7164–7178 *Nucleic Acids Research, 2021, Vol. 49, No. 12* *Published online 17 June 2021*
https://doi.org/10.1093/nar/gkab513

Alternative dimerization is required for activity and inhibition of the HEPN ribonuclease RnlA

Gabriela Garcia-Rodriguez [1,2], Daniel Charlier[3], Dorien Wilmaerts [4,5], Jan Michiels [4,5] and Remy Loris [1,2,*]

[1]Structural Biology Brussels, Department of Biotechnology, Vrije Universiteit Brussel, B-1050 Brussel, Belgium, [2]Molecular Recognition Unit, Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, B-1050 Brussel, Belgium, [3]Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium, [4]Center of Microbial and Plant Genetics, KU Leuven, Kasteelpark Arenberg 20, Box 2460, 3001 Leuven, Belgium and [5]Center for Microbiology, Vlaams Instituut voor Biotechnologie, Leuven, Belgium

Received January 28, 2021; Revised May 09, 2021; Editorial Decision May 30, 2021; Accepted June 03, 2021

ABSTRACT

The *rnlAB* toxin-antitoxin operon from *Escherichia coli* functions as an anti-phage defense system. RnlA was identified as a member of the HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding domain) superfamily of ribonucleases. The activity of the toxin RnlA requires tight regulation by the antitoxin RnlB, the mechanism of which remains unknown. Here we show that RnlA exists in an equilibrium between two different homodimer states: an inactive resting state and an active canonical HEPN dimer. Mutants interfering with the transition between states show that canonical HEPN dimerization via the highly conserved $RX_{4-6}H$ motif is required for activity. The antitoxin RnlB binds the canonical HEPN dimer conformation, inhibiting RnlA by blocking access to its active site. Single-alanine substitutions mutants of the highly conserved R255, E258, R318 and H323 show that these residues are involved in catalysis and substrate binding and locate the catalytic site near the dimer interface of the canonical HEPN dimer rather than in a groove located between the HEPN domain and the preceding TBP-like domain. Overall, these findings elucidate the structural basis of the activity and inhibition of RnlA and highlight the crucial role of conformational heterogeneity in protein function.

INTRODUCTION

When a *dmd*$^-$ mutant of phage T4 infects *Escherichia coli* cells, gene expression is halted at late stages by the rapid degradation of the phage mRNAs, resulting in the inability of this mutant to proliferate (1). RnlA was first identified as an *E. coli* endoribonuclease responsible for late gene silencing of the *dmd*$^-$ mutant T4 phage after infection (2) and later on for post-transcriptionally regulating endogenous *E. coli* gene expression, in particular the synthesis of adenylate cyclase (3). RnlA was subsequently discovered to be the toxin part of the *rnlAB* toxin-antitoxin (TA) module, which functions as an anti-phage system in *E. coli* cells, particularly upon infection of the *dmd*$^-$ mutant of the T4 phage (4). The endoribonuclease activity of RnlA is controlled by RnlB. RnlA is also directly inhibited by the phage antitoxin encoded by the *dmd* gene (5). The RnlA endoribonuclease constitutes one of the few examples of bacterial toxins with phage-encoded direct inhibitors, thus evidencing the co-evolution of bacteria and their phages.

Nucleases are widely represented amongst toxins of TA systems in bacteria. Endoribonucleases in particular, are most abundant among toxins of type II TA systems (a group of TA system in which both toxin and antitoxin are proteins) (6,7). One of the proposed causes for the extensive prevalence of endoribonucleases as toxins is attributed to the conservation of their RNA targets in all living systems, rendering them effective in a wide range of species in bacteria and archaea, and even in eukaryotes. Endoribonucleases from TA systems have been successfully employed as tools for genetic manipulation of yeasts (8), in the cell-ablation-based bio-containment strategy in *Arabidopsis thaliana* (9), and in the development of antiviral and anticancer gene therapy strategies (10,11).

RnlA and its plasmid-encoded homolog LsoA have a three-domain structure and a recent bioinformatics analysis suggests that they belong to the HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding domain) superfamily, where they define a distinct branch of ribonu-

*To whom correspondence should be addressed. Tel: +32 2 6291989; Email: remy.loris@vub.be
Present address: Gabriela Garcia-Rodriguez, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland.

© The Author(s) 2021. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

cleases. Their C-terminal domain, previously termed Dmd-binding domain (DBD), displays the conserved HEPN topology (12). The HEPN domains of both RnlA and LsoA contain two conserved arginines and a glutamate (R255, E258 and R318 in RnlA) thought to be active site residues (13,14). The other two domains have identical folds and were termed N-terminal (NTD) and N-repeated (NRD) domain, respectively (13). The function of the N-terminal domain is unknown, but the N-repeated domain is believed to contribute to substrate binding as the alternative phage-encoded antitoxin Dmd binds in a groove at the interface between the NRD and the HEPN domains close to the location of the above mentioned conserved Arg and Glu residues (14).

The HEPN domain was first described as the 110 residues-long domain at the C-terminus of the human protein Sacsin. HEPN domains are widely distributed amongst many bacterial and archaeal species, and also amongst animals (15). The HEPN domain consists of five α-helices, with three of them arranged in an up-and-down helical bundle and two shorter helices on one side (15). According to their domain architectures, gene-neighborhoods, phyletic patterns and operon organization, seven distinctive HEPN domain families can be discerned. Many function as metal-independent RNases that take part in cellular RNA maturation as well as in biological conflicts, such as virus-host interactions and as part of selfish genetic elements (12). All these HEPN proteins contain an $RX_{4-6}H$ motif as their most strongly conserved feature (where X represents any residue). The Arg and His in this motif were shown to be essential for the endoribonuclease activity in a number of widely divergent HEPN proteins (16,17).

When HEPN domains were first identified in bacteria and archaea, they were found to frequently co-localize with genes encoding minimal nucleotidyltransferases (MNTs) (15), which suggested a concerted action of both gene products. Interestingly, the $RX_{4-6}H$ motif is usually lost in HEPN-MNT fused aminoglycoside NTases (18,19) (e.g., PDB entries 1KNY and 4CS6), where the HEPN domain is believed to have preserved a nucleotide binding function upon the loss of its RNase activity (20). However, non-covalent HEPN-MNT pairs have since then been shown to form typical type II TA systems, for example, in *Shewanella oneidensis* (21,22), where the HEPN toxin functions as an RNase that is directly inhibited by its cognate MNT antitoxin. HEPN-MNT TA modules are since then identified as the most widely distributed class of mobile two-gene modules in prokaryotes, but their functions and mechanisms remain elusive (23).

Here, we analyze the functional mechanism of the *E. coli* toxin RnlA, which defines a family of ribonucleases within the HEPN superfamily. We show that, in contrast to what was suggested in earlier work (13), the dimer observed in the crystal structure of the free toxin corresponds to a non-active precatalytic conformation. Structural and mutagenesis studies show that RNase activity requires the formation of a canonical HEPN dimer and that the active site is located near this HEPN dimer interface and not at the interface between the NRD and the HEPN domain previously observed in the crystal structure of the free RnlA dimer (13,14).

MATERIALS AND METHODS

Cloning, expression and purification of His-RnlA and RnlA-RnlB-His

An expression clone for wild-type His-RnlA (*his-rnlAB*_pET28a) and the RnlA:RnlB-His complex (*rnlAB-his*_pET21b) in BL21(DE3)) was available from previous work (24). Site-directed mutagenesis of RnlA was carried out by PCR extension or full-plasmid amplification with partially overlapping primers (25) containing a single amino acid mutation, from the His-RnlA expressing plasmid *his-rnlAB*_pET28a. Previously the *rnlA* coding region was amplified by PCR to exchange the initial thrombin cleavage site between the N-terminus His-tag and the start of the *rnlA* gene for the TEV cleavage sequence. The resulting genes encoding the N-terminal His-tagged R255A, E258A, R318A, H323A, V206R and D245R RnlA mutants were cloned into pET28a plasmids under the control of the inducible T7 promoter and transformed into BL21(DE3) competent cells for further protein expression and purification. The RnlA truncation was generated by NEB Assembly from the *his-rnlAB_pET28a* plasmid. The resulting genes encoding the N-terminal His-tagged RnlAΔ1–91 in presence and absence of RnlB were cloned into a *pET21b* vector. Each construct was verified by DNA sequencing.

RnlA wild-type and RnlA:RnlB-His complex were produced and purified as described by Garcia-Rodriguez *et al.* (24). The endogenous *rnlAB* operon is expressed under the control of the inducible T7 promoter in BL21(DE3) cells. Placement of the affinity tag on RnlA leads to loss of RnlB antitoxin during Ni-NTA purification followed by gel filtration. A homogeneous preparation of RnlA:RnlB containing stoichiometric amounts of both proteins could only be obtained using a His tag placed C-terminal to RnlB.

For RnlA mutants and the RnlAΔ1–91 (TBP2-HEPN) truncated version, BL21(DE3) cells containing the *pET28a rnlA_mutant* or the *RnlAΔ1–91-rnlB pET21b* plasmids were grown overnight in LB medium supplemented with 50 μg ml^{-1} kanamycin or 100 μg ml^{-1} ampicillin, respectively, at 310 K. A 100-fold dilution was used to inoculate 1 l cultures of the same medium that were incubated at 310 K until OD_{600} reached 0.6, then the temperature was decreased to 293 K and expression of the desired proteins was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were harvested after 20 h of expression by centrifugation at 277 K and the pellets were resuspended in lysis buffer (20 mM Tris–HCl, pH 8, 500 mM NaCl, 2 mM 2-mercaptoethanol) supplemented with a protease inhibitor cocktail (cOmplete, Sigma-Aldrich/Merck). Cell extracts were injected into a Ni Sepharose column for affinity chromatography purification followed by size exclusion chromatography of the RnlA-mutant–containing fractions as a final step. The progress of the purification was verified by SDS-PAGE (26) and anti-Histidine tag Western blot (27).

Cell-free expression of RnlB antitoxin

Successful expression and purification of RnlB in *E. coli* failed and therefore synthesis of C-terminally His-tagged

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

RnlB (RnlB-His) in a cell-free expression system was carried out using WEPRO®7240H Expression Kit (CellFree Sciences Co., Ltd). The *rnlB* gene with a C-terminal His-tag sequence was cloned into the pEU-E01 vector under the SP6 promoter. The final *rnlB-his-pEU-E01* plasmid was purified, verified by DNA sequencing and subsequently used for 'large' scale transcription and translation in reaction volumes of 250 μl and 501 μl, respectively, according to the kit manufacturer's instructions. To isolate RnlB-His from the components of the wheat germ extract expression kit, a Ni-NTA purification in batch mode was carried out. 100 μl of pre-washed Ni Sepharose High Performance resin (GE Healthcare) was added to the final translation mixture including WEPRO®7240H. After 15 min incubation at room temperature on a shaking platform, the unbound fraction was separated by centrifugation and the Ni-bound proteins were washed two times with 20 mM Tris-HCl, pH 8, 500 mM NaCl, 1 mM tris(2-carboxyethyl) phosphine. RnlB-His bound to the Ni beads was eluted by adding 500 mM imidazole to the wash buffer. The eluted fraction was then injected onto a Superdex column for final purification. The progress was monitored by SDS-PAGE and anti-Histidine western blot.

