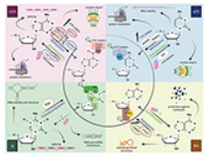
RNA Biology



ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/krnb20

Structural insights into dpCoA-RNA decapping by NudC

Wei Zhou, Zeyuan Guan, Fen Zhao, Yage Ye, Fang Yang, Ping Yin & Delin Zhang

To cite this article: Wei Zhou, Zeyuan Guan, Fen Zhao, Yage Ye, Fang Yang, Ping Yin & Delin Zhang (2021) Structural insights into dpCoA-RNA decapping by NudC, RNA Biology, 18:sup1, 244-253, DOI: 10.1080/15476286.2021.1936837

To link to this article: https://doi.org/10.1080/15476286.2021.1936837

RNA BIOLOGY
2021, VOL. 18, NO. S1, 244–253
https://doi.org/10.1080/15476286.2021.1936837

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Structural insights into dpCoA-RNA decapping by NudC

Wei Zhou[a], Zeyuan Guan[a], Fen Zhao[a], Yage Ye[a], Fang Yang[b], Ping Yin[a], and Delin Zhang [a]

[a]National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, China; [b]State Key Laboratory of Hybid Rice, College of Life Sciences, Wuhan University, Wuhan, Hubei, China

ABSTRACT

Various kinds of cap structures, such as m^7G, triphosphate groups, NAD and dpCoA, protect the 5′ terminus of RNA. The cap structures bond covalently to RNA and affect its stability, translation, and transport. The removal of the caps is mainly executed by Nudix hydrolase family proteins, including Dcp2, RppH and NudC. Numerous efforts have been made to elucidate the mechanism underlying the removal of m^7G, triphosphate group, and NAD caps. In contrast, few studies related to the cleavage of the RNA dpCoA cap have been conducted. Here, we report the hydrolytic activity of *Escherichia coli* NudC towards dpCoA and dpCoA-capped RNA *in vitro*. We also determined the crystal structure of dimeric NudC in complex with dpCoA at 2.0 Å resolution. Structural analysis revealed that dpCoA is recognized and hydrolysed in a manner similar to NAD. In addition, NudC may also remove other dinucleotide derivative caps of RNA, which comprise the AMP moieties. NudC homologs in *Saccharomyces cerevisiae* and *Arabidopsis thaliana* exhibited similar dpCoA decapping (deCoAping) activity. These results together indicate a conserved mechanism underpinning the hydrolysis of dpCoA-capped RNA in both prokaryotes and eukaryotes.

ARTICLE HISTORY
Received 16 February 2021
Revised 25 May 2021
Accepted 26 May 2021

KEYWORDS
Cap structures; m^7G; NAD cap; NudC; dpCoA decapping; adenine dinucleotide derivative caps

INTRODUCTION

RNA modifications play key roles in RNA function and metabolism, regulating and being associated with multiple cellular processes, and are related various diseases [1,2]. While most known RNA modifications occur in the internal regions, the diversity of 5′terminal modifications remains limited [3]. RNA capping at the 5′ end was considered a hallmark of eukaryotic gene expression until the discovery of an NAD cap in prokaryotic RNA [4]. In eukaryotes, a canonical 7-methylguanylate (m^7G) cap is always present at the 5′ end of mRNAs, protecting mRNA from degradation, and is crucial for pre-mRNA processing, mRNA translation initiation, nuclear transport, and stability [5,6]. In contrast to specialized eukaryotic m^7G caps, the NAD caps are found in prokaryotes and eukaryotes, ranging from bacteria [4,7] to yeast [8], mammals [9], and plants [10,11]. The NAD cap has been shown to influence RNA stability and turnover [12]. In addition to NAD, noncanonical caps such as dephospho-CoA (dpCoA), flavin adenine dinucleotide (FAD), nucleoside tetraphosphate (Np_4A) and UDP factors, including UDP-Glucose (UDP-Glc) and UDP-Nacetylglucosamine (UDP-GlcNAc), were subsequently identified at the 5′ end of RNAs [13–16]. The function of these noncanonical caps remains unknown, and researchers believe that the noncanonical cap may offer a strategy for cells to localize the capped RNAs to specific regions to form as-yet- unknown ribonucleoprotein complexes or to modulate the turnover or function of the targets of the RNAs [3].

Several kinds of hydrolases remove the 5′ end cap of RNA. For example, decapping protein 2 (Dcp2) and Nudt16 eliminate the m^7G cap from mRNA, resulting in the liberation of m^7GDP and 5′- monophosphate RNA [17–21]. The hydrolase protein NudC hydrolyzes NAD-capped RNA to NMN and monophosphorylated RNA in bacteria. Dcp2, Nudt16 and NudC all belong to the Nudix (nucleoside diphosphate linked to another moiety X) hydrolase family, which harbours a conserved Nudix motif for hydrolysing substrates [22]. In addition, another decapping enzyme, DXO/Rai1, which is involved in the quality control of the eukaryotic mRNA 5′ cap [5,23], has recently been reported to possess deNADding activity towards NAD-capped RNA in humans, acting by removing the entire NAD moiety [9].

RNA with a dpCoA modification at the 5′ end was first reported in 2009 [13], whereas specific dpCoA-modified RNAs have not yet been reported. Although researchers speculated that the dpCoA cap might play a role in RNA stability, RNA localization or gene regulation or even in mediating direct chemical reactions involving CoA groups, the physiological function of the dpCoA cap needs to be explored [13]. Studies on the function of the dpCoA cap have led to the discovery and identification of dpCoA decapping (deCoAping) enzymes. A recent study showed that mammalian DXO and SpRai1 can remove dpCoA and FAD caps from RNA *in vitro* [24]. It has also been demonstrated that Nudt2 and Nudt16 hydrolyse FAD-capped RNAs *in vitro* and that at least seven Nudix proteins possess deCoAping activity *in vitro* [25]. However, the molecular mechanism by

CONTACT Delin Zhang zdl@mail.hzau.edu.cn National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, China

Supplemental data for this article can be accessed here

which these Nudix proteins recognize and hydrolyse dpCoA caps remains unclear.

Here, we report that NudC can remove dpCoA caps from RNA *in vitro*, which increases the number of deCoAping enzymes. In addition, we determined the crystal structure of NudC in complex with dpCoA. Structural comparison of NudC complexed with different substrates suggests that the deCoAping enzyme may employ a similar mechanism to recognize the AMP moiety of adenine-containing noncanonical capped RNAs in bacteria, whereas its large substrate-binding pocket accommodates the remaining part of the dinucleotide derivative cap.

MATERIALS AND METHODS

NudC expression and purification

Proteins were expressed and purified as previously reported [26]. Briefly, 6× His -tagged full-length wild-type NudC in the pET15D vector (Novagen) was used to transform *Escherichia coli* BL21 (DE3). NudC expression was induced with 0.2 mM IPTG at 16°C overnight. The protein was purified with Ni-NTA Superflow resin (Qiagen), anion-exchange chromatography (Source 15Q, GE Healthcare), and size-exclusion chromatography (Superdex-200 Increase 10/300, GE Healthcare). The proteins were concentrated to 10 mg/mL and stored in a buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl and 5 mM dithiothreitol (DTT). The NPY1 gene from *Saccharomyces cerevisiae* (NCBI accession: NC_001139) and the Nudt19 gene from *Arabidopsis thaliana* (NCBI accession: Q94A82) were synthesized and optimized by Genewiz (GENEWIZ, Inc., China) for overexpression in BL21 (DE3) cells. The His-tagged NPY1 and Nudt19 proteins were purified through the same procedures as described above.

Crystallization, data collection and structure determination

Dephospho-CoA (Sigma) powder was dissolved to obtain a final concentration of 100 mM. Before crystallization, NudC (E178Q) was incubated with dpCoA at a molar ratio of 1:50 on ice for 30 min. Crystallization was carried out using the hanging-drop vapour-diffusion method at 18°C by mixing 1 µL of sample with an equal volume of reservoir solution. Crystals were obtained in 13% PEG 3350 and 0.1 M ammonium succinate. The crystals were then flash-frozen in liquid nitrogen using 20% (v/v) ethylene glycol as a cryoprotectant for data collection at the Shanghai Synchrotron Radiation Facility beamline BL17U1 [27].

