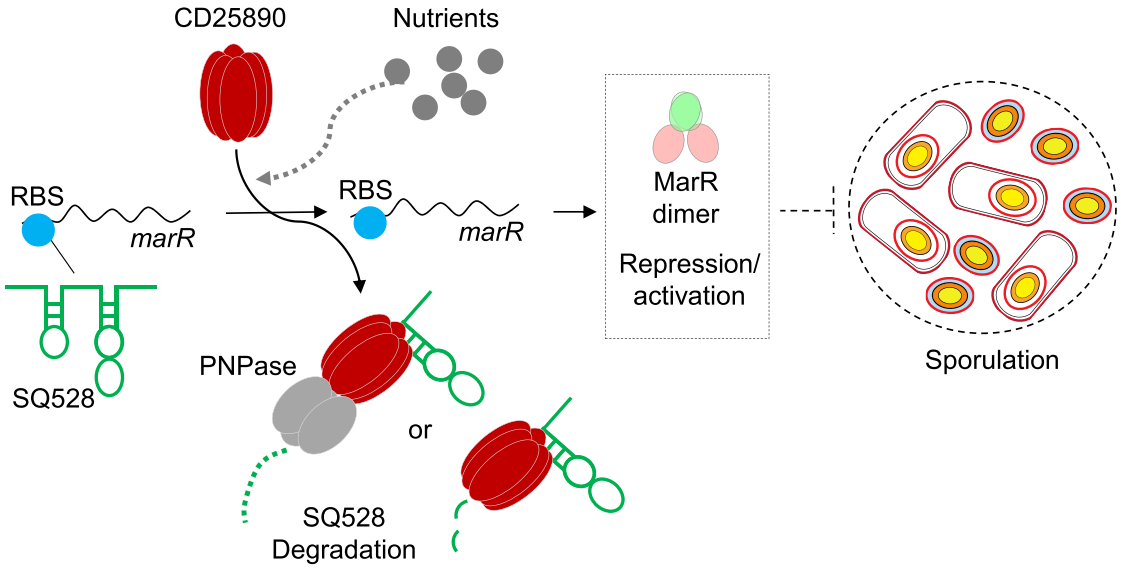


Figure 9. Model for the effect of the endoribonuclease CD25890 on sporulation in *C. difficile*. CD25890 alone or in complex with PNPase plays a role in the degradation of the sRNA *SQ528* in respond to different nutrient sources. Since *SQ528* is antisense to 5′ end of *CD08350*, coding for a transcription factor of the MarR family, CD25890 absence may lead to a decrease in the levels of this transcriptional regulator, which may work directly or indirectly as a repressor of *spo0A* transcription. In this way, the absence of CD25890 would lead to an increase in the number of cells that enter sporulation. Solid line, direct effect; dotted line, indirect effect.

electron microscopy we show that, like YicC and YloC [17, 18, 53], CD25890 is a hexamer (Figs 3 and 7; but see also below). Importantly, CD25890 is also detected as a hexamer in extracts prepared from *C. difficile* cells (Fig. 3). Together, our structural and functional data strongly suggests that the hexamer is the active form of the protein. CD25890 consists of an N-terminal domain involved in multimerization and a C-terminal domain that harbours the putative catalytic residues (Figs 4 and 5). The hexamer exhibits metal-dependent RNase activity against *SQ528*, which is completely abolished upon dual substitution of glutamates 254 and 258, in the C-terminal domain, with alanine (E254A/E258A; Fig. 5). Glutamate residues E216, E217, E252, and E281 of YicC, conserved among orthologues of the protein, are proposed to participate in catalysis via metal ion coordination (Supplementary Fig. S3). Since the E254A and E258A substitutions in CD25890 eliminate RNAase activity, we propose that the two replaced residues are also involved in metal coordination [17, 18, 53]. Strikingly, however, while E258 is homologous to YicC E252, CD25890 E254 is homologous to YicC E248, not previously implicated in catalysis (Supplementary Fig. S3). We suggest that residue E248, conserved among orthologs (Supplementary Fig. S3) is also important for catalysis in YicC. Conversely, the residues homologous to YicC E216, E217, and, E281, are likely to be required for the activity of CD25890 (E222, E223, and E287; Supplementary Fig. S3). Interestingly, even though CD25890$^{E254A/E258A}$ displays reduced catalytic activity, it exhibits enhanced binding to *SQ528* compared to the WT protein (Fig. 6). This effect may result from a decrease in the protein's overall negative charge. In any event, the reduced activity does not seem to result from deficient sRNA binding, but rather from a direct role of E254 and E258 in catalysis. In YicC, a cluster of arginine residues (R30, R251, and R280 [18]; Supplementary Fig. S3) has been shown to mediate RNA binding through electrostatic interactions. In the case of CD25890, the N-terminal domain is sufficient for binding to *SQ528*. However, additional positively charged residues in the C-terminal domain may contribute to nucleic acid binding in the context of the full-length protein.