Ribonuclease activity assays

RNase activity of RnlA and its mutants was monitored by following the cleavage of bacteriophage MS2 genomic RNA (Roche Diagnostics) via denaturing gel electrophoresis as described (28). RNase-free purifications of all proteins were completed before carrying out the activity assays, by preparing buffers with diethyl pyrocarbonate (DEPC)-pretreated water and/or Ultrapure RNase-free water, using RNase-free laboratory consumables and RNase-cleaned surfaces. Fifty picomoles of RnlA, the reconstituted RnlA:RnlB complex at a 1:2 molar ratio, and 50 pmol of the four alanine substitution RnlA mutants, were all incubated with 0.08 μg μl^{-1} of MS2 RNA in 10 μl final volume, in the presence of RNase Inhibitor (SUPERase In™ RNase Inhibitor, Invitrogen). All reactions were incubated for 1 h at 310 K in PBS. The reactions were stopped by addition of 10 μl of 2x loading dye (95% v/v formamide, 0.02% m/V SDS, 0.02% m/V bromophenol blue, 0.01% m/V xylene cyanol, 1 mM EDTA). Samples were loaded after 10 min incubation at 343 K onto 6% urea gels which had been pre-running at 100 V for 20 min in TBE buffer. The low range ssRNA ladder from New England Biolabs was used as a reference in all gels, which were ran at 150 V for 45 min and subsequently stained with ethidium bromide.

In vivo toxicity assay

Escherichia coli BL21(DE3) cells transformed with *rnlA_pET28*, *R255A_rnlA_pET28*, *E258A_rnlA_pET28a*, *R318A_rnlA_pET28a, H323A_rnlA_pET28a, V206R_rnlA, D245R_rnlA* or *pET28a* (empty vector) were inoculated in 5 ml LB with kanamycin at a final concentration of 40 μg ml^{-1}, at 310 K (200 rpm). After overnight incubation, cultures were diluted 100-fold in a honeycomb plate containing 297 μl LB medium, kanamycin (40 μg ml^{-1}) and IPTG at 1 mM, (time point 0 in Figures 3 and 6). Growth curves were obtained by incubating the plates in a Bioscreen C Device at 310 K under continuous shaking. Measurements with the interface-destabilizing mutants *V206R_rnlA* and *D245R_rnlA* were performed at 303 K to favor the stability of V206R RnlA. OD (595 nm) measurements were taken every 15 min for 12 h.

Primer extension assays

0.08 μg of total MS2 RNA from each ribonuclease activity reaction mix were used for primer extension analysis with primer 5′-ttagtagatgccggagtttgctgcg-3′. The primer was labeled on its 5′ end with (γ ^{32}P)-ATP (Perkin Elmer, 3000 $Ci.mmol^{-1}$) and T4 polynucleotide kinase (Thermofisher). One pmol of labeled primer was mixed with 1 μl of ribonuclease reaction mix, after the reactions were carried out as described in the previous section except for the following modifications. 36 pmol of RnlA were incubated for 10 min with increasing amounts of RnlB, to evaluate the effect of the antitoxin on the RnlA specific cleavage of MS2 RNA. Additionally, other RnlA-MS2 reactions were incubated for 1, 2, 5, 10 and 20 min to monitor the kinetics of the enzymatic reaction. All reactions were stopped by the addition of a 2.5 molar excess of RnlB antitoxin. A range of concentrations for the RnlAΔ1–91 truncated protein, the interface-destabilizing RnlA mutants V206R and D245R, as well as 50 pmol of RnlA single-alanine mutant proteins were also included in the primer extension analysis. These reactions were set up as described in the previous section, except that incubation of proteins and MS2 RNA was for 30 min. 36 and 50 pmol of wild-type RnlA, 100 pmol of RnlB only, and co-incubation of toxin and antitoxin prior to the addition of substrate were also included as references and negative controls. RnlAΔ1–91 and the V206R and D245R RnlA mutants pre-incubated with equimolar RnlB amounts were also included in the analysis.

Primer and template RNA mix were incubated for 5 min at 338 K and allowed to cool down to RT for ∼30 min to favor annealing. Recombinant M-MuLV reverse transcriptase from the ProtoScript® II kit (NEB) was used to synthesize first strand cDNA from the annealed primer-MS2 RNA mix, according to the kit's manufacturer. A set of sequence ladders for the MS2 coat protein was obtained by addition of the four ddNTPs (Sigma) in a final concentration of 1 mM each to the M-MulV reaction mix containing dNTPs at 0.25 mM each. Primer extension reactions were incubated at 315 K for 1 h after which the enzyme was inactivated for 5 min at 353 K and subsequently stored at 253 K. Formamide dye mix was added to the samples and after incubation at 365 K for 3 min they were loaded on a 6% denaturing polyacrylamide gel and ran for ∼2 h at 60 W (≈ 1500 V).

ssDNA binding assays

The 5′-biotin-labeled oligonucleotideTTGCTGCGAT TGCTGAGGGAATCGGGTTTCCATCTTT annealed to the 80-mer oligonucleotide of sequence TTCCGACTGCGAGCTTATTGTTAAGGCAATGCA AGGTCTCCTAAAAGATGGAAACCCGATTCCCT

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

CAGCAATCGCAGCAA with free 3′-OH followed by a 43-nucleotides long ssDNA segment (underlined). This fragment corresponds to the MS2 RNA sequence spanning over RnlA-specific cleavage sites 1 to 4 as identified in our primer extension assays. Biolayer interferometry (BLI) experiments were carried out in an Octet RED96 system with Streptavidin biosensors loaded with the primer-template DNA and blocked with BSA. The affinity of wild-type RnlA, the RnlAΔ1–91 truncate and single-alanine mutant proteins to this ssDNA oligonucleotide was measured by placing the sensors into solutions of several protein concentrations. PBS supplemented with 0.5% BSA and 0.02% Tween20 was used in all steps. Partial fitting of the data during association and dissociation to a 1:1 model (1 substrate molecule for 1 RnlA dimer) was performed with the Octet Data Analysis software (FORTÉBIO), followed by the steady state analysis with the theoretical R_{eq} values based on the curve fits. The analysis was performed in triplicate for each protein and the estimated K_D values averaged and reported with associated errors.

Macromolecular crystallography

Crystallization and data collection for His–RnlA and the RnlA:RnlB–His complex have been described (24). RnlA mutant R255A and the RnlAΔ1–91 truncate in 20 mM Tris–HCl pH 8, 150 mM NaCl, 1 mM TCEP were concentrated to 18 and 24 mg ml^{-1}, respectively. Crystallization conditions were screened using a Mosquito HTS robot from TTP Labtech (http://ttplabtech.com/) using 0.1 μl of protein solution and 0.1 μl of reservoir solution in a sitting drop equilibrated against 100 μl of reservoir solution. Crystals of R255A RnlA grew after one week in 100 mM Tris–HCl, pH 8.5, 8% PEG 8000. Crystals of RnlAΔ1–91 grew after 2–4 days in 0.2 M NaCl; 0.1 M Bis–Tris; pH 5.5; 25% PEG 3350. Data were collected at beamline Proxima 2A of the Soleil Synchrotron (Gif-Sur-Yvette, France) for R255A RnlA mutant and Proxima 1 beamline of the Soleil synchrotron for RnlAΔ1–91. All data were analyzed using XDS (29) and are summarized in Supplementary Table S1. Additionally, merged intensity data for RnlAΔ1–91 were analyzed by the DEBYE and STARANISO software via the web interface (http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi) to perform Bayesian estimation of structure amplitudes, and to apply an anisotropic correction to the data. The estimation of the unit cell content was performed by the MATTHEW COEF program implemented in the CCP4 program suite (30).

The structure of the RnlAΔ1–91 truncation was determined by molecular replacement with Phaser. TBP2 and HEPN domains from PDB ID 4I80 were used as search models. Coordinates were refined in *phenix.refine* (31), using an intensities-based maximum likelihood target including rigid body and TLS refinement with automated group determination as implemented in *phenix.refine*. The structures of wild-type His-RnlA as well as of the R255A mutant protein were determined by molecular replacement with Phaser (32,33) using the deposited co-ordinates of the RnlA structure (PDB entry 4I8O) (13) determined in a different space group. NCS and reference model (from the RnlAΔ1–91 structure) restraints were imposed in the refinement of RnlA WT and R255A structures. Rebuilding and addition of solvent atoms was done using Coot (34).

The structure of the RnlA:RnlB–His complex was determined by molecular replacement using Phaser-MR (32,33). A variety of search models was used ranging from the full RnlA dimer in PDB entry 4I8O to each of the individual domains of RnlA. Only the C-terminal RnlA domain (DBD, resid 203–357) resulted in an acceptable solution. Two copies of this domain were placed in a dimeric arrangement that did not reproduce the free RnlA dimer with LLG and TFZ values of 650 and 25.7, respectively. This partial solution did not provide sufficient phase information to allow building of the remaining parts of the structure. *phenix.mr_rosetta* (35) provided the essential tools for further model-building. The partial model and phases previously obtained from Phaser-MR were refined in Phenix and then used for a first phenix.mr_rosetta run. 9-mer and 3-mer fragments from the PDB, compatible with the sequence of RnlA were generated by the Robetta fragment server (http://robetta.bakerlab.org) and tallied as inputs to *phenix.mr_rosetta*). Several cycles of phenix refinement and manual model building in Coot improved the phases and provided a new partial model, consisting now of 445 residues which was used for another cycle of *phenix.mr_rosetta*, this time also including the Robetta generated fragments for the RnlB sequence. The overall best solution of the Autobuild run of this *phenix.mr_rosetta* job rendered R/R_{free} values of 0.29/0.32 with 570 residues built in. Further iterative cycles of refinement and both manual and automated model building using Coot and Buccaneer (36), respectively, lead to the completion of the RnlA:RnlB structure. All refinement statistics are given in Supplementary Table S1.

Small angle X-ray scattering

SEC-SAXS data were collected at beamlines BM29 (ESRF) and SWING [Soleil Synchrotron (Gif-Sur-Yvette, France)], using a setup with a SEC column (Shodex KW402.5–4F) placed just before the capillary to which the X-rays are directed. This setup helps to remove aggregates from the sample that could interfere with the measurements. Concentration series were collected in batch mode for RnlA wild-type and D245R mutant. Protein samples were prepared as described above for crystallization, concentrated and briefly centrifuged in a table-top centrifuge at 10 000 rpm for 2 mins, before loading on the SEC column or capillary. All samples were measured in 20 mM Tris, pH 8, 150 mM NaCl, 1 mM TCEP supplemented with 5% glycerol for stabilization of proteins in solution, with a constant column flow of 0.2 ml min^{-1}. The final scattering curve was generated for each sample in the SEC-SAXS runs after a range of scattering curves around the peak was normalized and averaged, as indicated in Supplementary Figures S11 and S12. The R_g values were derived from the Guinier approximation at small q values while the I_o parameter was estimated by extrapolation to $q = 0$ using the ATSAS suite (37). For the samples measured in batch mode, 10 successive 1 sec-

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

ond frames were collected for each concentration point in a series. The low angle data collected at lower concentration were extrapolated to infinite dilution and merged with the higher concentration data to yield the final scattering curves. The overall quality of the SAXS data was assessed by Guinier analysis, $P(r)$ vs r and Kratky plots (Supplementary Figures S11 and S12).