All data sets were collected at the Shanghai Synchrotron Radiation Facility beamline BL19U or BL17U1 and processed with the HKL3000 or HKL2000 package [28]. Data processing was performed with programmes from the CCP4 [29,30]. The data collection and structure refinement statistics are summarized in Table 1. The structure of the NudC-dpCoA complex was solved by molecular replacement (MR) with the PDB structure 2GB5 as the search model using the PHASER programme [31]. The structure was manually iteratively refined with the tools PHENIX [32] and COOT [33].

Table 1. Statistics of data collection and refinement.

	NudC-dpCoA
Data collection	
PDBID	7E44
Space group	P1
Unit cell (Å)[a]	A = 60.84 B = 61.65 C = 93.89
Wavelength (Å)	0.9792
Resolution range (Å)[a]	45–2.0(2.05–2.0)
Total reflections[a]	142,958 [10,460]
Unique reflections[a]	80,253 [5,889]
R_{merge} (%)[a]	5.1 (10.1)
R_{pim} (%)[a]	5.1 (10.1)
I/σ[a]	7.5 (4.2)
Completeness (%)[a]	97.0 (95.6)
Redundancy[a]	1.8 (1.8)
Wilson B factor (Å^2)	47.5
Refinement	
R-work (%)	23.28
R-free (%)	24.52
RMS(bonds) (Å)	0.014
RMS(angles) (Å)	1.400
Ramachandran favoured (%)	96.92

[a]Values in parentheses are for the highest resolution shell.

In vitro transcription of 5′-capped RNAs

NAD-capped and dpCoA-capped RNAs were synthesized *in vitro* by T7 polymerase as previously reported with minor modifications [26]. The double-stranded DNA template was prepared by gradually annealing two single-stranded DNA oligonucleotides

(5′-GATCACTAATACGACTCACTATTAGTCGTCGTCGTC-3′ and 5′-GACGACGACGACTAATAGTGAGTCGTATTAGTGATC-3′, transcription start site is underlined). The transcription reactions (1 mL) were performed in a buffer containing 100 mM Tris-HCl (pH 8.0), 20 mM $MgCl_2$, 2 mM spermidine, 0.01% Triton X-100, 40 mM DTT, 1 µM DNA template, 4 mM NAD (4 mM dpCoA), 4 mM rGTP, 4 mM rUTP and 0.1 mg/mL T7 RNA polymerase at 37°C for 4 h. Short NAD-RNA or dpCoA-RNA was loaded onto a 1 mL Mono Q column (GE Healthcare), followed by gradient elution with 0.5 M NaCl in 25 mM Tris-HCl (pH 8.0). The eluted fractions were examined with a NanoDrop spectrophotometer (Thermo) and by mass spectrometry (Ultraflex II MALDI-TOF, Bruker). Verified RNA was used in the *in vitro* NudC cleavage experiment.

Determination of dpCoA and dpCoA-capped RNA hydrolysis *in vitro*

The hydrolysis of NAD and NAD-capped RNA was analysed as previously reported [26]. The hydrolysis of dpCoA or dpCoA-capped RNA by NudC variants was performed in a 200 µL reaction containing 50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 150 mM NaCl, 2 µM NudC protein and 1 mM dpCoA or 100 µM dpCoA-capped RNA. The reactions were incubated at 37°C for 30 min and quenched by the addition of 10 µL of 50 mM EDTA (pH 8.0) for the following chromatographic analysis. For hydrolysis of dpCoA, a 10 µL reaction sample was applied to an Agilent 1200 Series HPLC system equipped with a multiwavelength detector using an Agilent

TC-C18 column (250 mm, 4.6 mm). The sample was then eluted with a gradient of 20% methanol in 20 mM ammonium acetate at a flow rate of 0.5 mL/min for 40 min. The elution profiles were recorded at 254 nm, and the peaks corresponding to the product or substrate were integrated. For hydrolysis of dpCoA-capped RNA, a 200 µL reaction sample was applied to a 1 mL Mono Q column (GE Healthcare) with a multiwavelength detector (260 mm, 230 mm) and then eluted with a gradient of 0.2 M NaCl in 25 mM Tris-HCl (pH 8.0). The elution profiles were recorded at 260 nm, and the peaks corresponding to the product or substrate were integrated. The activity of NudC variants was defined as the proportion of substrate hydrolysed during the reaction time. The relative activity was calculated by normalizing the measured activity to the activity of wild-type NudC.

Measurement of the hydrolytic rate of NudC and its homologs towards dpCoA-capped RNA *in vitro*

The hydrolysis of dpCoA-capped RNA by NudC and its homologs was performed in a 100 µL reaction containing 50 mM Tris–HCl (pH 8.0), 5 mM $MgCl_2$, 150 mM NaCl, 0.5 µM protein and 50 µM RNA substrates. The reactions were incubated at 37°C for 0, 5, 10, 30, 60, and 120 min and quenched by the addition of 900 µL of 50 mM sodium acetate trihydrate (pH 4.6), followed by centrifugation at 20,000 × g for 5 min and application in ion exchange chromatography (MonoQ 5/100 GL column, GE Healthcare). Elution profiles were recorded at 260 nm, and peaks were integrated using UNICORN software (GE Healthcare). The hydrolytic rate of the protein was calculated from the portion of nondigested substrate RNA during the reaction time.

RESULTS

NudC cleaves dpCoA and dpCoA-capped RNA *in vitro*

In addition to NAD, another nucleotide metabolite, dpCoA, has also been detected on the 5′ end of RNAs in at least two bacterial species by LC-MS/MS [13]. It has been reported that NudC can hydrolyse NAD, generating nicotinamide mononucleotide (NMN) and adenosine 5′-monophosphate (AMP). To examine whether NudC could hydrolyse dpCoA, we purified recombinant NudC and tested the hydrolytic activity of NudC towards dpCoA by high-performance liquid chromatography (HPLC). NudC hydrolysed dpCoA in the presence of Mg^{2+}, generating adenosine 5′-monophosphate (AMP) (Figure 1A, black lines). As expected, a catalytically inactive mutant of NudC (E178Q) lost its hydrolytic activity towards dpCoA (Figure 1A, blue lines).

Next, to test whether NudC can recognize and hydrolyse dpCoA-capped RNAs, we synthesized a short dpCoA-capped RNA (dpCoA-G-U) by *in vitro* transcription using T7 RNA polymerase. NudC or its mutant variants (E178Q) were incubated with the dpCoA-RNA molecules. The reaction products were subsequently analysed by anion-exchange chromatography. The results revealed that NudC hydrolysed dpCoA-capped RNA (dpCoA-G-U) to 5′-monophosphate RNA (p-A-G-U) (Figure 1B, black lines), whereas the inactive mutant of NudC (E178Q) exhibited impaired hydrolytic activity. (Figure 1B, red lines). These results together suggested that NudC possesses deCoAping activity *in vitro*.

DpCoA-RNA is preferred as a substrate over dpCoA *in vitro*

It has been reported that NudC is not efficient in hydrolysing NAD and that NudC prefers NAD-RNA over NAD *in vitro* [34,35]. Thus, we speculate that the biological substrate of NudC might be dpCoA-RNA rather than dpCoA. To test our hypothesis, we compared the rates of hydrolysis of dpCoA-RNA, dpCoA, NAD-RNA and NAD by NudC (Figure 1C). Consistent with the reported research, NudC hydrolysed NAD-RNA much faster than it hydrolysed free NAD. Similarly, dpCoA-RNA was hydrolysed faster than free dpCoA. In addition, NudC seems to prefer hydrolysing dpCoA-RNA over NAD-RNA *in vitro*. These results indicate that NudC may play a significant role in removing dpCoA caps in vivo.