The cryo-EM structure of YicC bound to a 26-mer RNA revealed a predominant closed complex configuration, resembling a barrel [18, 53]. This closed RNA complex exhibits pseudo-dimeric characteristics, with one trimer recessed and the other trimer extending outward. This inherent asymmetry allows the hexameric complex to present distinct sets of residues on each side of the RNA, with the RNA centrally bound within the barrel. In the absence of the substrate, the complex undergoes a conformational transition, opening into two trimers, forming a tight hexamer primarily at the N-terminal region of the protein. Based on this structural insight, the authors postulated that glutamate residues near the 5′ proximal cleavage site play an essential role in catalysis, likely through manganese coordination, reminiscent of the double metal-ion mechanism of RNase H [18].

Our AlphaFold-Multimer modelling consistently predicted a CD25890 hexamer with six-fold cylindrical symmetry (Fig. 7). This differs from structures of apo and RNA-bound YicC, where symmetry is broken outside of the hexamerization domain [18]. Further structural investigations will be necessary to fully elucidate the structure and dynamics of the CD25890 complex in the presence and absence of RNA, metals, and PNPase. Another RNase that binds RNA within a tunnel is PNPase, which forms a trimer in its active conformation. Importantly, YicC and YloC have been shown to interact with PNPase [15, 17]. Given that YicC and YloC are endoribonucleases, the association with the 3′ to 5′exonuclease PNPase results in a complex able to completely degrade an RNA substrate [15, 17]. We found that CD25890 also interacts with the *C. difficile* homologue of PNPase, as we detected complex formation by SEC and also using a bacterial two hybrid assay BACHT (Fig. 8). We observed interaction between PNPase and full-length CD25890, but not with isolated N- and C-terminal CD25890 domains. We hypothesize that the PNPase weakly interacts with the C-terminal domain of

CD25890, and that CD25890 hexamerization, mediated by its N-terminal domain, is required to stabilize the CD25890-PNPase complex. The interaction between the two enzymes result in the formation of a degradosome machine capable of fully degrading *SQ528* (Fig. 8). Since PNPase was previously identified as an essential gene in *C. difficile* [63], unlike *CD25890*, this suggests that PNPase has additional functions beyond those involving CD25890. The mechanism by which CD25890 influences sporulation under specific nutrient conditions, and the role of the MarR transcriptional regulator in this process, is currently under study (Fig. 9).

Acknowledgements

We thank Rute Pereira for a critical reading of the manuscript.

Author contributions: Diogo Martins (Conceptualization [equal], Formal analysis [equal], Writing—original draft [equal]), Bruno Salgueiro (Conceptualization [equal], Formal analysis [equal]), Daniel Sobral (Data curation [equal], Formal analysis [equal]), Marcos Gragera (Data curation [equal], Formal analysis [equal]), Zach Hensel (Conceptualization [equal], Investigation [equal], Writing—review & editing [equal]), Adriano O. Henriques (Conceptualization [equal], Formal analysis [equal], Funding acquisition [equal], Supervision [equal], Writing—review & editing [equal]), Célia V. Romão (Conceptualization [equal], Funding acquisition [equal], Supervision [equal], Writing—review & editing [equal]) and Mónica Serrano (Conceptualization [equal], Formal analysis [equal], Funding acquisition [equal], Supervision [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This project was supported by award PTDC/BIA-MIC/29293/2017 to M.S. and has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 857203 to C.V.R. This study was financially supported by the Portuguese Fundação para a Ciência e Tecnologia (FCT), Projects MOSTMICRO-ITQB with references UIDB/04612/2020, and LS4FUTURE Associated Laboratory (LA/P/0087/2020).

D.M. is the recipient of a PhD fellowship (PD/BD/143148/2019) within the scope of the PhD program INTERFACE funded by FCT. B.A.S. is the recipient of the FCT PhD4COVID grant SFRH/BD/08066/2020. This work benefited from access to the Instruct Image Processing Centre, an Instruct-ERIC centre. Financial support was provided by "Instruct-ERIC (PID 19564 and 20782)". The microscopy data were acquired at BIC, ITQB-NOVA, Oeiras, Portugal with equipment funded by FCT, project PPBI-POCI-01-0145-FEDER-022122. Mass spectrometry data were generated by the Mass Spectrometry Unit (UniMS), ITQB/iBET, Oeiras, Portugal.