Computation of the SAXS profile of the atomic resolution models and its comparison to the experimental data was performed in the FoXS webserver (38) (https://modbase.compbio.ucsf.edu/foxs/), which optimizes the excluded volume (c_1) and hydration layer density (c_2) to improve the fit of a model to the experimental SAXS data.

The MutiFoXS (38) tool of this webserver was used to address conformational heterogeneity in solution by considering multiple states contributing to the observed SAXS profile.

All-atom input models were generated by modeling missing fragments (N-terminal histidine tags and missing loops) with MODELLER (https://salilab.org/modeller/) (39). 10 000 conformations were generated by sampling the space of the φ and ψ main chain dihedral angles of selected flexible residues with a Rapidly exploring Random Trees (RRTs) algorithm. Residues in loops and connecting domains were defined as flexible while keeping the rest of the protein as rigid body. Subsequently, the computation of the SAXS profile for each sampled conformation was performed followed by the enumeration of the best-scoring multi-state models using the multi-state scoring function implemented in MultiFoXS (38). SAXS data collection and processing, as well as modeling parameters are summarized in Supplementary Tables S2–S5.

RESULTS

RnlA is a HEPN protein that adopts a non-canonical dimer in its resting state

In order to characterize the activity *in vitro* and to determine the conformation of isolated RnlA, the purification of the toxin was optimized and, subsequently, crystallization screens were performed. Expression of a bicistronic construct encoding N-terminally His-tagged RnlA (His-RnlA) followed by RnlB allows the purification and crystallization of His-RnlA, since RnlB is largely lost during purification possibly due to proteolysis. The resulting structure determined at 2.9 Å is very similar, although in a different crystal form, to an RnlA structure (PDB ID 4I8O) presented earlier (13) (RMSD = 0.907 Å over 4418 atoms from 668 residues). The RnlA monomer consists of three domains, previously denoted N-terminal domain (NTD), N-terminal repeated domain (NRD) and Dmd-antitoxin binding domain (DBD) (residues 1–92, 93–193 and 194–357, respectively, Figure 1A). A DALI search revealed that the NTD and NRD domains adopt a TATA box-binding protein (TBP)-like domain fold most similar to the N-terminal domain of *Aquifex aeolicus* RNase H3 (40) (PDB ID 3VN5, Z score of 6.2 - Supplementary Figure S1A). Therefore, they will be referred to as TBP1 and TBP2 from hereon. The C-terminal domain adopts a HEPN nuclease fold and will be referred to as the HEPN domain. The closest structural relatives for the HEPN domain, aside from LsoA, comprise the TA toxins SO_3166 and HI0074 from *Shewanella oneidensis* and *Haemophilus influenzae,* respectively (21,41), but also other HEPN proteins including human Sacsin (42) are readily identified in a DALI search.

A sequence alignment of the RnlA family (PF19034) in the *pfam* database (http://pfam.xfam.org) reveals the presence of an $RX_{4\text{-}6}H$ motif (R318 and H323 in RnlA) that was earlier identified among members of the HEPN superfamily and that is considered to be part of the active site of HEPN ribonucleases. This minimal motif is located at the C-terminus of helix α5 of the HEPN domain (α13 in full length RnlA) but this helix differs in length in the different HEPN subfamilies (16–18). In addition, an arginine and glutamate residue (R255 and E258) that were suggested as possible catalytic residues in LsoA (14) are equally conserved in the RnlA/LsoA subfamily, but not in other HEPN subfamilies (Supplementary Figure S1B).

Available high-resolution structures show homodimerization of HEPN domains via the juxtaposition of the conserved $RX_{4\text{-}6}H$ motifs in the dimerization interface (16–18). This homodimerization mode of HEPN domains renders a cleft where the active site has been located in other HEPN proteins (16,17,21). We will refer to this dimerization mode as 'canonical HEPN dimerization' from here onwards. Strikingly, in the structure of isolated RnlA presented here, a non-canonical HEPN dimer is formed via interactions between helixes α8 and α10. Remarkably, this non-canonical HEPN dimer places the conserved $RX_{4\text{-}6}H$ motif in helix α13 on opposite sides of the dimerization interface. We will refer to this alternative conformation as the resting or unbound state as it is not involved in binding of substrate nor antitoxin (see later). Although the dimer interface is relatively small for a protein of this molecular weight (1638 $Å^2$ buried), the PISA analysis of the full-length RnlA structure outputs this dimer conformation as the only stable assembly among the different contacts in the crystal, with a theoretical free Gibson's energy gain of –15 kcal mol^{-1}. Moreover, the presence of this non-canonical dimer in solution is confirmed by SAXS (Figure 1A, C and Supplementary Figure S11), and is also in agreement with previous results (13).

Further evidence for the stability of this non-canonical HEPN dimer stems from a truncation mutant of RnlA (RnlAΔ1–91) that lacks the N-terminal residues 1–91 corresponding to the TBP1 domain. The crystal structure of RnlAΔ1–91 shows the same non-canonical HEPN dimer observed in the crystal structure of isolated RnlA (Figure 1D, E – RMSD of 1.56 Å over 435 common Cα atoms with the full-length dimer) and was confirmed in solution using SAXS. The theoretical SAXS profile for an all-atom model of the RnlAΔ1–91 crystallographic dimer agrees with the experimental SAXS profile with $\chi^2 = 3.1$ (Figure 1F and Supplementary Figure S11).

Taken together, these results indicate that isolated RnlA exhibits a non-canonical HEPN dimer conformation both in the crystal and in solution, which might not correspond to the conformation adopted by RnlA for catalysis. Next, we set out to explore the activity and specificity of isolated RnlA *in vitro*.

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

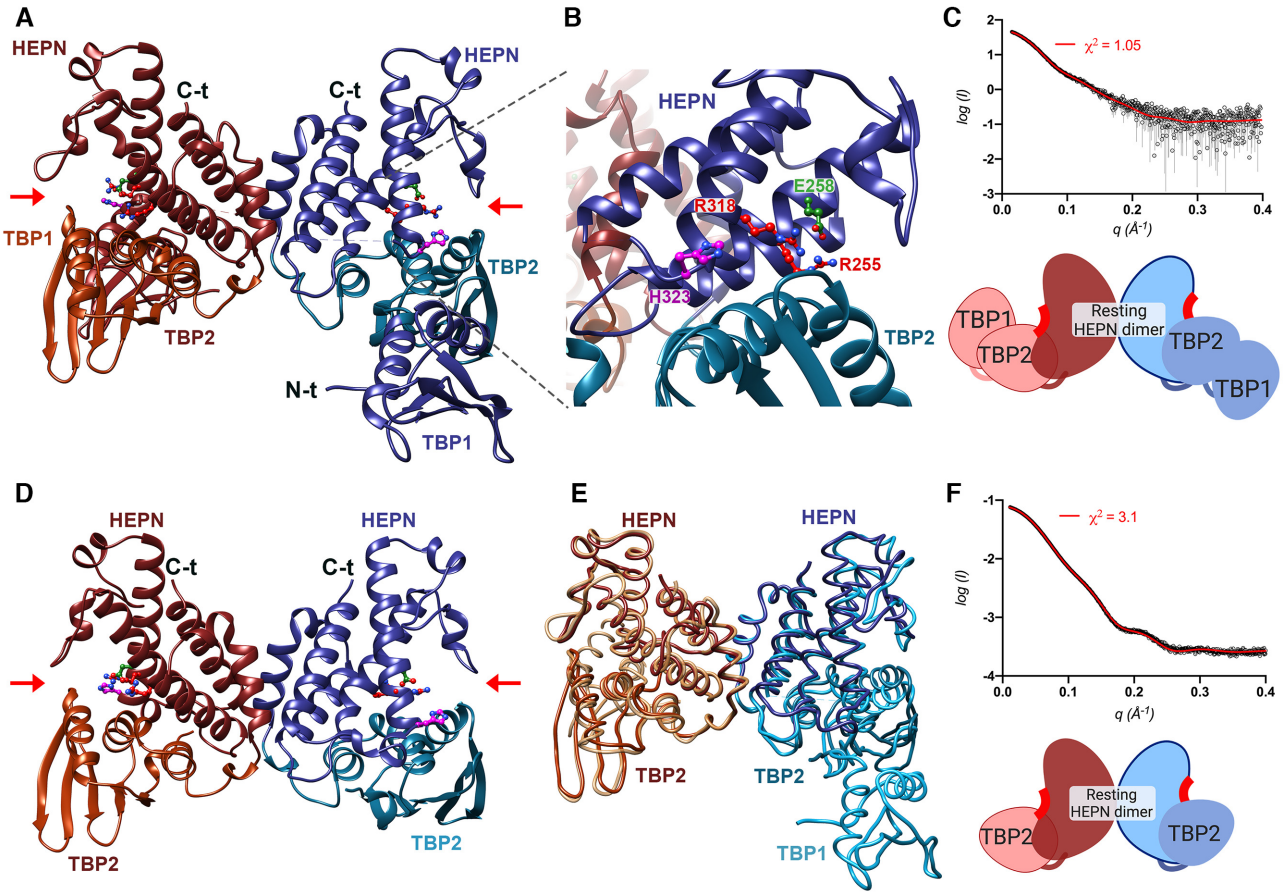

Figure 1. Unbound RnlA exists as a resting dimer both in the crystal and in solution. (**A**) Ribbon representation of the RnlA dimer in its free state. The three domains are labeled in both chains (TBP1, TBP2, HEPN). Chains are colored in red and blue; TBP2 domains are colored in a different shade of red and blue, respectively, to mark a clear distinction between domains. Red arrows indicate the position of the conserved catalytic residues in the TBP2-HEPN interface of each chain. (**B**) Ball and stick representation of the proposed catalytic residues R255, E258, R318 and H323 in the TBP2-HEPN interface. (**C, F**) Measured SAXS profiles (empty circles) for the isolated RnlA and RnlAΔ1–91 proteins with the theoretical profiles (red line) calculated for all-atom models of their respective crystallographic structures. χ^2 values for the comparisons between experimental and calculated SAXS profiles are indicated. Cartoon representations of the domain organization of RnlA full-length and the truncated RnlAΔ1–91 version are included below the curves. (**D**) Ribbon representation of the crystallographic structure of the RnlAΔ1–91 truncated variant. Coloring is as in panel A. (**E**) Trace representation of the superposition of full-length RnlA dimer and the RnlAΔ1–91 truncated variant. RMSD between both structures is 1.56 Å over 435 Cα.