Structural basis for recognition of dpCoA by NudC

To dissect the molecular basis of the recognition and removal of the RNA 5′-end dpCoA caps mediated by NudC, we used the NudC (E178Q) mutant, which showed catalytic inactivity towards dpCoA to co-crystalize with dpCoA and determined the crystal structure of the NudC (E178Q) mutant in complex with dpCoA at 2.0 Å resolution. The complex structure shows a tight symmetric homodimer (Figure 2A). Two dpCoA molecules are observed in the homodimer of NudC, residing at the dimer interface, and only one of them has high-quality electron density for the entire molecule. (Figure S1). Structural superposition showed that the two NudC protomers adopt a nearly identical conformation, with a root-mean-square-deviation (RMSD) of 0.329 Å for the 223 Cα atoms (Figure S2A). Each NudC protomer is composed of an N-terminal domain (NTD), a C-terminal domain (CTD), and a zinc-binding motif that connects the NTD with the CTD (Figure 2B). The NTD is arranged crosswise between the two NudC protomers of the homodimer, comprising one α-helix (α1) and seven β-strands (β1-β7). The CTD adopts a canonical Nudix fold with a six-stranded mixed sheet (β11-β16) sandwiched between two remote perpendicular α-helices (α2 and α3). The conserved Nudix motif (residues 159–181) is located in a loop-helix-loop structure formed by helix α2 and the adjacent loops. The zinc motif is composed of three antiparallel β-strands (β8-β10) and a zinc ion that is coordinated with four cysteine residues (C98, C101, C116, and C119). The first three conserved cysteine residues have been reported to be important for the dimerization of NudC [26].

The residues mediating dimer formation or interaction with dpCoA in the catalytic pocket were analysed by PDBePISA and Ligplot [36,37] (Table S1 and Figure S2B-C). The dimer interface is primarily formed by hydrogen bonds of 18 residues from each protomer and consists of two regions, from residues R84 to Q126 and E162 to S191 (Table S1); these residues form an extensive interaction

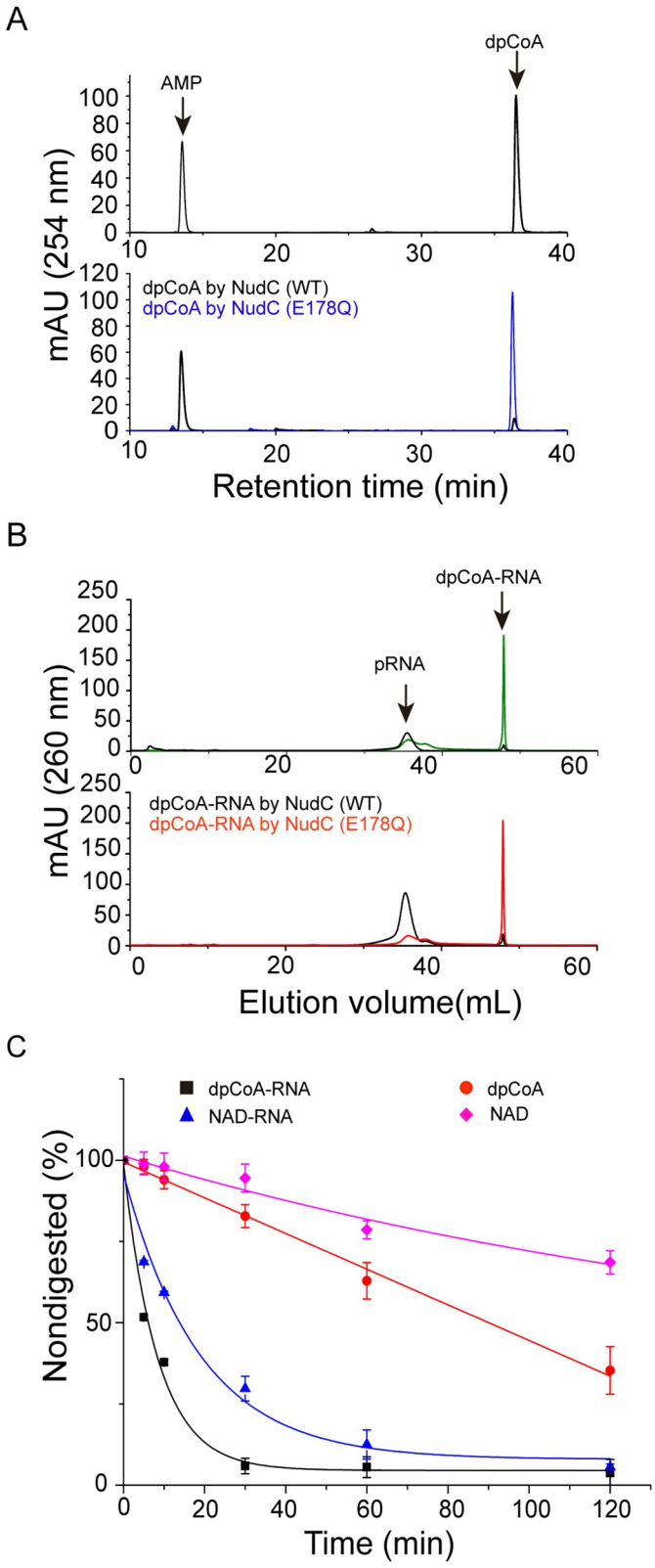


Figure 1. NudC can hydrolyse dpCoA and dpCoA-capped RNA *in vitro*. (A) NudC-mediated hydrolysis of dpCoA was detected by high-performance liquid chromatography (HPLC). (B) NudC-mediated decapping of dpCoA-RNA was detected by ion-exchange chromatography. The elution profiles of theAMP and dpCoA standards were recorded at 254 nm. dpCoA-RNA and 5′-monophosphate RNA were detected at 260 nm. All the standard samples are shown in the upper panel. (C) Comparison of the hydrolysis rates of dpCoA-RNA (50 μM), NAD-RNA (50 μM), NAD (1 mM) and dpCoA (1 mM) by 0.2 μM NudC. The data points represent the mean ± standard deviation (s.d.); n = 3.

network that is difficult to disrupt [26]. An electrostatic surface view of the dpCoA-binding pocket of NudC showed that the entrance to the catalytic pocket is extensively positively charged (Figure 2C). DpCoA is stabilized in the pocket mainly through hydrophobic interactions (Figure 3A). A close-up view of the dpCoA-binding region revealed a hydrophobic pocket, which accommodates the sulphhydryl group of dpCoA. Notably, several hydrophobic residues were identified in this pocket region, including I132, A158, W194, M201, T239, and V240 (Figure 3B and C). In addition to hydrophobic interactions, the hydroxyl group of T155 forms a hydrogen bond with the carbonyl oxygen of dpCoA (Figure 3C). The adenine base of the AMP moiety is stacked via π-π interactions with the side chain of F160 from one protomer and Y124 from the other protomer in dimeric NudC (Figure 3B and C). The 2-hydroxyl group of the adenine ribose interacts with E111 and Y124 on the same protomer to buttress the stacking interactions with the adenine base (Figure 3B and C).

