Published in final edited form as:

Wiley Interdiscip Rev RNA. 2019 March ; 10(2): e1509. doi:10.1002/wrna.1509.

Non-canonical features and modifications on the 5'-end of bacterial sRNAs and mRNAs

Nikita Vasilyev, Ang Gao, Alexander Serganov*

Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA

Abstract

While many eukaryotic transcripts contain cap structures, it has been long thought that bacterial RNAs do not carry any special modifications on their 5' ends. In bacteria, primary transcripts are produced by transcription initiated with a nucleoside triphosphate and are therefore triphosphorylated on 5' ends. Some transcripts are then processed by nucleases that yield monophosphorylated RNAs for specific cellular activities. Many primary transcripts are also converted to monophosphorylated species by removal of the terminal pyrophosphate for 5'-end-dependent degradation. Recent studies surprisingly revealed an expanded repertoire of chemical groups on 5'-ends of bacterial RNAs. In addition to mono- and triphosphorylated moieties, some mRNAs and sRNAs contain cap-like structures and diphosphates on their 5'-ends. Although incorporation and removal of these groups have become better understood in recent years, the physiological significance of these modification remain obscure. This review highlights recent studies aimed at identification and elucidation of novel modifications on the 5' ends of bacterial RNAs and discusses possible physiological applications of the modified RNAs.

Keywords

RNA decay; 5'-end-dependent degradation; cap; RNA processing

Introduction

To enrich the poor chemical repertoire of similar building blocks and add features recognized by partner molecules, many cellular RNAs undergo modifications and processing. Over a hundred modifications have been identified in various types of RNA in all three domains of life (S. Li & Mason, 2014; Marbaniang & Vogel, 2016; X. Wang & He, 2014). Most often, these modifications include chemical alterations in nucleobases and methylation of sugars (Decatur & Fournier, 2002; Motorin & Helm, 2011; Sokolowski, Klassen, Bruch, Schaffrath, & Glatt, 2018). RNA modifications have considerable influence on RNA folding and tertiary structure stabilization and are essential for key biological processes, especially for certain steps in translation, such as aminoacylation and codon-

*Corresponding author: alexander.serganov@nyumc.org.

Competing financial interests

The authors declare no competing financial interests.

anticodon recognition (Kruger, Pedersen, Hagervall, & Sorensen, 1998; Madore et al., 1999; Muramatsu et al., 1990; Wilson & Roe, 1989).

RNA processing ensures the proper conversion of pre-RNAs into mature RNAs with specific length and content. While nucleobase modifications are typically associated with ubiquitous rRNAs and tRNAs, textbook RNA processing is usually ascribed to posttranscriptional events on pre-mRNAs transcribed by RNA polymerase II in eukaryotic organisms. Physical separation of transcription and translation as well as the mosaic nature of genes add complexity to the mRNA-related activities in eukaryotic cells. Pre-mRNAs undergo splicing, a two-step process involving excision of non-coding segments (introns) and ligation of coding regions (exons) to create one continuous reading frame (Chandler, 2011). Other common changes include extensive alterations at the termini. The 3' end of the mRNA is typically extended by a poly(A) sequence and the 5' end becomes converted to a "cap" structure, an N⁷-methylated guanosine linked to the 5' RNA end through a 5'-to-5' triphosphate linkage (Fig. 1(a)) (Adhikari, Xiao, Zhao, & Yang, 2016; Katahira, 2015; Ramanathan, Robb, & Chan, 2016; Tian & Manley, 2013, 2017). In multicellular eukaryotes, formation of the cap involves several consecutive steps starting from the removal of the terminal phosphate by an RNA triphosphatase and ending by methylation of the terminal guanosine at the N⁷ position and further methylation of the first or first two sugars at the 2' position (Ghosh & Lima, 2010). Both poly(A) and cap bind regulatory proteins, inhibit degradation of mRNA, promote export from the nucleus, participate in translation, and contribute to other cellular activities (Brook & Gray, 2012; Katahira, 2015; Ramanathan et al., 2016; Topisirovic, Svitkin, Sonenberg, & Shatkin, 2011).