RnlA is a ribonuclease with broad sequence specificity

To characterize the endoribonuclease activity of RnlA, the 3569 nucleotides-long RNA from bacteriophage MS2 was used in an *in vitro* RNase assay. Wild-type RnlA efficiently degrades this substrate, revealing multiple bands on a denaturing gel ranging from 500 to <50 nucleotides in size (Supplementary Figure S2). Moreover, RnlA is fully inhibited by the addition of a twofold excess of RnlB antitoxin prior to the addition of MS2 substrate. RnlB on its own does not show any ribonuclease activity under the same conditions.

The specificity of wild-type RnlA was investigated by primer extension with a 5′-^{32}P-labeled primer annealing to the coat protein mRNA on the MS2 genome. Compared to control reactions lacking RnlA, several RnlA-cleavage specific bands are observed from 1 min of incubation with RnlA onwards. The intensities of the higher molecular weight RnlA-specific bands decrease as the incubation time increases, which indicates the subsequent degradation of RNA fragments by the endoribonuclease activity of RnlA (Figure 2A). When RnlA is incubated with increasing amounts of RnlB prior to the addition of MS2 RNA substrate, the RnlA-specific cleavage bands disappear or show background intensity, thereby further confirming the presence of RnlA specific activity (Figure 2B).

Ten strong RnlA cleavage sites can be identified in the primer extension analysis. They are indicated by numbered arrows on the sequencing gel and corresponding internucleotide positions on the MS2 cDNA sequence fragment (Figure 2B, D). Most but not all of the cleavage sites occur within or at the ends of base-paired RNA stretches, according to the predicted secondary structure of this region of the MS2 RNA performed in the RNAfold WebServer (http://rna.tbi.univie.ac.at/). Cleavage sites 1–4 are remarkably close on the sequence, with cleavages 2 to 4 occurring every second nucleotide. This region of the MS2 RNA might be particularly prone to RnlA cleavage. In terms of sequence specificity, no consensus could be generated by simple sequence inspection. RnlA seems to be able to cleave

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

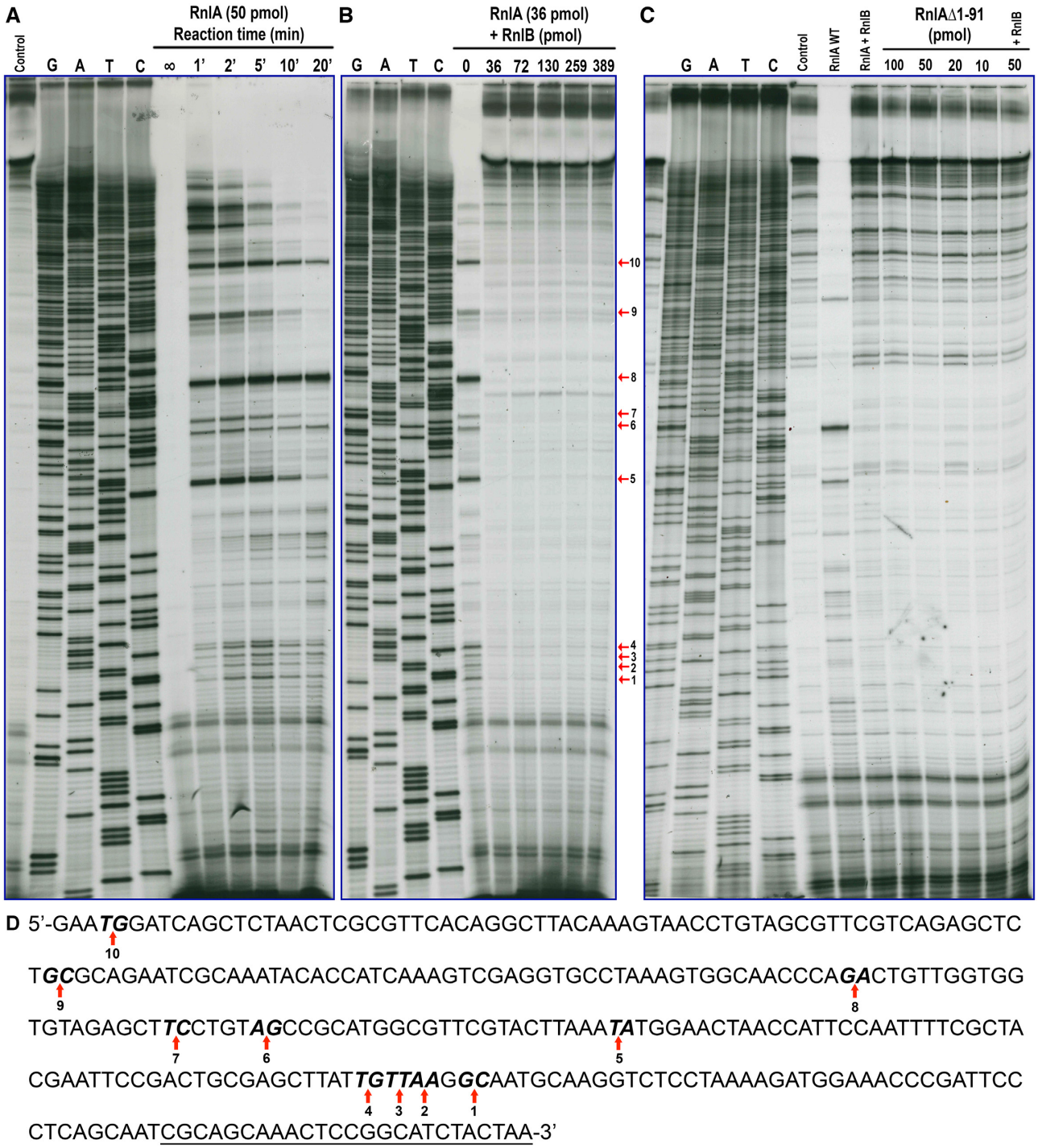

Figure 2. Specificity of RnlA ribonuclease activity. Autoradiographs of the coat-protein MS2-RNA primer extension analysis. (**A**) RnlA activity time-progress. Purified His-RnlA was incubated with MS2 RNA substrate for increasing time periods prior to primer extension analysis with 5′-^{32}P-labeled primer. All reactions were stopped by a 2.5 molar excess of RnlB antitoxin, except for one marked as infinite time. (**B**) Effect of RnlB on the endoribonuclease activity of RnlA. Increasing amounts of RnlB were incubated with RnlA for 10 min prior to the addition of MS2 substrate. Subsequent primer extension reactions were carried out as in A. RnlA-specific cleavage bands are marked by arrows and numbered. (**C**) Decreasing amounts of RnlAΔ1–91 truncated protein was incubated with MS2 RNA substrate for 30 min prior to primer extension analysis performed as in Ai. Control reactions consisting of no protein, RnlA, RnlA preincubated with RnlB antitoxin and RnlAΔ1–91 preincubated with RnlB were included for comparison in the same primer extension analysis. All reactions were carried out in parallel. Sequence ladders were obtained by the addition of the four ddNTPs to primer extension reactions. (**D**) cDNA sequence annotation of coat-protein region on MS2 RNA. The 10 RnlA-specific cleavage bands identified in B are marked by arrows and bases are in bold. The 3′ under-scripted end represents the labeled primer.

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

both downstream of purines and pyrimidines, which agrees with the broad sequence specificity previously suggested (5). Based on these data, it cannot be excluded that substrate specificity of RnlA is either determined by RNA structure rather than sequence, or that RnlA is an exonuclease where the strongest bands in the primer extension analysis correspond to pauses due to stable structure in the substrate.

Taken together these results indicate that RnlA displays ribonuclease activity in absence of any cofactors, and that stoichiometric quantities of RnlB antitoxin are sufficient to fully inhibit RnlA *in vitro*. Moreover, RnlA shows broad substrate sequence specificity, which could instead be determined by RNA secondary structure and/or the function of the TBP-like domains, which have been shown to serve as substrate binding domains in other RNases (40,46). Therefore, we next explored the role of the TBP1 domain in the activity and substrate binding properties of RnlA.

The N-terminal domain of RnlA is required for substrate binding

To assess the function of the N-terminal TBP1 domain in the ribonuclease activity of RnlA, primer extension experiments were performed with the RnlAΔ1–91 truncate lacking this domain. Surprisingly, no activity could be detected at concentrations varying from 10 to 100 pmol, nor did the addition of RnlB antitoxin influence the outcome (Figure 2C). These results are unexpected, considering the presence of a supposedly intact catalytic site in the RnlAΔ1–91 variant.

Therefore, substrate-binding experiments were performed with the RnlAΔ1–91 truncation variant. A 43-nucleotides ssDNA segment comprising RnlA-specific cleavage sites 1–4 on the MS2 RNA sequence was used in Biolayer Interferometry (BLI) experiments (Supplementary Figure S3). While the differences in conformational preferences between DNA and RNA will certainly lead to significant differences in affinity, the use of ssDNA has the advantage that cleavage is fully prevented. This approach is valid as only the protein and not the substrate is varied in this experiment and we are interested in relative rather than absolute affinities. A similar approach was earlier adopted to study substrate recognition of *Staphylococcus aureus* MazF (28).

We observe a two-order decrease in binding affinity when the TBP1 domain is absent compared to the full-length protein (Table 1, Supplementary Figure S3). These findings indicate that the TBP1 is required for substrate binding and explain why no activity is observed when RnlAΔ1–91 is used. They also provide a potential explanation for the seeming lack of RNA specificity of RnlA, namely that RNA sequence specificity is at least in part embedded in the TBP1 domain, but that the sequence that is recognized as such is too far removed from where the HEPN domain actually cuts to be readily identified. Next, in order to identify the HEPN domain residues involved in catalysis, single-alanine substitution mutants of the four residues identified as highly conserved in the RnlA/LsoA family of HEPN proteins, were generated and tested in RNase activity and substrate binding experiments.

Table 1. TBP1 is required for substrate binding and Arg255 is involved in substrate binding

Protein	K_D (M)
RnlA	$6.30 \times 10^{-6} \pm 1.07 \times 10^{-6}$
R255A RnlA	$1.97 \times 10^{-4} \pm 1.13 \times 10^{-4}$
E258A RnlA	$7.17 \times 10^{-6} \pm 1.91 \times 10^{-6}$
R318A RnlA	$6.93 \times 10^{-6} \pm 1.37 \times 10^{-6}$
RnlAΔ1–91	$4.83 \times 10^{-4} \pm 3.29 \times 10^{-4}$

Binding of RnlA and its single alanine mutants to ssDNA. Data were collected in an Octet RED96 instrument and analyzed with the manufacturer data analysis tool v. 9.0 (FORTÉBIO). K_D values are derived from the steady state analysis of the data fits to a 1:1 model and R_{Max} values are extrapolated from the fits to this model.