To verify the structural findings, we classified the residues into three categories: 1) residues involved in catalysis; 2) residues participating in AMP moiety recognition; and 3) residues participating in the remaining recognition. We mutated the relevant residues and examined their effects on the hydrolytic activity towards dpCoA. Glu residues in the core of the Nudix motif, REx_2EE, play an important role in binding essential divalent cations and catalysing the hydrolysis of diphosphates linked to some other moiety [22]. As expected, the E178Q mutant abolished the catalytic activity (Figure 3D). Alanine substitution of two residues, Y124 and F160, which are responsible for recognition of the AMP moiety, significantly diminished hydrolysis activity. The E111A mutant exhibited no effect on NudC activity, suggesting a minor role for the dpCoA ribose moiety in substrate recognition. This observation is in accordance with the role of the AMP moiety in NAD recognition by NudC (Figure 3D), indicating conserved AMP moiety- mediated recognition of different substrates by NudC. In addition, I132, A158, G159, and A241 seem to have specific effects on NudC activity, since mutation of these residues markedly reduced NudC activity towards both substrates, which indicates that the hydrophobic substrate-binding cavity is highly important for accommodating substrates (Figure 3D). Large side-chain substitution may also hinder proper entry of the substrate into the cavity. T155A, T239A, and V240A had almost no effect on NudC activity (Figure 3D), indicating that the recognition of dpCoA does not occur through the sulphhydryl group of dpCoA. Surprisingly, we found that W194A and M201A had strong effects on NAD hydrolysis but not on dpCoA (Figure 3D). This may have resulted from the structural difference between NAD and dpCoA. The benzene ring of nicotinamide has a strong hydrophobic interaction with the side chain of W194, thus rendering the NMN part inflexible in the bound structure. Similar results were observed for the hydrolysis of dpCoA-capped RNA. As expected, Y124A, F160A, I132D, and A241K significantly diminished the hydrolysis of dpCoA-capped RNA (Figure S3). Taken together, these results suggest that the hydrophobic substrate-binding cavity of NudC is important for accommodating dpCoA-RNA and that cap

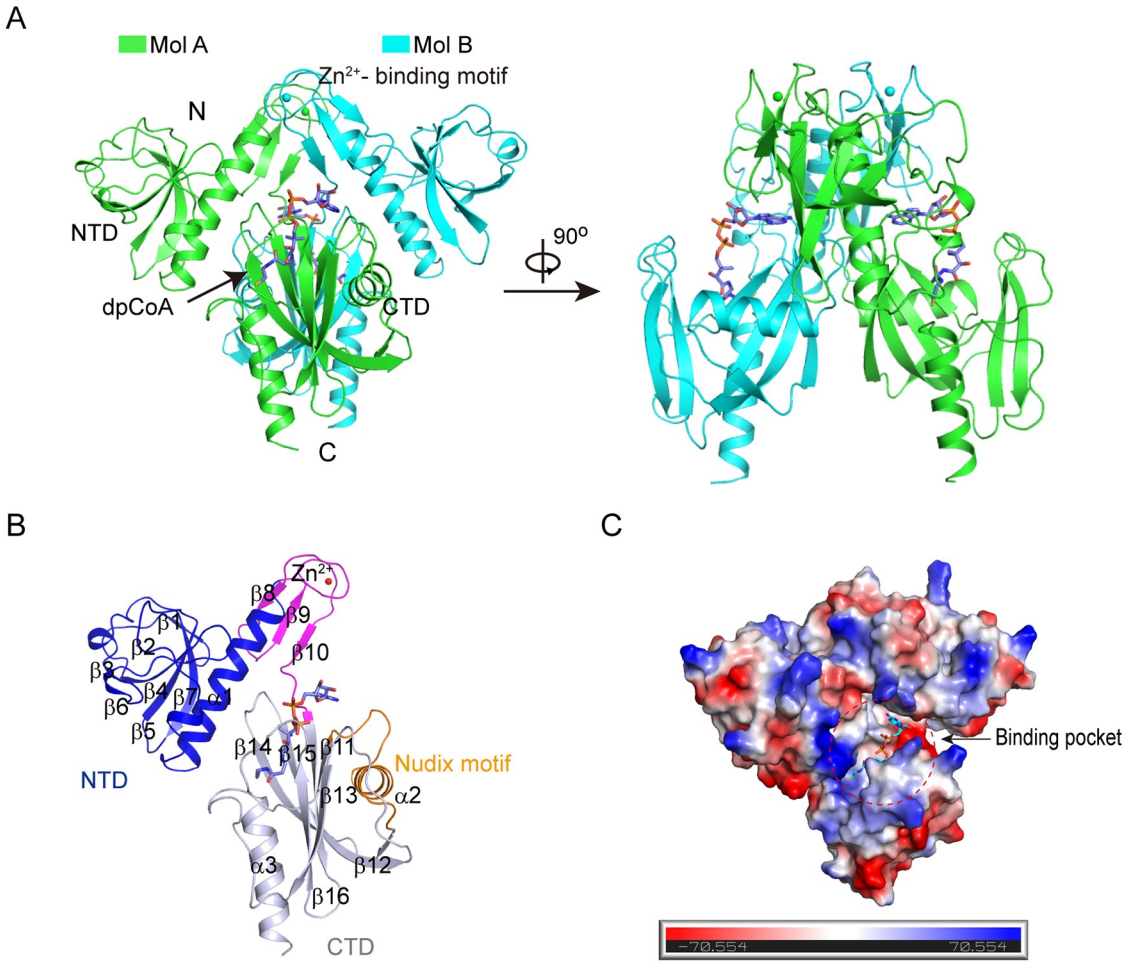


Figure 2. Crystal structure of NudC in complex with dpCoA. (A) Overall structure of NudC in complex with dpCoA. The NudC-dpCoA complex contains two monomers, which are represented as Mol A (coloured green) and Mol B (coloured cyan). (B) Crystal structure of the NudC monomer. Each monomer has an N-terminal domain (NTD) and a C-terminal domain (CTD). The two domains are connected by a zinc-binding motif. DpCoA bound to the CTD domain is shown as a slate stick representation. The atoms are coloured according to the elements as follows: carbon, slate; nitrogen, blue; oxygen, red; phosphorus, orange. (C) Electrostatic surface view of the dpCoA-binding pocket of NudC.

recognition by NudC occurs mainly through binding with the AMP moiety.

Comparison of the substrates within the NudC complex

To further investigate whether NudC binds dpCoA in a manner similar to NAD, we compared the structures of apo NudC (PDB: 2GB5), NudC-NAD (PDB: 5ISY), and NudC-dpCoA (Figure 4A). Overall, the three structures of NudC exhibit no substantial divergence. The root-mean-square deviation (RMSD) value of dpCoA-bound and apo NudC is 0.574 Å over 473 Cα atoms, while that of NAD-bound and dpCoA-bound NudC is 0.495 Å over 443 Cα atoms. However, some regions of NudC in the three structures exhibit wobbling where dpCoA-bound NudC is nearly superimposed upon NAD-bound NudC but differs from apo NudC. These regions comprise a loop between β12 and β13 (residues 146–157), and a portion located near the C-terminal helix (α3) of NudC (residues 239–256) (Figure 4B and C). In the areas where differences are concentrated, the RMSD value of dpCoA-bound and apo NudC (0.383 Å over 21 Cα atoms) is larger than that of NAD-bound and apo NudC (0.216 Å over 17 Cα atoms). Notably, they are all located in the vicinity of the substrate-binding pocket, raising the possibility that NudC undergoes conformational changes upon the binding of different substrates, although the orientations of the residues in the pocket side remain unchanged (Figure 4D). We traced residues on the cavity region and found that some of them, including R69, E219, and H149, are highly flexible with weak electron density. The flexibility of the region may result in different cavity sizes. Therefore, we examined the size of the cavity in the apo NudC, NudC-NAD, and NudC-dpCoA structures (Figure 5A-C). The volume of the dpCoA cavity (approximately 376 $Å^3$, calculated using POCASA1.1 [38]) (Figure 5B) is larger than that of the NAD cavity (232 $Å^3$) (Figure 5C). As exemplified by R69, H149, E174, S199, E219, and P235, the distances between the side chains of R69 and P235 in the dpCoA cavity are almost twofold higher than those in the NAD cavity (Figure 5E and F), thereby leading to a larger cavity. The larger cavity of dpCoA indicates to us that the rigidity of the substrate might be influenced upon binding to NudC. Therefore, we examined the rigidity of dpCoA and NAD in the NudC-bound structures by molecular dynamic simulation (Figure S4). The results showed that