Bacteria do not have a nucleus, and many basic cellular processes, such as translation, differ significantly from their eukaryotic counterparts. Therefore, processing of bacterial RNAs is expected to diverge from processing of eukaryotic transcripts. Indeed, bacteria do not possess splicing machinery analogous to a spliceosome of eukaryotes and do not undergo spliceosome-mediated maturation, although some transcripts become spliced through ribozyme-driven splicing (Hausner, Hafez, & Edgell, 2014; E. R. Lee, Baker, Weinberg, Sudarsan, & Breaker, 2010). However, a 3' poly(A) sequence, initially believed to be a characteristic unique to eukaryotic mRNAs, has been identified in bacterial mRNAs (Mohanty & Kushner, 2011). This similarity suggests parallels between bacterial and eukaryotic RNA processing although the same features can serve different purposes in eukaryotes and bacteria (Belasco, 2010). For example, instead of protecting mRNA from 3'-exonucleases, bacterial poly(A) enhances 3'-to-5' degradation by adding a non-structured tail that facilitates nuclease loading on the 3' end of RNA (Mohanty & Kushner, 2011).

While splicing and polyadenylation were observed in bacteria many years ago (Gopalakrishna, Langley, & Sarkar, 1981; Kuhsel, Strickland, & Palmer, 1990; Xu, Kathe, Goodrich-Blair, Nierzwicki-Bauer, & Shub, 1990), processing of the 5' terminus was thought to exclusively involve endonucleolytic cleavage of primary transcripts to yield mature rRNAs and tRNAs with a monophosphorylated end (Condon, 2007; Deutscher, 2009). The presence or absence of the 5' cap in mRNAs (Fig. 1(b)) was considered as one of the fundamental features that distinguishes eukaryotic mRNAs translated in cap-dependent manner from the cap-independent translation initiation in bacteria, which typically requires

loading of the 30S ribosomal subunit on mRNA by complementary interactions of the 16S rRNA with the Shine-Dalgarno sequence (Shine & Dalgarno, 1974). However, several studies in the last decade demonstrated that the 5' end of bacterial RNAs, including mRNAs, can be modified as well. Most likely, these modifications affect the degradation of RNAs from the 5' end, the process which has been recently appreciated as one of the important means for RNA decay in bacteria. This review discusses the non-canonical features and modifications on the 5' end of bacterial mRNAs and sRNAs in two sections. The first section details phosphorylation states of the 5' end of bacterial RNAs and the molecular mechanisms responsible for removal of the phosphates. The second section describes non-canonical moieties on the 5' end of bacterial RNAs and the mechanism of their introduction and removal. The review highlights the early studies on the 5' end of bacterial mRNAs as well as most recent works that revealed that the 5' end of bacterial mRNAs is a much busier entity than it has been thought for many years.

PHOSPHORYLATION STATES OF THE 5' END OF BACTERIAL RNAS

Early studies of the 5' ends of bacterial mRNAs

Early identification of the chemical groups on the 5' ends of bacterial mRNA were primarily driven by studies aimed to determine directionality of transcription. These research efforts used chromatography and fingerprinting approaches to analyze the chemical composition of natural bacterial mRNAs and RNAs transcribed *in vitro* by bacterial RNA polymerases (Bremer, Konrad, Gaines, & Stent, 1965; Jorgensen, Buch, & Nierlich, 1969; Konrad, Toivonen, & Cook, 1976). These methods are based on the complete digestion of RNA by alkaline hydrolysis or nuclease cleavage, separation of the resulting fragments by one- or two-dimensional chromatography, and comparison of migration of the fragments with authentic standards. These experiments revealed that the 5' end of bacterial mRNA carries three phosphates, the 3' end does not contain a free phosphate, and adjacent nucleotides are separated by a single phosphate. Since transcription involves polymerization of ribonucleoside triphosphates, these results were consistent with the RNA chain being initiated from the nucleoside triphosphate.

Although alkaline hydrolysis of bacterial mRNAs predominantly identified pppGp and pppAp moieties on the 5' ends, it also revealed nucleoside diphosphates and triphosphates, which are likely equivalent to pNp and ppNp, the 5' mono- and diphosphorylated nucleosides bearing a 2'-3' cyclophosphate (Jorgensen et al., 1969). The presence of these 5' mono- and diphosphorylated species was noted but not discussed possibly because experimental procedures could have caused hydrolysis of the triphosphate moiety (Konrad et al., 1976).