R255, E258, R318 and H323 are essential for activity

Single-alanine substitution mutants of the four putative catalytic residues of RnlA suggested above do not show any activity in the presence of MS2 RNA under the same conditions used for wild-type RnlA (Supplementary Figure S2). Together with the high conservation scores for these residues in the RnlA/LsoA family and their involvement in the toxicity of LsoA *in vivo* (14), this finding constitutes a strong indication of a role for R255, E258, R318 and H323 in catalysis or substrate binding.

Additionally, these mutants were included in the primer extension analysis to further confirm the lack of activity observed in the previous MS2 RNA cleavage assay. As seen on Figure 3A, cDNA amplification of MS2 RNA after incubation with each of the single-alanine RnlA mutant proteins shows the same background intensity pattern as negative controls in absence of RnlA or presence of RnlB incubation. This result confirms the absence of endoribonuclease activity for these RnlA mutants and of RnlB.

To validate our single alanine mutants *in vivo*, wild-type and mutants were expressed from *pET28a* in *E. coli* BL21(DE3). We observe an increase in lag time for growth of the cells over-expressing wild-type RnlA compared to the empty vector control. This result is in agreement with earlier measurements that indicate that RnlA causes growth defects in *E. coli* (5). No such growth delay is observed for R255A, E258A, R318A or H323A RnlA expressing cells (Figure 3B).

In order to explore the role of these four highly conserved residues in substrate binding, BLI experiments in the setup described above (see previous section) were performed with the mutant proteins. Wild-type RnlA binds to this substrate mimic with a K_D in the μM range. For the mutants E258A and R318A, this affinity does not drastically change, indicating that these two residues are not crucial for substrate binding but may have a catalytic role (Table 1 and Supplementary Figure S3). On the other hand, for mutant R255A the affinity for the substrate mimic is weakened by two orders of magnitude, which points to the involvement of this arginine residue in substrate binding (Table 1 and Supplementary Figure S3A, B). BLI analysis with mutant H323A RnlA was not possible due to protein instability at the concentrations required for the assay.

Overall, these experiments demonstrate the catalytic role of the $RX_{4\text{-}6}H$ conserved motif in the HEPN domain of

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

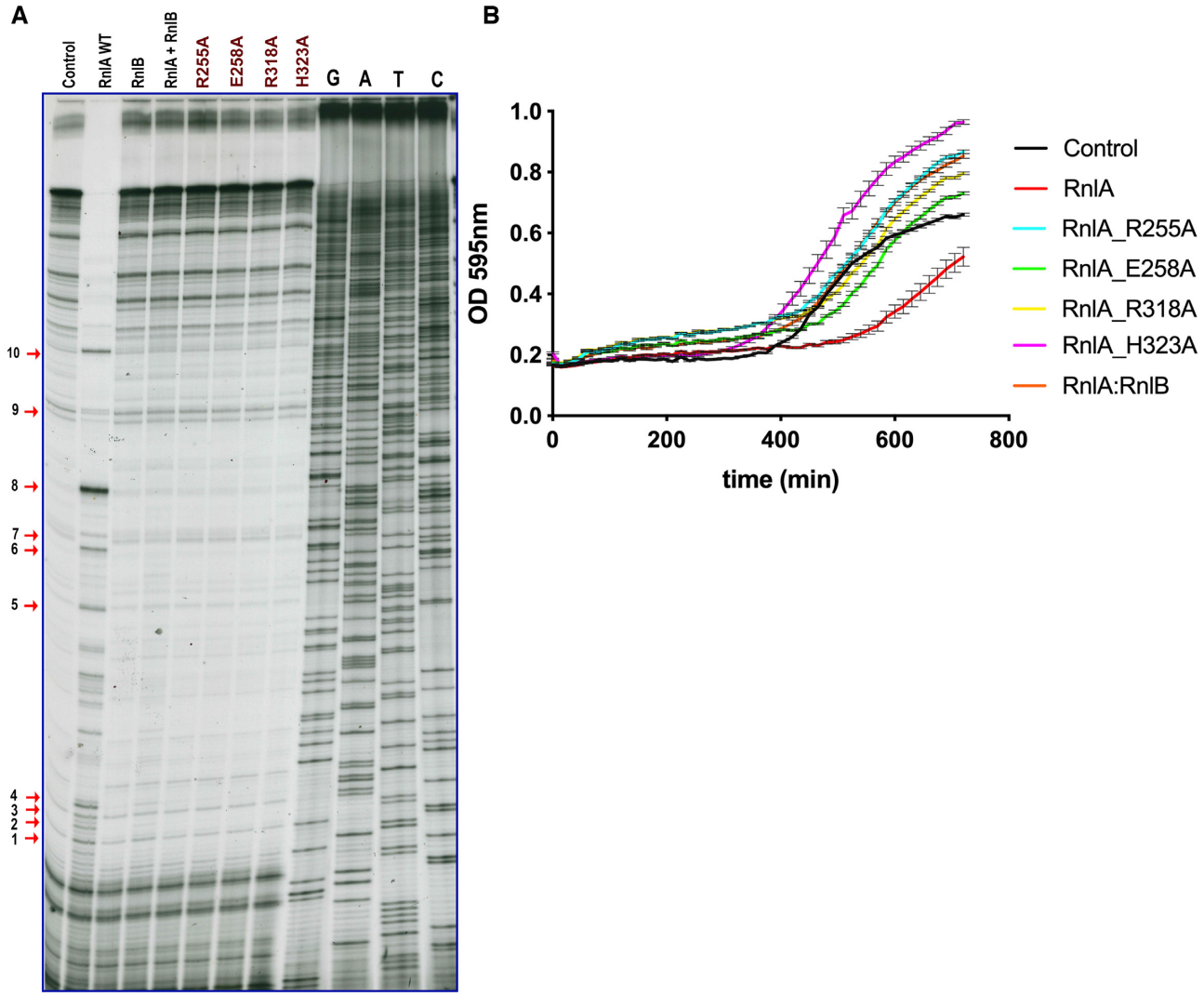

Figure 3. HEPN conserved motif defines the catalytic site of RnlA. Autoradiographs of the coat-protein MS2-RNA primer extension analysis. (**A**) R255A, E258A, R318A and H323A RnlA mutant proteins were incubated with MS2 RNA substrate for 30 min prior to primer extension analysis performed as in Figure 2. Control reactions consisting of no protein, RnlA wild-type, RnlB antitoxin and RnlA:RnlB complex were included for comparison in the same primer extension analysis. All reactions were carried out in parallel. Sequence ladders were obtained by the addition of the four ddNTPs to primer extension reactions. (**B**) BL21 (DE3) *E. coli* cells containing *pET28a* plasmids with wild-type *rnlA*, *R255A_rnlA*, *E258A_rnlA*, *R318A_rnlA*, *H323A_rnlA,* and the *rnlAB operon*, were incubated in the presence of IPTG at 37°C and OD (595 nm) was monitored every 15 min. Cells transformed with *pET28* (empty vector) were included as a control. Data represent mean ± SEM originating from $N > 3$ biological repeats.

RnlA and pinpoint the active site of the enzyme to the region involving these residues. Subsequently, in order to gain further insight into the mechanism of RnlA and the inhibition of RnlA activity by direct interaction with its antitoxin RnlB, crystallization of the complex was attempted.

RnlB antitoxin binds a canonical HEPN dimer of RnlA

To understand how RnlB antitoxin inhibits RnlA, we determined the crystal structure of the RnlA:RnlB complex at 2.6 Å resolution. This structure shows one RnlA dimer interacting with two monomeric RnlB molecules (Figure 4A). Furthermore, this hetero-tetramer is representative of the RnlA:RnlB assembly in solution as confirmed by SAXS (Supplementary Figures S4 and S11). RnlB is entirely folded and consists of a four-stranded mixed β-sheet sandwiched between two α-helices (Figure 4, Supplementary Figure S5A).

The most striking feature of the RnlA:RnlB complex is that here RnlA forms a dimer that is fundamentally different from what is seen in the structure of isolated RnlA. The dimer is again formed via the HEPN domain but resembles the V-shaped canonical dimer that has been observed in different HEPN subfamilies (Supplementary Figure S9). This novel arrangement brings R255, E258, R318 and H323 from helix α13 in both monomers together, with R318 and H323 remaining exposed but R255 becoming buried and mediating inter-monomer contacts. In this arrangement, E258 seems to help to correctly position the side chain of H323.

The new contact interface formed between the HEPN domains results in the burial of 4070 $Å^2$, which is significantly