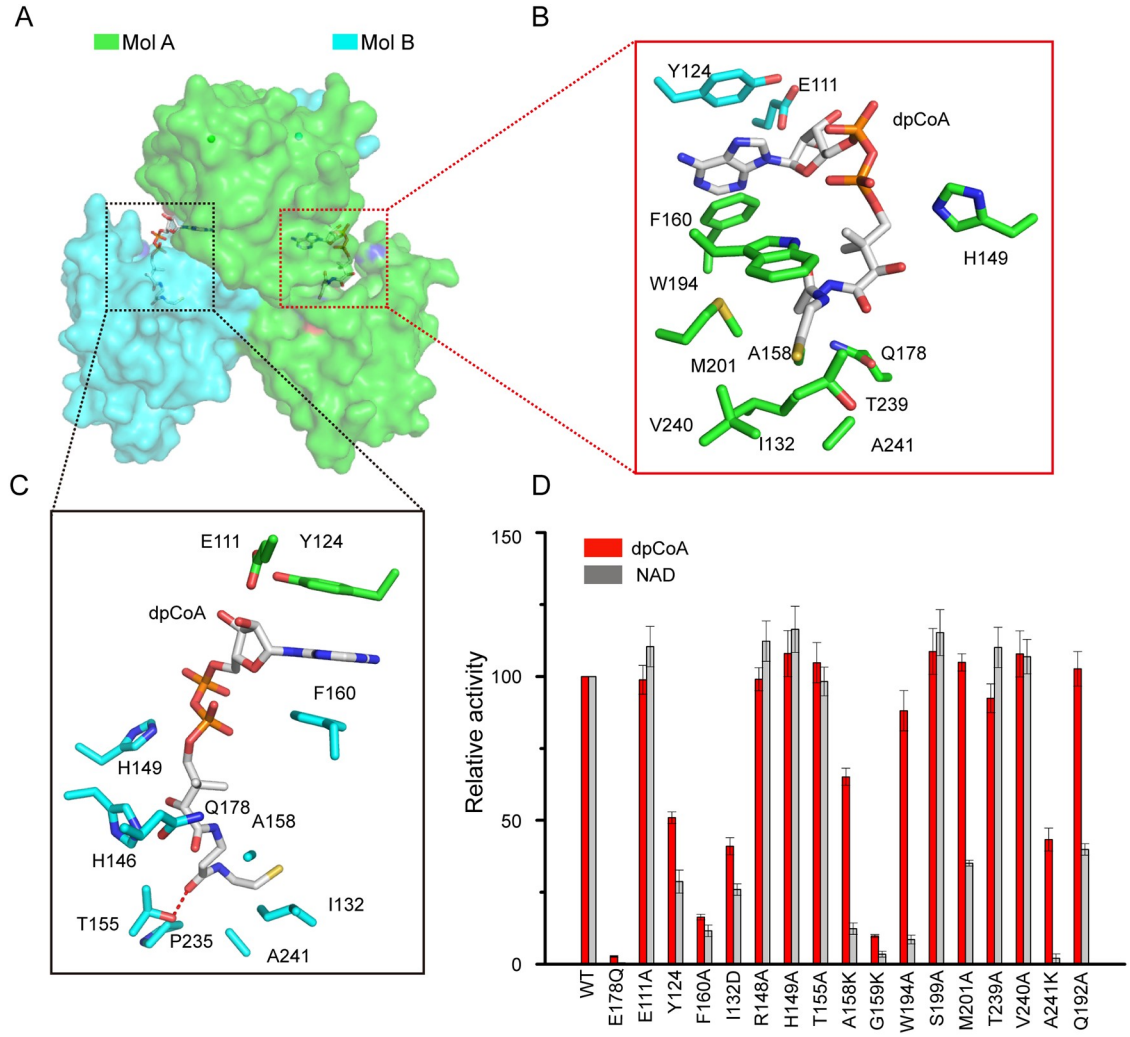


Figure 3. Close-up view of the dpCoA binding site. (A) Surface representation of the NudC-dpCoA complex. (B-C) Close-up view of the dpCoA binding pocket. The residues in the pocket are shown as sticks with nitrogen in blue and oxygen in red. The hydrogen bonds are represented by red dashed lines. The Q178 residue corresponds to E178 of the wild-type protein. (D) Hydrolytic activity of NudC and the mutants towards dpCoA and NAD. The activity was measured by HPLC. Each experiment was independently repeated three times; average values and standard deviations are shown.

dpCoA is more dynamic than NAD, with an average RMSD value of 2.401 Å, rather than 1.880 Å, over 256 Cα atoms within 200 ns. These observations suggest that when the adenine base of dpCoA is stacked with two residues (Y124 and F160), the rest of dpCoA could be flexible and the nicotinamide part of NAD could be relatively rigid. The larger dpCoA cavity may have resulted in a more flexible structure of dpCoA compared to NAD. Above all, the large cavity of apo NudC (599 $Å^3$) (Figure 5A) indicates that in addition to dpCoA and NAD, other substrates may also be accommodated in the cavity.

NudC homologs also exhibit deCoAping activity

Given the prevalence of dpCoA in the three kingdoms, the dpCoA capping/decapping processes for RNA might be evolutionarily conserved. To examine the decapping activity of NudC homologs towards dpCoA-capped RNA, we used NPY1 from *S. cerevisiae* and Nudt19 from *A. thaliana* for the hydrolysis assay (Figure S5). Both NudC homologs hydrolysed the substrate to monophosphorylated RNA *in vitro* (Figure 6A-B). To determine whether the reactivity of NPY1 and Nudt19 resembles that of NudC, we compared the rates of hydrolysis of dpCoA-RNA by NudC, NPY1, and Nut19 (Figure 6C). We found that NudC could hydrolyse dpCoA-RNA at a faster rate than the other two enzymes, especially Nudt19, with which half of the substrate still remained at 30 min (Figure 6C). A recent study found that mammalian DXO and seven Nudix proteins possess deCoAping activity *in vitro*, which corroborates our speculation [24,25]. These results suggested that dpCoA-mediated RNA processing might be a universal mechanism in eukaryotes.

Discussion

Recent studies have shown diverse noncanonical caps on the 5′ end of RNAs [14]. Among these noncanonical caps, NAD-capped RNAs have been detected in bacteria, yeast, mammals and plants [12]. The prokaryotic NudC protein and its homologs in eukaryotes can eliminate 5′ NAD modification from