The 5' ends of bacterial mRNAs can be monophosphorylated

Monophosphate, as a 5' end modification of bacterial RNAs, was identified in studies of mRNA degradation in *E. coli*. This bacterium contains several endoribonucleases and 3' exoribonucleases but lacks a dedicated enzyme with 5' exoribonucleolytic activity. Many *E. coli* mRNAs are produced with a stem-loop at the 3' end that protects RNA from degradation by 3' exonucleases (Fig. 2(a)). Therefore, degradation of *E. coli* mRNAs must

begin with endonucleolytic cleavage at the internal site to produce two fragments (Apirion, 1973; Belasco & Higgins, 1988). The 5' fragment lacking protective stem-loop structure at the 3' end can be degraded by 3' exonucleases. The 3' fragment must undergo additional rounds of endonucleolytic and exonucleolytic cleavages. The most important endoribonuclease that makes initial cleavage is single-strand-specific RNase E (Apirion, 1978; Babitzke & Kushner, 1991; Melefors & von Gabain, 1991; Mudd, Krisch, & Higgins, 1990; Ono & Kuwano, 1979; Taraseviciene, Miczak, & Apirion, 1991). Interestingly, the rate of RNase E cleavage was found to depend on the properties of the 5' end (Arnold, Yu, & Belasco, 1998; Bouvet & Belasco, 1992; Emory, Bouvet, & Belasco, 1992). It turned out that RNase E strongly prefers unpaired 5' ends bearing a single phosphate (Mackie, 1998). How this feature of the enzyme can be accommodated in the degradation pathway if primary transcripts are produced with triphosphorylated 5' ends, aside of some RNAs matured by endonucleolytic cleavage of the primary transcripts?

Subsequent studies have identified the enzyme capable of converting 5'-triphosphorylated RNAs into monophosphorylated species. This enzyme, termed RNA pyrophosphohydrolase or RppH, cleaves a pyrophosphate off the triphosphorylated 5' end *in vitro* and produces RNA substrates for RNase E cleavage (Celesnik, Deana, & Belasco, 2007; Deana, Celesnik, & Belasco, 2008) (Fig. 2(b)). The substantial abundance of monophosphorylated mRNAs was demonstrated *in vivo* by a newly developed Phosphorylation Assay by Ligation of Oligonucleotides (PABLO), the method that uses T4 DNA ligase to link adjacent ends of nucleic acids bearing 3' hydroxyl and 5' monophosphate (Deana et al., 2008). Thus, the 5'-monophosphorylated end generated by *E. coli* RppH can be considered as the first posttranscriptional modification of the primary transcripts in bacteria. This modification was later described for *B. subtilis* (Richards et al., 2011), the bacterium which does not possess RNase E, typically found in Gram-negative species, but has a specialized 5' exoribonuclease with preference for monophosphorylated 5'-end (Even et al., 2005; Mathy et al., 2007). Therefore, the 5'-end-dependent RNA degradation pathway triggered by RppH is likely ubiquitous in bacteria.

Aside from involvement into RNA degradation, the phosphorylation state of the 5' end of mRNA may define efficiency of other cellular processes. A recent bioinformatics study showed that many bacterial genomes contain leaderless genes lacking a 5' untranslated region and Shine-Dalgarno ribosome-binding site (Zheng, Hu, She, & Zhu, 2011). In some bacteria, for example, *Mycobacterium* species, nearly one-quarter of transcripts are leaderless and leadered and leaderless translations are comparably robust (Shell et al., 2015). While proximity of the start codon to the 5' terminus of the leaderless mRNA appears to be a strong positive determinant (Krishnan, Van Etten, & Janssen, 2010), the 5'-terminal phosphate has been shown to be an important determinant of the translation initiation step as well (Giliberti, O'Donnell, Etten, & Janssen, 2012). Removal of the 5' phosphates virtually abolishes translation of the leaderless *cd* mRNA in *E. coli* and significantly reduces ribosome binding to the 5'-hydroxylated mRNA *in vitro*.