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

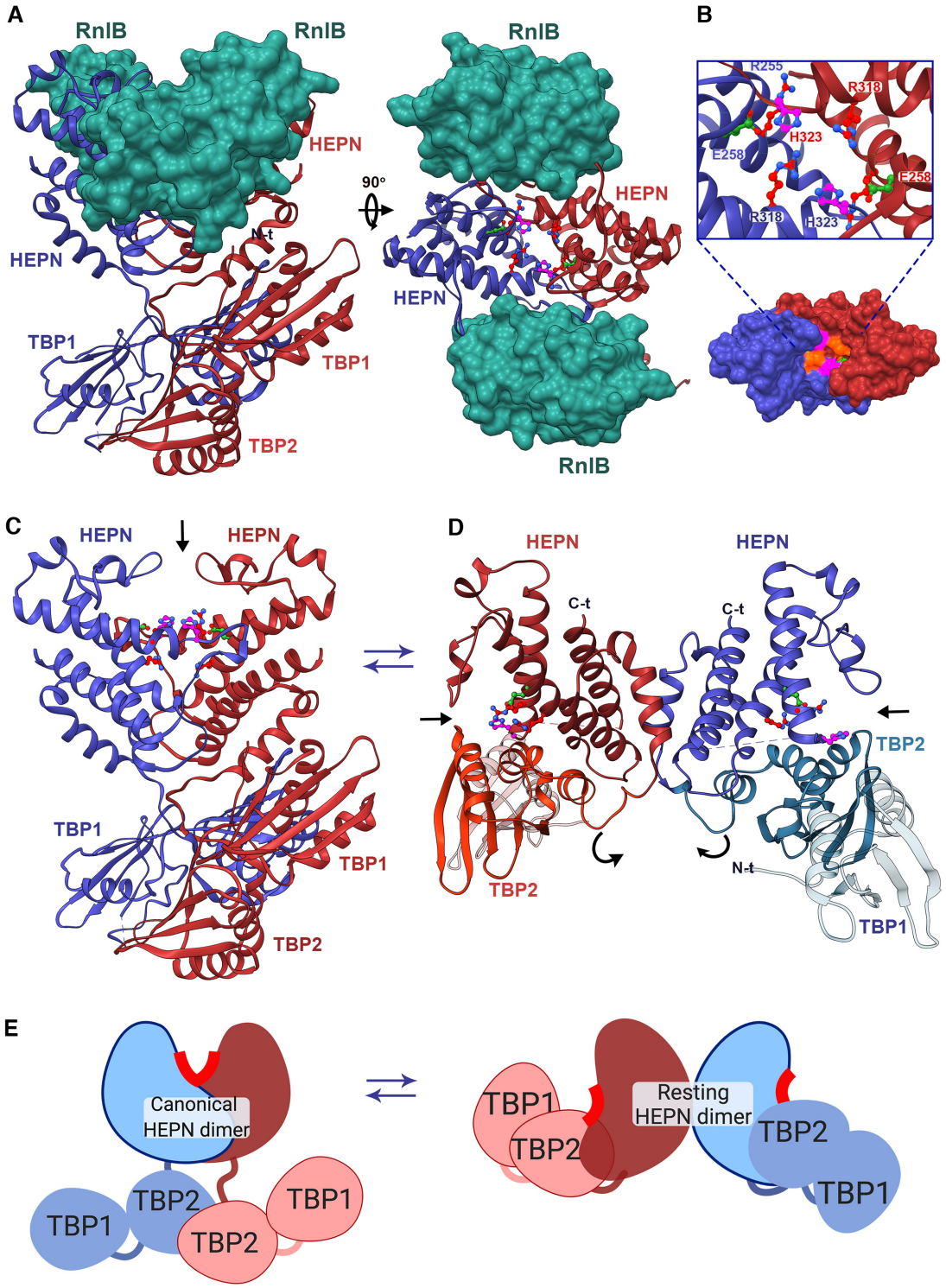

Figure 4. RnlB recognizes a canonical HEPN RnlA dimer. (**A**) Ribbon representation of the RnlA:RnlB hetero-tetrameric complex. Chains in the RnlA dimer are colored in red and blue, while the two RnlB monomers are shown in green surface representation. The three domains (TBP1, TBP2, HEPN) are indicated. Top view of the RnlA:RnlB complex. (**B**) Catalytic residues are shown in ball and stick, colored by heteroatom and labelled. Surface representation of the canonical HEPN dimer in RnlA showing the catalytic cleft where the conserved catalytic residues reside. Ribbon representation of RnlA dimers in the canonical HEPN dimer (**C**) and resting (**D**) conformations. Coloring is the same as in panel A. Black arrows indicate the position of the catalytic residues in each dimer. Catalytic residues are represented as ball and stick and colored by heteroatom. (**E**) Cartoon representation of the domain organization and dimerization modes of the canonical HEPN RnlA dimer and the resting RnlA dimer.

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

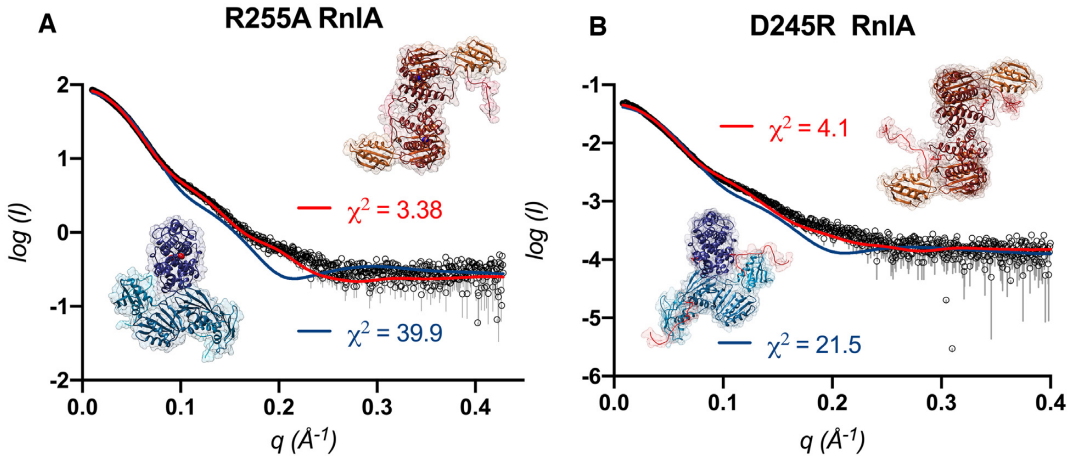

Figure 5. Mutations in residues involved in dimerization shift the equilibrium in solution between the resting RnlA conformation and the canonical HEPN RnlA dimer. Small angle X-Ray Scattering (SAXS) measurements of R255A RnlA (**A**) and D245R RnlA (**B**). All panels show the comparison of the experimental data (open circles) to the theoretical profile for the corresponding all-atoms model obtained by the program MODELLER. MultiFoXS-calculated scattering curves for resting RnlA (red curves) and canonical HEPN RnlA dimer (blue curves) conformations are compared to the experimental data for R255A and D245R RnlA. χ^2 values for comparisons to resting RnlA models are shown in red and to canonical HEPN RnlA models are shown in blue.

higher than the 1638 Å^2 of the HEPN dimerization interface in the unbound state of RnlA (Figure 4B,C). At the same time, the TBP2 domains detach from the HEPN domains and also associate into a domain-swapped dimeric arrangement that buries another 2776 Å^2. This results in a large 12-stranded cross-domain β-sheet (Supplementary Figure S6). Finally, each RnlB monomer further stabilizes the novel arrangement by bridging the HEPN dimer, thereby ordering the otherwise disordered loop T326-D336 of RnlA.

The surprising finding that RnlA adopts a canonical HEPN dimer, which is also bound by RnlB antitoxin, prompted the question of whether RnlA needs to adopt this conformation to perform catalysis. Remarkably, canonical HEPN dimerization renders a cleft between HEPN domains with the conserved catalytic residues placed at its base seemingly posed for catalysis, as observed in other HEPN ribonucleases (16,17,21). Therefore, we next performed mutagenesis of RnlA aiming at interfering with each dimerization mode without disturbing residues required for catalysis, and evaluated the resulting mutant proteins in activity assays.

The canonical HEPN dimer is the active conformation of RnlA

In order to determine whether canonical HEPN dimerization is required for activity, mutations in the HEPN domain were designed to destabilize the resting or the canonical HEPN dimer conformations, respectively. V206R RnlA was designed to interfere with the formation of the resting state dimer, as observed in the structure of isolated RnlA. Importantly, this residue is exposed to the solvent in the HEPN canonical dimer but is buried at the interface of the resting HEPN dimer. On the other hand, D245R RnlA destabilizes the canonical HEPN dimer conformation but should not interfere with the formation of the resting RnlA dimer. Additionally, the already mentioned R255A mutation would remove interactions at the canonical HEPN dimer interface and may thus stabilize the resting dimer conformation. SAXS data for both D245R and R255A RnlA as well as the crystal structure of R255A RnlA show that these mutants indeed adopt the resting RnlA dimer conformation (Figure 5, Supplementary Figures S7 and S12). The V206R mutant unfortunately is too insoluble to determine its dominant solution state by SAXS.

In vitro ribonuclease activity and primer extension assays were performed with these mutant proteins. V206R RnlA displays the same RnlA wild-type cleavage pattern in the primer extension experiments. Importantly, its activity is inhibited by the addition of RnlB antitoxin (Figure 6A), showing that it indeed can adopt the canonical HEPN dimer conformation, which is the conformation bound by the antitoxin. On the other hand, ribonuclease activity of D245R RnlA could not be detected at any concentration within a range of 10–50 pmol of protein (Figure 6A). As expected, the addition of RnlB antitoxin to this mutant protein did not cause a change in its activity.

We next tested these mutants in our *in vivo* growth assay. In agreement with these results, the V206R mutant is active *in vivo*, inducing the same growth defect as the wild-type protein. In contrast, D245R RnlA shows no active phenotype and growth of this mutant follows that of the empty vector control (Figure 6B).

Taken together, these results indicate that HEPN canonical dimerization is required for activity, in contradiction to what had been previously reported for RnlA (13). HEPN canonical dimerization leads to the formation of a composite active site generated by the juxtaposition of the highly conserved residues located within a catalytic cleft (Figure 4B and Supplementary Figure S8B).

DISCUSSION

The HEPN domain is widespread across the proteomes of bacteria, archaea and higher eukaryotes. It shows a conserved topology despite high sequence divergence. HEPN