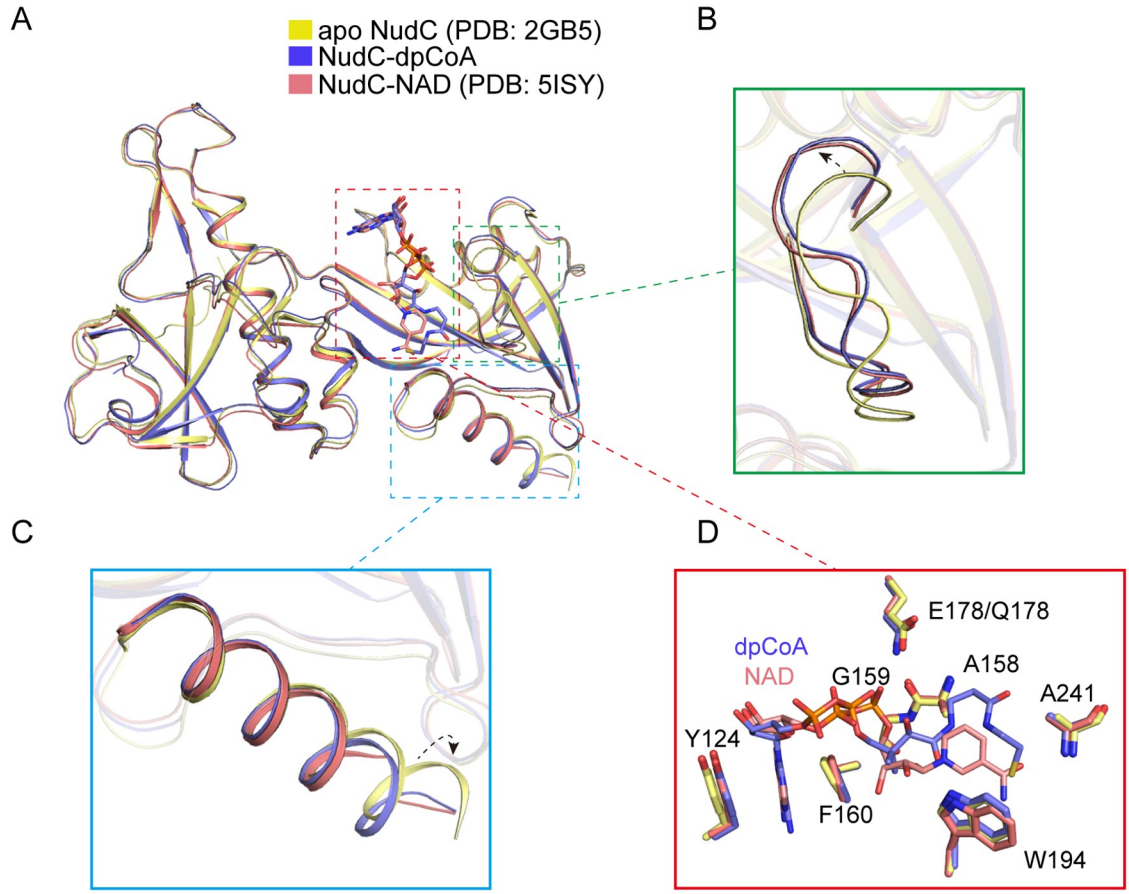


Figure 4. Comparation of NudC apo, NudC-dpCoA and NudC-NAD structures. (A) Cartoon representation of the apo NudC (PDB: 2GB5, Mol A), NudC-dpCoA (Mol A) and NudC-NAD (PDB: 5ISY, Mol A) structures. Apo NudC is coloured yellow, NudC-dpCoA is coloured slate, and NudC-NAD is coloured salmon. Alignments of different structures are displayed around the overall structural alignment framed by squares. (B) Close-up view of loop (147–156 residues), showing that this loop has a changed orientation in the three structures. (C) Close-up view of α3 (239–256 residues), showing that this helix also has a changed orientation in the NudC-dpCoA and NudC-NAD structures. (D) Close-up view of the residues in the binding pocket of the three structures. Residues in the pocket remain at almost the same position in the three structures. The Q178 residue corresponds to E178 of the wild-type protein.

RNA. Here, we show that NudC can cleave dpCoA-capped RNA *in vitro*. The crystallographic and biochemical data presented here support the role of NudC as a bacterial decapping enzyme for at least two kinds of noncanonical caps, namely, NAD and dpCoA caps. Compared to NudC, mammalian Nudt12, a homolog of NudC, also acts as a decapping enzyme for M^7G-capped RNA [39]. The overall structure of NudC-dpCoA is quite similar to the recently reported structure of human Nudt12 in complex with M^7GTP and two Cd^{2+} ions (PDB:6SCX) (Figure S6C) and to that of mouse Nudt12 in complex with AMP and three Mg^{2+} ions (PDB:6O3P) (Figure S6D) [39,40]. The major difference between NudC in complex with dpCoA and the other two structures is the absence of metal ions and some alternative loop conformations. The noncanonical caps in these structures have a common adenine. This adenine base is stacked with a tyrosine from one protomer and a phenylalanine from the other protomer. The adenine moiety is buried in a hydrophobic pocket formed at the dimer interface. Interestingly, the adenine base recognized via π-stacked interactions was also observed in Nudt16 with FAD [25], m^7GpppA [41], inosine monophosphate(IMP) [42], and ADP-ribose [43], suggesting that the stacking effect of the adenine base may be a universal mechanism underpinning substrate recognition by most Nudix proteins, including NudC.

Although NudC acts as a decapping enzyme, it binds to RNA in a non-specific manner [35]. The surface pockets of NudC are extensively positively charged around the entrance to the catalytic centre, which suggests a plausible role in RNA binding (Figure 2C). We found that NudC seemed to prefer hydrolysing dpCoA-RNA over NAD-RNA *in vitro* (Figure 1C). Thus, it is likely that NudC mediates the RNA metabolism of dpCoA-RNA in a different manner than NAD-RNA *in vivo*, and this needs to be further validated.

Given that NudC-dpCoA and NudC-NAD (as well as Nudt12-M^7GTP and Nudt12-AMP) exhibited high structural similarity, we propose a model for the recognition and cleavage of adenine dinucleotide derivative-capped RNA (NAD-capped RNA and dpCoA-capped RNA) by NudC (Figure 7). First, before entry of the substrate, two NudC protomers form a homodimer and expose the interface and catalytic pocket. After the substrate enters the pocket of one protomer, the side chain of F160 from the other protomer stacks with Y124 from the pocket protomer to stabilize the adenine moiety of the substrate. Once the

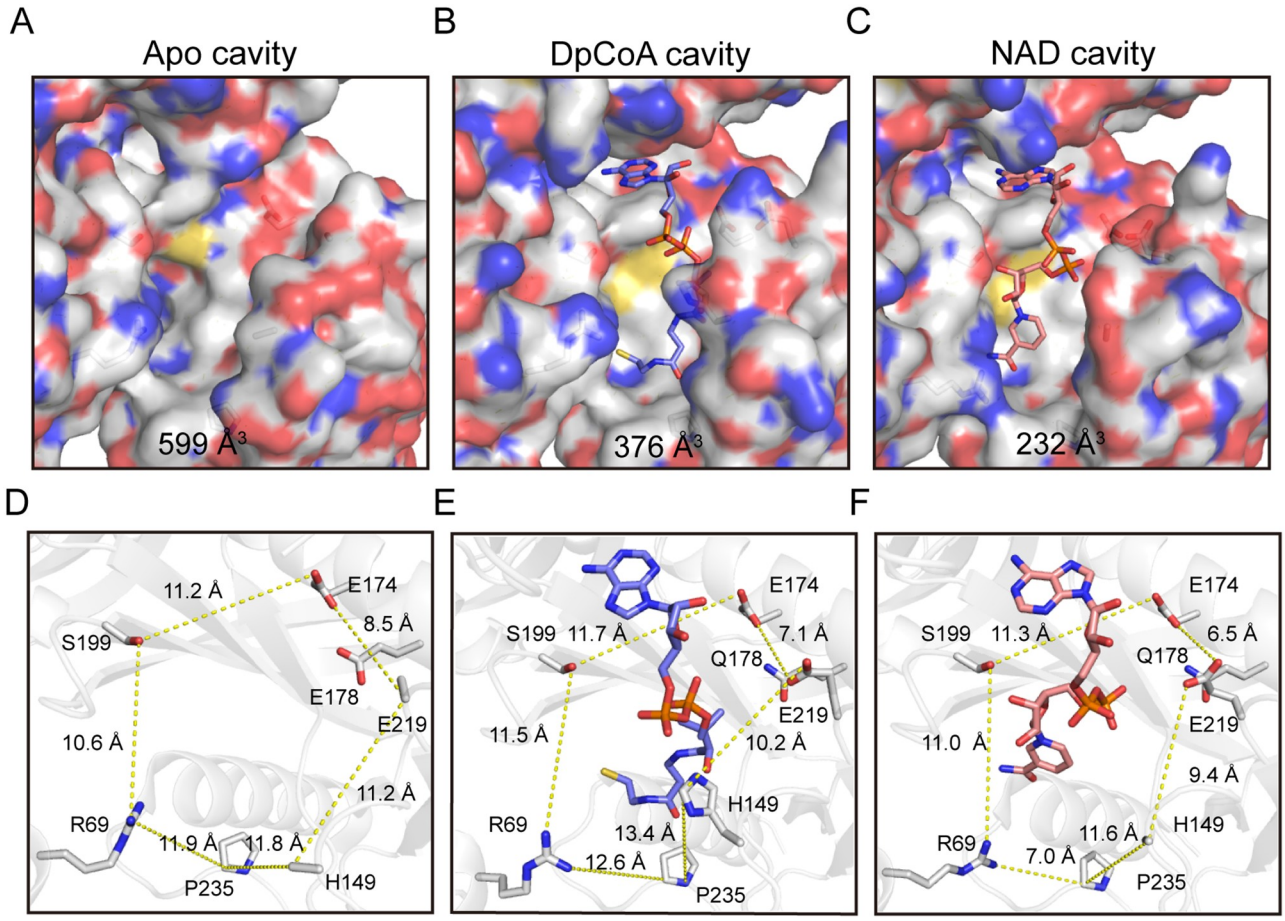


Figure 5. DpCoA and NAD cavites in the complex. (A-C) Cavity volumes of apo NudC, NudC-dpCoA and NudC-NAD. (D-F) Residues distances in the apo NudC and dpCoA/NAD-binding cavities. The Q178 residue corresponds to E178 of the wild-type protein.

adenine moiety is recognized by the two residues (Y124 and F160), the diphosphate linkage is hydrolysed, which results in liberation of the 5′ monophosphate adenosine-RNA and the remaining part (e.g. NAD, the remaining part is NMN). The proposed model needs further experimental validation in the future, as the pocket of NudC in the CTD is large enough to accommodate various substrates.