The difference in the 5'-end phosphorylation of cellular RNAs formed the basis for pioneering global identification of transcription start sites and RNA processing events (Hor, Gorski, & Vogel, 2018; Wurtzel et al., 2012). The primary transcripts are generated with the

triphosphate on the 5'-end while the RNAs processed by nuclease cleavage as well as degradation intermediates contain monophosphorylated or hydroxylated RNAs, which can be, for example, degraded by the Terminator exonuclease (Sharma et al., 2010). Comparison of the libraries prepared with the exonuclease-treated and untreated RNA samples can distinguish the primary transcripts from the processed ones. Similarly, profiling the RNA ends obtained with and without activity of a particular processing enzyme could identify relevant processing sites with a single-nucleotide resolution (Chao et al., 2017; Le Rhun et al., 2017).

A diphosphorylated 5' end is another modification in *E. coli* mRNAs

Although RppH-dependent conversion of triphosphorylated mRNAs to monophosphorylated species seemed to explain the initial events in the 5'-end-dependent RNA degradation pathway in *E. coli*, a couple of puzzling observations prompted a more detailed investigation of the pathway. First, deletion of the chromosomal copy of *rppH* gene affected the degradation of hundreds but not all mRNAs (Luciano et al., 2012). Second, overexpression of RppH did not significantly affect mRNA degradation (Luciano et al., 2012). These results were partially clarified by genetic and biochemical experiments that demonstrated specific requirements of RNA substrates for efficient RppH catalysis (Foley, Hsieh, Luciano, & Belasco, 2015). It turned out that the enzyme requires at least two and prefers three or more unpaired nucleotides at the 5' RNA end. RppH is also more active on the RNAs that have purine in second position and prefers guanine over adenine there.

To understand the substrate specificity of *E. coli* RppH, its three-dimensional structure has been determined in the *apo* and RNA-bound conformations (Bi, Li, Fan, Xia, & Jin, 2009; Vasilyev & Serganov, 2015). The structures revealed a bipartite mode of RNA recognition by RppH (Fig. 3(a)) (Vasilyev & Serganov, 2015). The enzyme binds the triphosphorylated RNA end using direct and Mg^{2+} -mediated interactions (Fig. 3(b)). The enzyme also binds the second nucleotide of the RNA and the backbone of adjacent nucleotides in a cleft formed on the surface of the protein. The cleft contains amino acids that make cation- π interactions (R27) and semi-specific hydrogen bonds (K140 and S32) with the nucleobase of guanosine in the second position of the RNA. These contacts explain the preference of RppH for purines at this position (Foley et al., 2015).

The structure revealed multiple interactions between RppH and the α and β phosphates at the RNA 5' end that ensure correct positioning of the RNA substrate in the active site (Vasilyev & Serganov, 2015) (Fig. 3(c)). The structure also identified a water molecule likely activated by a glutamate and two adjacent Mg^{2+} cations for in-line attack on the β phosphorus resulting in cleavage of the bond between α and β phosphates. Interestingly, the structure did not reveal extensive interactions with the γ phosphate and biochemical experiments showed the principal ability of RppH to remove a phosphate from the diphosphorylated RNA, thereby raising doubts on importance of the γ phosphate for catalysis. The latter observation together with the limited effect of the RppH overproduction on the rate of mRNA decay in cells suggested existence of an additional rate-limiting step during conversion of triphosphorylated RNA to monophosphorylated species (Luciano et al., 2012) and prompted further investigation of RppH specificity.

The next study explained some of these observations. This work confirmed that RppH can function on diphosphorylated RNAs *in vitro* (Luciano, Vasilyev, Richards, Serganov, & Belasco, 2017) and demonstrated that RppH activity was more than tenfold greater on diphosphorylated RNAs compared to their triphosphorylated counterparts. These data suggested that the diphosphorylated RNAs are the natural substrates of RppH in *E. coli*. To show that diphosphorylated RNAs are present in cells and are targeted by RppH, the Joel Belasco laboratory developed a new method termed PACO (Phosphorylation Assay by Capping Outcome) that allowed them to estimate the fraction of diphosphorylated mRNA species in the total RNA extracted from cells (Luciano et al., 2017). This new method takes advantage of yeast mRNA guanylyltransferase which uses diphosphorylated RNA as a substrate for capping. A combination of PACO and PABLO (Celesnik, Deana, & Belasco, 2008) made it possible to precisely determine the phosphorylation state for representative *E. coli* mRNAs (Luciano et al., 2017). Surprisingly, the triphosphorylated primary transcripts were found to be much less abundant than their di- and monophosphorylated counterparts, while *rppH* knock out resulted in further increase of fraction of diphosphorylated RNAs. Overall, the study concluded that RppH indeed targets diphosphorylated RNAs and appears to be the only enzyme converting diphosphorylated RNAs to monophosphorylated products in cells, whereas preceding step of the γ phosphate removal is attributed to yet unidentified enzyme(s). The diphosphate at the RNAs 5' end thus can be considered as a new posttranscriptional modification of mRNAs in *E. coli* (Fig. 2(c)). Retrospectively, the 5'-diphosphorylated RNA digestion products reported in earlier studies (Bremer et al., 1965; Jorgensen et al., 1969; Konrad et al., 1976) might have originated from genuine diphosphorylated RNA rather than side reactions during the purification procedure.