Data availability

Raw sequencing data have been submitted to SRA with the accession number PRJNA1183315.

References

1. Serrano M, Martins D, Henriques AO. *Clostridioides difficile* sporulation. In: Mastrantonio P, Rupnik M (eds.), *Updates on Clostridioides difficile in Europe, Advances in Experimental Medicine and Biology*. Vol. **1435**. Cham: Springer International Publishing, 2024, 273–314.
2. Grossman AD. Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet* 1995;**29**:477–508. https://doi.org/10.1146/annurev.ge.29.120195.002401
3. Janoir C, Denève C, Bouttier S *et al.* Adaptive strategies and pathogenesis of *Clostridium difficile* from *in vivo* transcriptomics. *Infect Immun* 2013;**81**:3757–69. https://doi.org/10.1128/IAI.00515-13
4. Pereira FC, Saujet L, Tomé AR *et al.* The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet* 2013;**9**:e1003782. https://doi.org/10.1371/journal.pgen.1003782
5. Abt MC, McKenney PT, Pamer EG. *Clostridium difficile* colitis: pathogenesis and host defence. *Nat Rev Micro* 2016;**14**:609–20. https://doi.org/10.1038/nrmicro.2016.108
6. Kordus SL, Thomas AK, Lacy DB. *Clostridioides difficile* toxins: mechanisms of action and antitoxin therapeutics. *Nat Rev Micro* 2022;**20**:285–98. https://doi.org/10.1038/s41579-021-00660-2
7. Kuehne SA, Cartman ST, Heap JT *et al.* The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* 2010;**467**:711–3. https://doi.org/10.1038/nature09397
8. Smits WK, Lyras D, Lacy DB *et al.* *Clostridium difficile* infection. *Nat Rev Dis Primers* 2016;**2**:16020. https://doi.org/10.1038/nrdp.2016.20
9. Deakin LJ, Clare S, Fagan RP *et al.* The *Clostridium difficile* spo0A gene is a persistence and transmission factor. *Infect Immun* 2012;**80**:2704–11. https://doi.org/10.1128/IAI.00147-12
10. Cassona CP, Ramalhete S, Amara K *et al.* Spores of *Clostridioides difficile* are toxin delivery vehicles. *Commun Biol* 2024;**7**:839. https://doi.org/10.1038/s42003-024-06521-x
11. Smits W, Browne HP, Choudhary JS *et al.* Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism. *BMC Genomics* 2014;**15**:160.
12. Fimlaid KA, Bond JP, Schutz KC *et al.* Global analysis of the sporulation pathway of *Clostridium difficile*. *PLoS Genet* 2013;**9**:e1003660. https://doi.org/10.1371/journal.pgen.1003660
13. Underwood S, Guan S, Vijayasubhash V *et al.* Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J Bacteriol* 2009;**191**:7296–305. https://doi.org/10.1128/JB.00882-09
14. Martins D, DiCandia MA, Mendes AL *et al.* CD25890, a conserved protein that modulates sporulation initiation in *Clostridioides difficile*. *Sci Rep* 2021;**11**:7887. https://doi.org/10.1038/s41598-021-86878-9
15. Chen J, To L, De Mets F *et al.* A fluorescence-based genetic screen reveals diverse mechanisms silencing small RNA signaling in *E. coli*. *Proc Natl Acad Sci USA* 2021;**118**:e2106964118. https://doi.org/10.1073/pnas.2106964118
16. Huang L, Tam KS, Xie W. Structural and biochemical studies of the novel hexameric endoribonuclease YicC. *ACS Chem Biol* 2023;**18**:1738–47. https://doi.org/10.1021/acschembio.3c00091
17. Ingle S, Chhabra S, Chen J *et al.* Discovery and initial characterization of YloC, a novel endoribonuclease in *Bacillus subtilis*. *RNA* 2022;**28**:227–38. https://doi.org/10.1261/rna.078962.121
18. Wu R, Ingle S, Barnes SA *et al.* Structural insights into RNA cleavage by a novel family of bacterial RNases. *Nucleic Acids Res* 2024;**52**:10705–16. https://doi.org/10.1093/nar/gkae717