Overall, our work provides the molecular basis of dpCoA-capped RNA recognition by NudC. The architecture of the NudC-dpCoA complex also provides a structural explanation for the conservation of residues that might participate in recognition of dpCoA and dpCoA-capped RNA in organisms ranging from *E. coli* to humans. Furthermore, the ability of NudC eukaryotic homologs to cleave dpCoA-capped RNAs suggests that the mechanism of dpCoA-cap removal from the 5′ terminus of RNA is probably evolutionarily conserved.

Acknowledgments

We thank the staffs of the BL17U1/BL19U1 beamline of the National Center for Protein Sciences Shanghai (NCPSS) at the Shanghai Synchrotron Radiation Facility for assistance during data collection, and research associates at the Center for Protein Research, Huazhong Agricultural University, for technical support. We thank Dr. Zhou Gong for molecular dynamic simulation analysis, Dr. Junjie Yan and Dr. Chen Wang for critical reading of the manuscript. This work was supported by the National Natural Science Foundation of China [31770878], the open funds of the State Key Laboratory of Hybrid Rice [KF201703], the open funds of the National Key Laboratory of Crop Genetic Improvement [ZK201907], the open funds of the State Key Laboratory for Biology of Plant Diseases and Insect Pests [SKLOF201904],
the Fundamental Research Funds for the Central Universities [2662018JC023], and the Huazhong Agricultural University Scientific & Technological Self-innovation Foundation [2662017QD034].

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [31770878].

Accession number

The coordinates and the structure factors have been deposited to the Protein Data Bank with the accession code: 7E44.

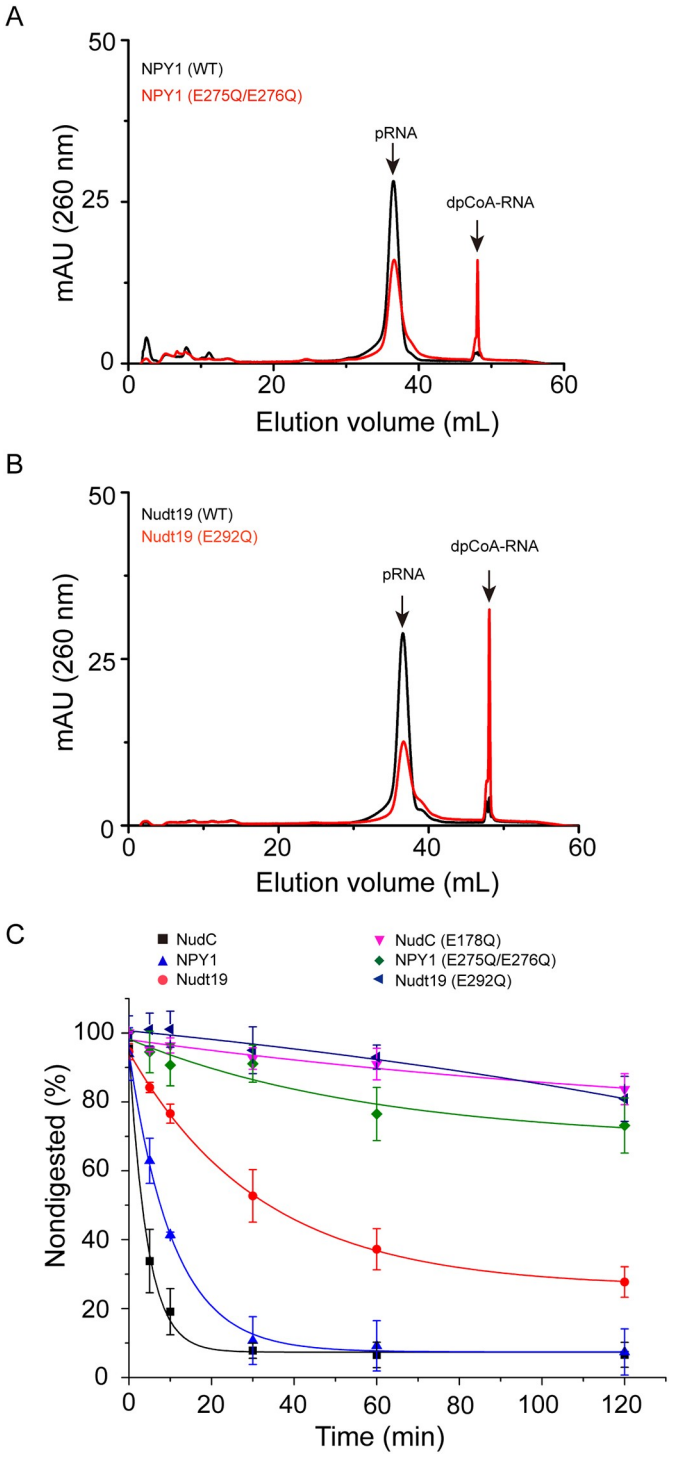


Figure 6. Cleavage of dpCoA-capped RNA by NudC homologs. (A) The NudC homolog from *Saccharomyces cerevisiae*, NPY1, exhibits hydrolytic activity towards dpCoA-capped RNA *in vitro*. The E275Q/E276Q mutant of NPY1 shows impaired RNA cleavage activity. (B) The NudC homolog from *Arabidopsis thaliana*, Nudt19, exhibits hydrolytic activity towards dpCoA-capped RNA *in vitro*. The E292Q mutant of Nudt19 shows impaired RNA cleavage activity. (C) Comparison of the rates of hydrolysis of dpCoA-RNA (50 μM) by 0.5 μM NudC, 0.5 μM NudC (E178Q), 0.5 μM NPY1, 0.5 μM NPY1 (E275Q/E276Q), 0.5 μM Nudt19 and 0.5 μM Nudt19 (E292Q). The data points represent the mean ± standard deviation (s.d.); n = 3. Cleavage was detected by ion-exchange chromatography.