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

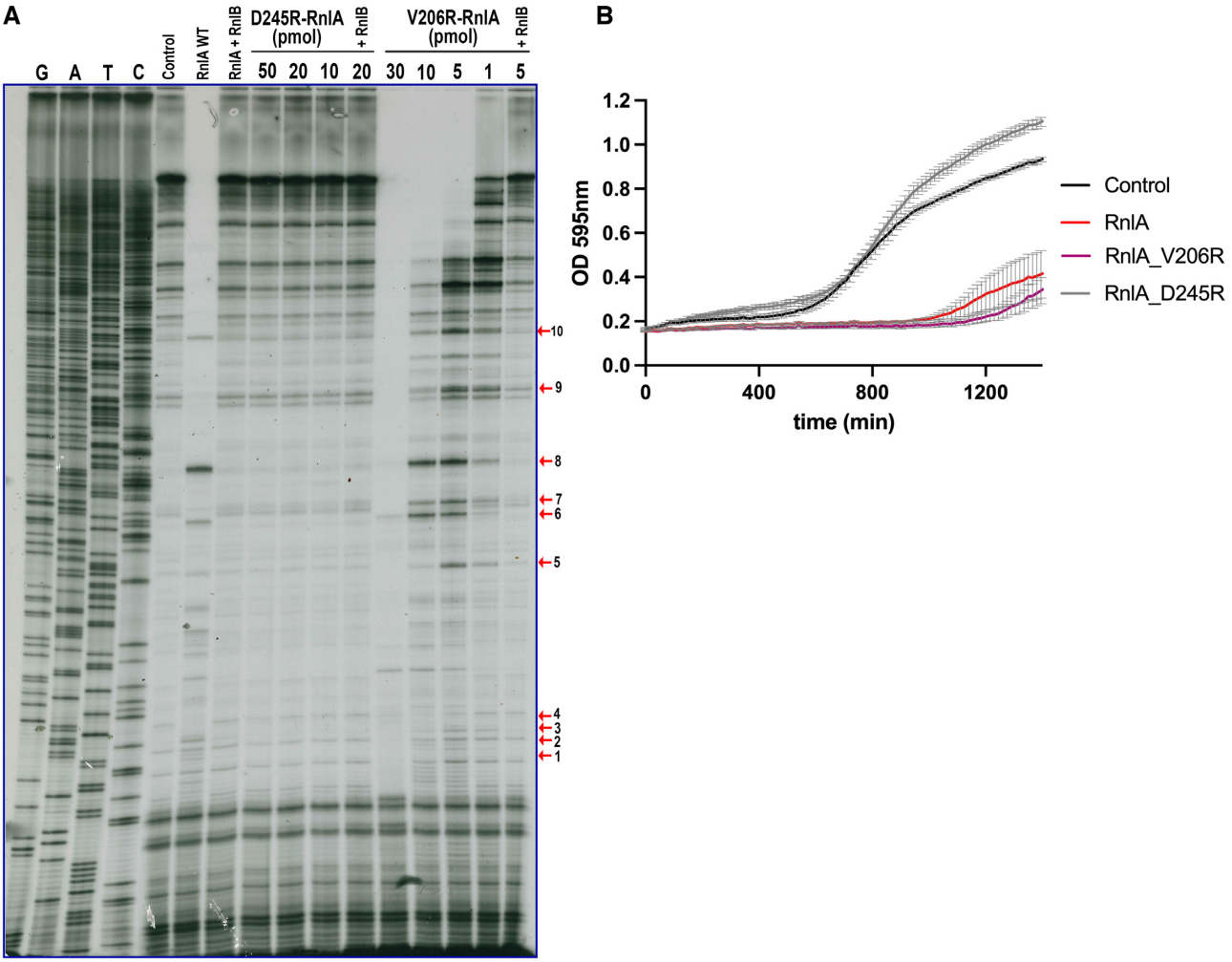

Figure 6. Canonical HEPN dimerization is required for RnlA activity. (**A**) Autoradiographs of the coat-protein MS2-RNA primer extension analysis. Decreasing amounts of interface-destabilizing mutants D245R and V206R RnlA were incubated with MS2 RNA substrate for 30 min prior to primer extension analysis performed as in Figure 2. V206R and D245R RnlA mutants preincubated with equimolar RnlB amounts were also included in the analysis. RnlA-specific cleavage bands are marked by arrows and numbered. Control reactions consisting of no protein, RnlA wild-type and RnlA:RnlB complex were included for comparison in the same primer extension analysis. All reactions were carried out in parallel. Sequence ladders were obtained by the addition of the four ddNTPs to primer extension reactions. (**B**) BL21 (DE3) *E. coli* cells containing *pET28a* plasmids with wild-type *rnlA*, *V206R_rnlA* and *D245R_rnlA* were incubated in the presence of IPTG at 30°C and OD (595 nm) was monitored every 15 min. Cells transformed with *pET28a* (empty vector) were included as a control. Data represent mean ± SEM originating from $N > 3$ biological repeats.

domains are RNases that are characterized by a conserved $RX_{4\text{-}6}H$ active site motif. They are found in bacterial anticodon nucleases (ACNases) RloC (43) and PrrC (44) involved in the restriction of T4 phage, in the kinase-extension nuclease (KEN) domain of RNase L in vertebrates (16), in domains in the abortive infection genes (Abi) system and in domains encoded by genes related to the CRISPR_Cas loci (12). Several subfamilies of HEPN domains have been identified based on their structural variability outside the common HEPN core. RnlA and LsoA represent one of these subfamilies.

HEPN domains are dimerization units. From the crystal structures of different HEPN domain containing proteins, a canonical dimer association emerges. This association results in a V-shaped structure that brings together the $RX_{4\text{-}6}H$ motifs of both chains and forms a positively charged surface to which the RNA substrate docks. The details of this association vary between subfamilies, resulting in a range of more open or closed architectures in addition to a few cases that deviate from this canonical arrangement (Supplementary Figure S9). RnlA, although being a HEPN ribonuclease, was believed not to adopt the canonical HEPN dimer. The crystal structure of RnlA in its unbound state shows an alternative non-canonical HEPN domain-mediated dimer (13) while the homologous LsoA in complex with Dmd does not dimerize (14). In these structures the $RX_{4\text{-}6}H$ motif as well as R255 and E258 (or their equivalents in LsoA) are located at the interface between the HEPN domain and the TBP2 domain (Figure 1A, B; Figure 4D and Supplementary Figure S8A). As the equivalents of R255, E258 and R318 in LsoA are essential for its activity *in vivo* and the inhibitor Dmd locates in the groove between the HEPN and the TBP2 domains of LsoA, it was assumed that this corresponded to an active conformation

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

of the enzyme where Dmd inhibits the protein by blocking the active site (14) (Supplementary Figure S10B, C).

We now observe that RnlA adopts a second HEPN-mediated dimer that resembles a canonical HEPN dimer, which has been proposed to represent the active conformation of other HEPN ribonucleases (Figure 4A–C and Supplementary movie 1). This dimer brings together the $RX_{4-6}H$ motifs from both monomers, as is the case for HEPN ribonucleases in general (17,21) (Figure 4A–C and Supplementary Figures S8B and S9). Thus, the change in quaternary structure between the unbound and RnlB-bound form of RnlA is drastic and to our knowledge has not been observed in any other protein. Not only does the dimer interface completely change with residues belonging to opposite sides of the HEPN domain monomer being involved, it also coincides with further domain rearrangements: the TBP2 domain detaches from the HEPN domain (thus making available its alternative dimer interface) and also dimerizes via swapping of its N-terminal β-strand (Supplementary Figure S6). Stabilization of the canonical HEPN dimer conformation leads to an active enzyme both in *vivo* and *in vitro*. In this arrangement, R255, which is not conserved among the HEPN superfamily, would not be a catalytic residue as suggested earlier (13), but contribute to stabilization of the canonical HEPN dimer (Figure 4A, B and Supplementary Figure S8B), in agreement with the observed defect in substrate binding of the R255A mutant.

RnlB binds to this active conformation of RnlA. Yet, RnlB does not bind to the active site groove of the HEPN domain. Rather, it is positioned adjacent to it, seemingly blocking entrance and exit to the symmetric substrate-binding groove, resulting in the inhibition of RnlA activity (Figure 4A and Supplementary Figure S5C). The side chains of E258, R318 and H323 remain at least partially exposed upon RnlB binding. This contrasts with SO_3166 from *Shewanella oneidensis*, the only HEPN protein in a toxin-antitoxin context for which a crystal structure is available, besides RnlA and its homolog LsoA (21). SO_3166 together with SO_3165 constitute a type II HEPN-MNT TA system. SO_3166 consists of a single HEPN domain that adopts the canonical HEPN-dimer arrangement, similar to the conformation in the RnlB-inhibited form of RnlA (Supplementary Figure S9). The presumed active site groove formed upon SO_3166 homodimerization is occupied by α-helix 4 of the antitoxin SO_3165, which obviously results in inhibition as the $RX_{4-6}H$ motifs are inaccessible to the substrate.

The crystal structure of the RnlA:RnlB complex shows that RnlB exclusively interacts with the HEPN domain, in contradiction with pull-down assays presented earlier (45), which suggested that RnlB interacts both with the NRD/TBP2 domain and the DBD/HEPN domain. Possibly, non-specific binding to the Ni-NTA matrix may be involved in such experiments, which unfortunately did not include appropriate controls (45).

Although the canonical HEPN dimer of RnlA has a much larger interface than the alternative non-canonical resting dimer, the latter is the dominant species in solution in absence of any ligand as evidenced by SAXS. The stability of this 'resting' dimer likely comes from interactions between the HEPN and TBP2 domains, which need to be broken before the canonical HEPN dimer can be formed. Using site-specific mutants that stabilize one of the two possible conformations, we could show that RnlA is active only when it forms the canonical HEPN dimer. A mutant that stabilizes the resting state by preventing canonical HEPN dimerization is inactive. Whether the alternative resting dimer, that exists in the crystal and in solution, has any function *in vivo*, for example, by reducing activity and thus toxicity of the protein, remains unclear.

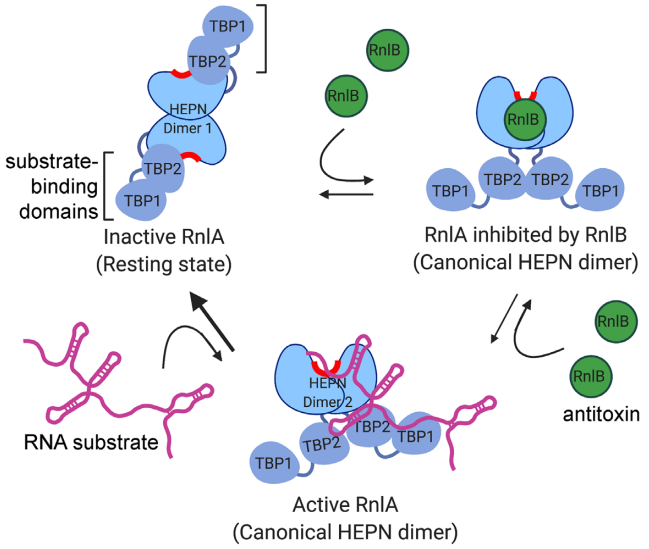

Figure 7. Model for the activity and inhibition of HEPN ribonuclease RnlA. RnlA drastically shifts domain orientation and dimerization interface between different states. RnlA free resting state involves the formation of HEPN dimer 1, which positions the conserved catalytic residues in the TBP2–HEPN interface of each chain in the dimer. The TBP1 and TBP2 adopt a TBP-like domain fold which is involved in substrate binding in other RNases. The TBP1 domain of RnlA is required for substrate binding and activity. The catalytically active RnlA conformation requires HEPN canonical dimerization (HEPN Dimer 2), which involves juxtaposition of the conserved catalytic residues in the formation of a composite catalytic cleft upon dimerization. The TBP1 and TBP2 assist in substrate binding and potentially in positioning the RNA substrate for catalysis. RnlB antitoxin recognizes the HEPN canonical dimer and binds on either side of the catalytic cleft, blocking substrate access. RnlB shifts the equilibrium of the enzyme from the free resting state to the RnlA canonical HEPN dimer.

Finally, our data for the first time point towards a functional role of the N-terminal TBP1 domain of RnlA. Deletion of this domain leads to an inactive enzyme that is defective in substrate binding. Possibly, this domain provides an anchor point on the larger RNA substrate, bringing the catalytic HEPN domain close to its substrate. As any specific recognition sequence for the TBP1 domain would be relatively far from the actual cutting site(s), this could explain the apparent lack of sequence specificity of RnlA (Figure 7). In agreement with such a function is that TBP domains are found associated with other RNases as well. In particular, the N-terminal domain of RNase H3, which is otherwise unrelated to RnlA, is such a domain. In the crystal structure of *Thermovibrio ammonificans* RNase H3 (46) (PDB ID: 4py5), which is the only one in complex with a substrate, the 15 bp hybrid DNA/RNA duplex is sandwiched between the TBP domain and the RNase H3 catalytic domain. This TBP domain thus seems to have a role in substrate position-

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

ing, strengthening our assumption for a similar role in RnlA and LsoA.

CONCLUSION

In contrast to what was concluded in earlier studies, RnlA is a bona-fide HEPN ribonuclease that requires canonical HEPN dimerization for activity. Drastic conformational changes mediate transitions between the active HEPN canonical form and the resting state that is adopted as the major conformation in solution when no RnlB or substrate is bound. This resting state may help to tune down its toxicity *in vivo*. The N-terminal TBP1 domain is required for substrate binding while interactions between the TBP2 and HEPN domains shift the conformational equilibrium of the protein towards the resting state. The antitoxin RnlB binds RnlA in its active state and prevents substrate binding via blocking the entrance and exit of the substrate binding groove (Figure 7). This study therefore reconciles the mechanism and regulation of RnlA with that of other HEPN ribonucleases.