19. Hussain HA, Roberts AP, Mullany P. Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Δerm) and demonstration that the conjugative transposon Tn916ΔE enters the genome of this strain at multiple sites. *J Med Microbiol* 2005;**54**:137–41. https://doi.org/10.1099/jmm.0.45790-0
20. Karimova G, Pidoux J, Ullmann A *et al.* A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* 1998;**95**:5752–6. https://doi.org/10.1073/pnas.95.10.5752
21. Wilson KH, Kennedy MJ, Fekety FR. Use of sodium taurocholate to enhance spore recovery on a medium selective for *Clostridium difficile*. *J Clin Microbiol* 1982;**15**:443–6. https://doi.org/10.1128/jcm.15.3.443-446.1982
22. Putnam EE, Nock AM, Lawley TD *et al.* SpoIVA and SipL are *Clostridium difficile* spore morphogenetic proteins. *J Bacteriol* 2013;**195**:1214–25. https://doi.org/10.1128/JB.02181-12
23. Isidro J, Santos A, Nunes A *et al.* Imipenem resistance in *Clostridium difficile* ribotype 017, Portugal. *Emerg Infect Dis* 2018;**24**:741–5. https://doi.org/10.3201/eid2404.170095
24. Fagan RP, Fairweather NF. *Clostridium difficile* has two parallel and essential sec secretion systems. *J Biol Chem* 2011;**286**:27483–93. https://doi.org/10.1074/jbc.M111.263889
25. Plácido D, Fernandes CG, Isidro A *et al.* Auto-induction and purification of a *Bacillus subtilis* transglutaminase (Tgl) and its preliminary crystallographic characterization. *Protein Expr Purif* 2008;**59**:1–8. https://doi.org/10.1016/j.pep.2007.12.004
26. De L, Rosa-Trevín JM, Quintana A *et al.* Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol* 2016;**195**:93–9.
27. Scheres SHW. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 2012;**180**:519–30. https://doi.org/10.1016/j.jsb.2012.09.006
28. Zhang K. Gctf: real-time CTF determination and correction. *J Struct Biol* 2016;**193**:1–12. https://doi.org/10.1016/j.jsb.2015.11.003
29. Wagner T, Merino F, Stabrin M *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun Biol* 2019;**2**:218. https://doi.org/10.1038/s42003-019-0437-z
30. Punjani A, Rubinstein JL, Fleet DJ *et al.* cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 2017;**14**:290–6. https://doi.org/10.1038/nmeth.4169
31. Punjani A, Zhang H, Fleet DJ. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* 2020;**17**:1214–21. https://doi.org/10.1038/s41592-020-00990-8
32. Mirdita M, Schütze K, Moriwaki Y *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;**19**:679–82. https://doi.org/10.1038/s41592-022-01488-1
33. Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-Multimer. bioRxiv, https://doi.org/10.1101/2021.10.04.463034. 10 March 2022, preprint: not peer reviewed.
34. Bryant P, Pozzati G, Elofsson A. Improved prediction of protein–protein interactions using AlphaFold2. *Nat Commun* 2022;**13**:1265. https://doi.org/10.1038/s41467-022-28865-w
35. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* 2008;**3**:1101–8. https://doi.org/10.1038/nprot.2008.73
36. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;**9**:671–5. https://doi.org/10.1038/nmeth.2089
37. Pogliano J, Osborne N, Sharp MD *et al.* A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during *Bacillus subtilis* sporulation. *Mol Microbiol* 1999;**31**:1149–59. https://doi.org/10.1046/j.1365-2958.1999.01255.x
38. Stringer C, Wang T, Michaelos M *et al.* Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* 2021;**18**:100–6. https://doi.org/10.1038/s41592-020-01018-x
39. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**9**:676–82. https://doi.org/10.1038/nmeth.2019
40. Soutourina OA, Monot M, Boudry P *et al.* Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. *PLoS Genet* 2013;**9**:e1003493. https://doi.org/10.1371/journal.pgen.1003493
41. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;**11**:R25. https://doi.org/10.1186/gb-2010-11-3-r25
42. Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* 2007;**23**:2881–7. https://doi.org/10.1093/bioinformatics/btm453
43. Fuchs M, Lamm-Schmidt V, Lenče T *et al.* A network of small RNAs regulates sporulation initiation in regulates sporulation initiation in *Clostridioides difficile*. *EMBO J* 2023;**42**:e112858. https://doi.org/10.15252/embj.2022112858
44. Soutourina O, Dubois T, Monot M *et al.* Genome-wide transcription start site mapping and promoter assignments to a sigma factor in the human enteropathogen *Clostridioides difficile*. *Front Microbiol* 2020;**11**:1939. https://doi.org/10.3389/fmicb.2020.01939
45. Saujet L, Monot M, Dupuy B *et al.* The key sigma factor of transition phase, SigH, controls sporulation, metabolism, and virulence factor expression in *Clostridium difficile*. *J Bacteriol* 2011;**193**:3186–96. https://doi.org/10.1128/JB.00272-11
46. Rigden DJ. *Ab initio* modeling led annotation suggests nucleic acid binding function for many DUFs. *OMICS* 2011;**15**:431–8. https://doi.org/10.1089/omi.2010.0122
47. Sinturel F, Pellegrini O, Xiang S *et al.* Real-time fluorescence detection of exoribonucleases. *RNA* 2009;**15**:2057–62. https://doi.org/10.1261/rna.1670909
48. Yokoi H, Onishi H. Effects of several metal ions, especially zinc ions, on RNA degradation by halophilic nuclease H in solution or adsorbed on flocculated cells of halophilic *Micrococcus*. *Agric Biol Chem* 1989;**53**:1817–22.
49. Butzow JJ, Eichhorn GL. Different susceptibility of DNA and RNA to cleavage by metal ions. *Nature* 1975;**254**:358–9. https://doi.org/10.1038/254358a0
50. Nurmohamed S, Vaidialingam B, Callaghan AJ *et al.* Crystal structure of *Escherichia coli* polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradosome assembly. *J Mol Biol* 2009;**389**:17–33. https://doi.org/10.1016/j.jmb.2009.03.051
51. Shi Z, Yang W-Z, Lin-Chao S *et al.* Crystal structure of *Escherichia coli* PNPase: central channel residues are involved in processive RNA degradation. *RNA* 2008;**14**:2361–71. https://doi.org/10.1261/rna.1244308
52. Carpousis AJ. The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* 2007;**61**:71–87. https://doi.org/10.1146/annurev.micro.61.080706.093440
53. Huang L, Tam KS, Xie W. Structural and biochemical studies of the novel hexameric endoribonuclease YicC. *ACS Chem Biol* 2023;**18**:1738–47. https://doi.org/10.1021/acschembio.3c00091
54. Peng JW, Yuan H, Tan XS. Crystal structure of the multiple antibiotic resistance regulator MarR from *Clostridium difficile*. *Acta Crystallogr F Struct Biol Commun* 2017;**73**:363–8. https://doi.org/10.1107/S2053230X1700766X
55. Alekshun MN, Levy SB. Alteration of the repressor activity of MarR, the negative regulator of the *Escherichia coli marRAB* locus, by multiple chemicals *in vitro*. *J Bacteriol* 1999;**181**:4669–72. https://doi.org/10.1128/JB.181.15.4669-4672.1999