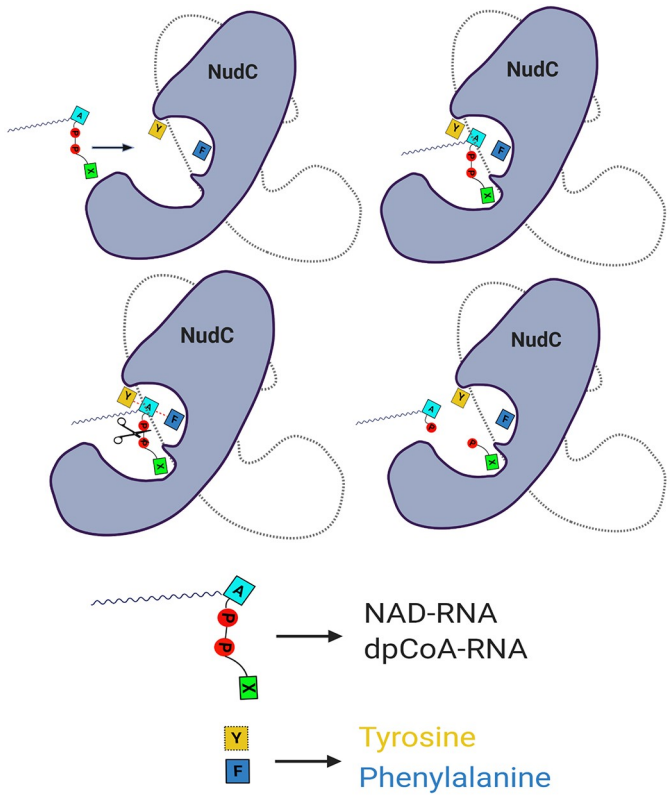


Figure 7. Proposed model for the decapping process of NudC. When the substrate RNA enters the catalytic pocket of NudC, the phenylalanine in the catalytic pocket and tyrosine in the other protomer recognize the substrate RNA via π-π interactions with the adenine base. The substrate RNA is then cleaved at the diphosphate linkage, setting free the 5′ monophosphate adenosine-RNA. X in the green square represents the nicotinamide of NAD or the sulphhydryl group of dpCoA.

ORCID

Delin Zhang http://orcid.org/0000-0002-5299-0718

References

[1] Harcourt EM, Kietrys AM, Kool ET. Chemical and structural effects of base modifications in messenger RNA. Nature. 2017;541(7637):339–346.
[2] Roundtree IA, Evans ME, Pan T, et al. Dynamic RNA modifications in gene expression regulation. Cell. 2017;169(7):1187–1200.
[3] Jäschke A, Höfer K, Nübel G, et al. Cap-like structures in bacterial RNA and epitranscriptomic modification. Curr Opin Microbiol. 2016;30:44–49.
[4] Chen YG, Kowtoniuk WE, Agarwal I, et al. LC/MS analysis of cellular RNA reveals NAD-linked RNA. Nat. Chem. Biol. 2009;5(12):879–881.
[5] Jiao X, Xiang S, Oh C, et al. Identification of a quality-control mechanism for mRNA 5′-end capping. Nature. 2010;467(7315):608–611.
[6] Ghosh A, Lima CD. Enzymology of RNA cap synthesis. Wiley Interdiscip Rev RNA. 2010;1(1):152–172.
[7] Cahová H, Winz ML, Höfer K, et al. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. Nature. 2015;519(7543):374–377.
[8] Walters RW, Matheny T, Mizoue LS, et al. Identification of NAD + capped mRNAs in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. 2017;114(3):480–485.
[9] Jiao X, Doamekpor SK, Bird JG, et al. 5′ End nicotinamide adenine dinucleotide cap in human cells promotes RNA decay through DXO-mediated deNADding. Cell. 2017;168(6):1015–1027.e1010.
[10] Wang Y, Li S, Zhao Y, et al. NAD(+)-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be

translated. Proc. Natl. Acad. Sci. U. S. A. 2019;116 (24):12094–12102.
[11] Zhang H, Zhong H, Zhang S, et al. NAD tagSeq reveals that NAD (+)-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 2019;116(24):12072–12077.
[12] Kiledjian M. Eukaryotic RNA 5′-End NAD(+) capping and DeNADding. Trends Cell Biol. 2018;28(6):454–464.
[13] Kowtoniuk WE, Shen Y, Heemstra JM, et al. A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. Proc. Nat.l Acad. Sci. U. S. A. 2009;106(19):7768–7773.
[14] Wang J, Alvin Chew BL, Lai Y, *et al.* Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. Nucleic Acids Res. 2019;47(20):e130. .
[15] Luciano DJ, Levenson-Palmer R, Belasco JG. Stresses that raise Np (4)A levels induce protective nucleoside tetraphosphate capping of bacterial RNA. Mol Cell. 2019;75(5):957–966.e958.
[16] Luciano DJ, Belasco JG. Np4A alarmones function in bacteria as precursors to RNA caps. Proc. Natl. Acad. Sci. U. S. A. 2020;117 (7):3560–3567.
[17] Dunckley T. The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. EMBO J. 1999;18(19):5411–5422.
[18] Lykke-Andersen J. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. Mol. Cell. Biol. 2002;22(23):8114–8121.
[19] Wang Z, Jiao X, Carr-Schmid A, et al. The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc. Natl. Acad. Sci. U. S. A. 2002;99(20):12663–12668.
[20] Li Y, Song M, Kiledjian M. Differential utilization of decapping enzymes in mammalian mRNA decay pathways. RNA. 2011;17 (3):419–428.
[21] Song MG, Li Y, Kiledjian M. Multiple mRNA decapping enzymes in mammalian cells. Mol Cell. 2010;40(3):423–432.
[22] McLennan AG. The Nudix hydrolase superfamily. Cell. Mol. Life Sci. 2006;63(2):123–143.
[23] Jiao X, Chang JH, Kilic T, et al. A mammalian pre-mRNA 5′ end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. Mol Cell. 2013;50(1):104–115.
[24] Doamekpor SK, Grudzien-Nogalska E, Mlynarska-Cieslak A, et al. DXO/Rai1 enzymes remove 5′-end FAD and dephospho-CoA caps on RNAs. Nucleic Acids Res. 2020;48(11):6136–6148.
[25] Sharma S, Grudzien-Nogalska E, Hamilton K, et al. Mammalian Nudix proteins cleave nucleotide metabolite caps on RNAs. Nucleic Acids Res. 2020;48(12):6788–6798.
[26] Zhang D, Liu Y, Wang Q, et al. Structural basis of prokaryotic NAD-RNA decapping by NudC. Cell Res. 2016;26(9):1062–1066.
[27] Wang Q, Zhang K-H, Cui Y, *et al.* Upgrade of macromolecular crystallography beamline BL17U1 at SSRF. Nucl. Sci. Tech. 2018;29(5):68.
[28] Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997;276:307–326.
[29] Collaborative Computational PN. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 1994;50(5):760–763.
[30] Winn MD, Ballard CC, Cowtan KD, *et al.* Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 2011;67(4):235–242.
[31] McCoy AJ, Grosse-Kunstleve RW, Adams PD, et al. Phaser crystallographic software. J Appl Crystallogr. 2007;40 (4):658–674.
[32] Adams PD, Grosse-Kunstleve RW, Hung LW, et al. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr. 2002;58 (11):1948–1954.
[33] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60 (12):2126–2132.
[34] Frick DN, Bessman MJ. Cloning, purification, and properties of a novel NADH pyrophosphatase. Evidence for a nucleotide pyrophosphatase catalytic domain in MutT-like enzymes. J. Biol. Chem. 1995;270(4):1529–1534.
[35] Höfer K, Li S, Abele F, et al. Structure and function of the bacterial decapping enzyme NudC. Nat. Chem. Biol. 2016;12 (9):730–734.
[36] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995;8(2):127–134.
[37] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 2007;372(3):774–797.
[38] Yu J, Zhou Y, Tanaka I, et al. Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. Bioinformatics. 2010;26(1):46–52.
[39] Wu H, Li L, Chen KM, et al. Decapping enzyme NUDT12 partners with BLMH for cytoplasmic surveillance of NAD-Capped RNAs. Cell Rep. 2019;29(13):4422–4434.e4413.
[40] Grudzien-Nogalska E, Wu Y, Jiao X, et al. Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. Nat. Chem. Biol. 2019;15(6):575–582.
[41] Scarsdale JN, Peculis BA, Wright HT. Crystal structures of U8 snoRNA decapping nudix hydrolase, X29, and its metal and cap complexes. Structure. 2006;14(2):331–343.
[42] Trésaugues L, Lundbäck T, Welin M, et al. Structural basis for the specificity of human NUDT16 and its regulation by inosine monophosphate. PLoS One. 2015;10(6):e0131507.
[43] Thirawatananond P, McPherson RL, Malhi J, et al. Structural analyses of NudT16–ADP-ribose complexes direct rational design of mutants with improved processing of poly(ADP-ribosyl)ated proteins. Sci. Rep. 2019;9(1):5940.