DATA AVAILABILITY

All crystallographic data have been deposited into the Protein Data Bank (www.rcsb.org) The accession codes are the following: RnlA:RnlB, 6Y2P; RnlA, 6Y2Q; R255A RnlA, 6Y2R, RnlAΔ1–91, 7AEX. The solution scattering data has equally been deposited at the www.sasbdb.org database with the following accession codes: RnlA:RnlB, SASDHW7; RnlA, SASDHX7; RnlAΔ1–91 (TBP2-HEPN), SASDKL2; R255A RnlA, SASDHY7; D245R RnlA, SASDKS9.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Sarah Haesaerts for technical assistance and the beamline staff of the Soleil macromolecular crystallography beamlines for help during data collection and processing.

FUNDING

FWO-Vaanderen [G.0B25.15N]; Onderzoeksraad of the Vrije Universiteit Brussel [SRP13]. Funding for open access charge: VIB.
Conflict of interest statement. None declared.

REFERENCES

1. Kai,T., Selick,H.E. and Yonesaki,T. (1996) Destabilization of bacteriophage T4 mRNAs by a mutation of gene 61.5. *Genetics*, **144**, 7–14.
2. Otsuka,Y. (2004) A novel endoribonuclease, RNase LS, in *Escherichia coli*. *Genetics*, **169**, 13–20.
3. Iwamoto,A., Lemire,S. and Yonesaki,T. (2008) Post-transcriptional control of Crp–cAMP by RNase LS in *Escherichia coli*. *Mol. Microbiol.*, **70**, 1570–1578.
4. Koga,M., Otsuka,Y., Lemire,S. and Yonesaki,T. (2011) *Escherichia colirnlA* and *rnlB* compose a novel toxin-antitoxin system. *Genetics*, **187**, 123–130.
5. Otsuka,Y. and Yonesaki,T. (2012) Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins: T4 phage Dmd as an antitoxin. *Mol. Microbiol.*, **83**, 669–681.
6. Harms,A., Brodersen,D.E., Mitarai,N. and Gerdes,K. (2018) Toxins, targets, and triggers: an overview of toxin-antitoxin biology. *Mol. Cell*, **70**, 768–784.
7. Ramisetty,B.C. and Santhosh,R.S. (2017) Endoribonuclease type II toxin–antitoxin systems: functional or selfish? *Microbiology*, **163**, 931–939.
8. Yang,J., Jiang,W. and Yang,S. (2009) mazF as a counter-selectable marker for unmarked genetic modification of *Pichia pastoris*. *FEMS Yeast Res.*, **9**, 600–609.
9. Bakar,F.A., Yeo,C.C. and Harikrishna,J.A. (2015) Expression of the *Streptococcus pneumoniaeyoeB* chromosomal toxin gene causes cell death in the model plant *Arabidopsis thaliana*. *BMC Biotechnol.*, **15**, 26.
10. Chono,H., Matsumoto,K., Tsuda,H., Saito,N., Lee,K., Kim,S., Shibata,H., Ageyama,N., Terao,K., Yasutomi,Y. *et al.* (2010) Acquisition of HIV-1 tesistance in T lymphocytes using an ACA-specific *E. coli* mRNA interferase. Hum. *Gene Ther.*, **22**, 35–43.
11. Shimazu,T., Degenhardt,K., Nur-E-Kamal,E., Zhang,J., Yoshida,T., Zhang,Y., Mathew,R., White,E. and Inouye,M. (2007) NBK/BIK antagonizes MCL-1 and BCL-XL and activates BAK-mediated apoptosis in response to protein synthesis inhibition. *Genes Dev.*, **21**, 929–941.
12. Anantharaman,V., Makarova,K.S., Burroughs,A.M., Koonin,E.V. and Aravind,L. (2013) Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct*, **8**, 15.
13. Wei,Y., Gao,Z.-Q., Otsuka,Y., Naka,K., Yonesaki,T., Zang,H. and Dong,Y.-H. (2013) Structure-function studies of *Escherichia coli* RnlA reveal a novel toxin structure involved in bacteriophage resistance. *Mol. Microbiol.*, **90**, 956–965.
14. Wan,H., Otsuka,Y., Gao,Z.-Q., Wei,Y., Chen,Z., Masuda,M., Yonesaki,T., Zhang,H. and Dong,Y-H. (2016) Structural insights into the inhibition mechanism of bacterial toxin LsoA by bacteriophage antitoxin Dmd: Structure-function of toxin-antitoxin LsoA-Dmd. *Mol. Microbiol.*, **101**, 757–769.
15. Grynberg,M., Erlandsen,H. and Godzik,A. (2003) HEPN: a common domain in bacterial drug resistance and human neurodegenerative proteins. *Trends Biochem. Sci.*, **28**, 224–226.
16. Han,Y., Donovan,J., Rath,S., Whitney,G., Chitrakar,A. and Korennykh,A. (2014) Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. *Science*, **343**, 1244–1248.
17. Molina,R., Stella,S., Feng,M., Sofos,N., Jauniskis,V., Pozdnyakova,I., López-Méndez,B., She,Q. and Montoya,G. (2019) Structure of Csx1-cOA4 complex reveals the basis of RNA decay in Type III-B CRISPR-Cas. *Nat. Commun.*, **10**, 4302.
18. Pedersen,L.C., Benning,M.M. and Holden,H.M. (1995) Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry*, **34**, 13305–13311.
19. Chen,Y., Näsvall,J., Wu,S., Andersson,D.I. and Selmer,M. (2015) Structure of AadA from *Salmonella enterica*: a monomeric aminoglycoside (3′′)(9) adenyltransferase. *Acta Crystallogr.*, **D71**, 2267–2277.
20. Niewoehner,O. and Jinek,M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA*, **22**, 318–329.
21. Jia,X., Yao,J., Gao,Z., Liu,G., Dong,Y.-H., Wang,X. and Zhang,H. (2018) Structure–function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN–MNT toxin–antitoxin system. *J. Biol. Chem.*, **293**, 6812–6823.
22. Yao,J., Guo,Y., Zeng,Z., Liu,X. and Wang,X. (2015) Identification and characterization of a HEPN-MNT family type II toxin-antitoxin in *Shewanella oneidensis*: a HEPN-MNT family type II TA in *S. oneidensis*. *Microb. Biotechnol.*, **8**, 961–973.
23. Makarova,K.S., Wolf,Y.I. and Koonin,E.V. (2009) Comprehensive comparative-genomic analysis of Type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct*, **4**, 19.

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021

24. Garcia-Rodriguez,G., Talavera Perez,A., Konijnenberg,A., Sobott,F., Michiels,J. and Loris,R. (2020) The *Escherichia coli* RnlA–RnlB toxin–antitoxin complex: production, characterization and crystallization. *Acta Crystallogr.*, **F76**, 31–39.
25. Xia,Y., Chu,W., Qi,Q. and Xun,L. (2015) New insights into the QuikChangeTM process guide the use of Phusion DNA polymerase for site-directed mutagenesis. *Nucleic Acids Res.*, **43**, e12.
26. Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
27. Towbin,H., Staehelin,T. and Gordon,J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 4350–4354.
28. Zorzini,V., Buts,L., Sleutel,M., Garcia-Pino,A., Talavera,A., Haesaerts,S., De Greve,H., Cheung,A., van Nuland,N.A.J. and Loris,R. (2014) Structural and biophysical characterization of *Staphylococcus aureus* SaMazF shows conservation of functional dynamics. *Nucleic Acids Res.*, **42**, 6709–6725.
29. Kabsch,W. (2010) XDS. *Acta Crystallogr.*, **D66**, 125–132.
30. Kantardjieff,K.A. and Rupp,B. (2003) Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.*, **12**, 1865–1871.
31. Afonine,P.V., Grosse-Kunstleve,R.W., Echols,N., Headd,J.J., Moriarty,N.W., Mustyakimov,M., Terwilliger,T.C., Urzhumtsev,A., Zwart,P.H. and Adams,P.D. (2012) Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallogr.*, **D68**, 352–367.
32. McCoy,A.J., Grosse-Kunstleve,R.W., Storoni,L.C. and Read,R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr.*, **D61**, 458–464.
33. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
34. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr.*, **D66**, 486–501.
35. Terwilliger,T.C., Dimaio,F., Read,R.J., Baker,D., Bunkóczi,G., Adams,P.D., Grosse-Kunstleve,R.W., Afonine,P.V. and Echols,N. (2012) *phenix.mr_rosetta*: molecular replacement and model rebuilding with Phenix and Rosetta. *J. Struct. Funct. Genomics*, **13**, 81–90.
36. Cowtan,K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr.*, **D62**, 1002–1011.
37. Franke,D., Petoukhov,M.V., Konarev,P.V., Panjkovich,A., Tuukkanen,A., Mertens,H.D.T., Kikhney,A.G., Hajizadeh,N.R., Franklin,J.M., Jeffries,C.M. *et al.* (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.*, **50**, 1212–1225.
38. Schneidman-Duhovny,D., Hammel,M., Tainer,J.A. and Sali,A. (2016) FoXS, FoXSDock and MultiFoXS: single-state and multi-state structural modelling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.*, **44**, W424–W429.
39. Webb,B. and Sali,A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.*, **86**, 2.9.1–2.9.37.
40. Jongruja,N., You,D.-J., Angkawidjaja,C., Kanaya,E., Koga,Y. and Kanaya,S. (2012) Structure and characterization of RNase H3 from *Aquifexaeolicus*. *FEBS J.*, **279**, 2737–2753.
41. Lehmann,C., Lim,K., Chalamasetty,V.R., Krajewski,W., Melamud,E., Galkin,A., Howard,A., Kelman,Z., Reddy,P.T., Murzin,A.G. *et al.* (2003) The HI0073/HI0074 protein pair from *Haemophilus influenzae* is a member of a new nucleotidyltransferase family: structure, sequence analyses, and solution studies. *Proteins*, **50**, 249–260.
42. Kozlov,G., Denisov,A.Y., Girard,M., Dicaire,M.-J., Hamlin,J., McPherson,P.S., Brai,B. and Gehring,K. (2011) Structural basis of defects in the Sacsin HEPN domain responsible for autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS). *J Biol. Chem.*, **286**, 20407–20412.
43. Davidov,E. and Kaufmann,G. (2008) RloC: a wobble nucleotide-excising and zinc-responsive bacterial tRNase. *Mol. Microbiol.*, **69**, 1560–1574.
44. Blanga-Kanfi,S., Amitsur,M., Azem,A. and Kaufmann,G. (2006) PrrC-anticodon nuclease: functional organization of a prototypical bacterial restriction RNase. *Nucleic Acids Res.*, **34**, 3209–3219.
45. Naka,K., Qi,D., Yonesaki,T. and Otsuka,Y. (2017) RnlB Antitoxin of the *Escherichia coli* RnlA-RnlB toxin–antitoxin module requires RNase HI for inhibition of RnlA toxin activity. *Toxins*, **9**, 29.
46. Figiel,M. and Nowotny,M. (2014) Crystal structure of RNase H3–substrate complex reveals parallel evolution of RNA/DNA hybrid recognition. *Nucleic Acids Res.*, **42**, 9285–9294.

Downloaded from https://academic.oup.com/nar/article/49/12/7164/6303458 by guest on 13 September 2021