56. Beggs GA, Brennan RG, Arshad M. MarR family proteins are important regulators of clinically relevant antibiotic resistance. *Protein Sci* 2020;**29**:647–53. https://doi.org/10.1002/pro.3769
57. Deochand DK, Grove A. MarR family transcription factors: dynamic variations on a common scaffold. *Crit Rev Biochem Mol Biol* 2017;**52**:595–613. https://doi.org/10.1080/10409238.2017.1344612
58. McClennen G, Angelini LL, Fox G *et al*. Discovery of a novel MarR-type transcriptional regulator that controls cell death in *Bacillus subtilis* biofilms. bioRxiv, https://doi.org/10.1101/2023.11.22.568295, 23 November 2023, preprint: not peer reviewed.
59. Grove A. Regulation of metabolic pathways by MarR family transcription factors. *Comput Struct Biotechnol J* 2017;**15**:366–71. https://doi.org/10.1016/j.csbj.2017.06.001
60. Huang H, Sivapragasam S, Grove A. The regulatory role of *Streptomyces coelicolor* TamR in central metabolism. *Biochem J* 2015;**466**:347–58. https://doi.org/10.1042/BJ20130838
61. Huang H, Grove A. The transcriptional regulator TamR from from S*treptomyces coelicolor* controls a key step in central metabolism during oxidative stress. *Mol Microbiol* 2013;**87**:1151–66. https://doi.org/10.1111/mmi.12156
62. Randazzo P, Aubert-Frambourg A, Guillot A *et al*. The MarR-like protein PchR (YvmB) regulates expression of genes involved in pulcherriminic acid biosynthesis and in the initiation of sporulation in *Bacillus subtilis*. *BMC Microbiol* 2016;**16**:190. https://doi.org/10.1186/s12866-016-0807-3
63. Dembek M, Barquist L, Boinett CJ *et al*. High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. *mBio* 2015;**6**:e02383–14. https://doi.org/10.1128/mBio.02383-14

Received: November 12, 2024. **Revised:** May 29, 2025. **Editorial Decision:** June 16, 2025. **Accepted:** June 19, 2025