Is the diphosphorylated state of mRNA common and important?

How common are diphosphorylated 5' ends in bacterial RNA? So far, the abundance of 5' diphosphorylated species has been shown experimentally only for two *E. coli* mRNAs degraded through the RppH-mediated pathway (Luciano et al., 2017). One would expect that at least the RNAs targeted for decay by RppH may be initially converted to diphosphorylated species; therefore, this modification is anticipated to be common for hundreds of bacterial RNAs. Since recognizable homologs of *E. coli* RppH are widespread among species representing all classes of Proteobacteria, except Deltaproteobacteria, the RppH-mediated 5'-end-dependent degradation pathway is likely to be common for many bacteria (Foley et al., 2015). *E. coli* RppH and its homologs show high degree of similarity for residues involved in RNA binding and catalysis, suggesting similar substrate requirements across the *E. coli* type RppHs (Foley et al., 2015; Vasilyev & Serganov, 2015). Thus, the presence of another enzyme(s) that removes the γ phosphate from hundreds of RNAs seems likely for a large number of bacterial species.

Can stable diphosphorylated RNAs be present in bacterial species that do not contain the *E. coli* type RppH? *Bacillus subtilis* RppH has low sequence identity with *E. coli* RppH, and despite sharing similar structural folds, the two enzymes have divergent substrate specificities and catalytic mechanisms (Foley et al., 2015). In contrast to *E. coli* RppH, *B. subtilis* RppH strictly requires guanosine at the second position of its substrate RNAs (Hsieh, Richards, Liu, & Belasco, 2013) and the enzyme forms a tight pocket for specific

recognition of the nucleobase of this guanosine (Piton et al., 2013). Unlike *E. coli* RppH, which cleaves off pyrophosphate from triphosphorylated RNA in a single step *in vitro*, *B. subtilis* RppH removes pyrophosphate in two sequential orthophosphate cleavage steps (Richards et al., 2011). The *B. subtilis* enzyme does not appear to have a preference for diphosphorylated species (our unpublished data), and the diphosphorylated intermediate generated by the first cleavage likely undergoes the second cleavage by repositioning the substrate in the active site without release of the intermediate. Therefore, *B. subtilis* RppH may not need an additional enzyme for initial processing of 5' ends and in principle can itself convert triphosphorylated RNA ends into monophosphorylated species for degradation. *B. subtilis* appears to contain another yet unidentified RNA pyrophosphohydrolase (Hsieh et al., 2013). Whether diphosphorylated intermediates are present in *B. subtilis* and required by the second RppH remains to be investigated.

The ability of RppH to convert triphosphorylated substrates directly to monophosphorylated products raises the question as to why this processing evolved in *E. coli* as a two-step process involving multiple enzymes and a diphosphorylated intermediate rather than a simple one-step process catalyzed solely by RppH (Luciano, Vasilyev, Richards, Serganov, & Belasco, 2018). Indeed, conversion of triphosphorylated RNAs to monophosphorylated species could just as easily have been accelerated by increasing the cellular concentrations of RppH and its stimulating factor DapF (see below). A first possibility is that some cellular enzymes, such as NTPases or kinases, might not be able to distinguish mononucleotide from polynucleotide substrates thereby generating diphosphorylated RNA byproducts. If so, *E. coli* may have simply adapted an RNA decay pathway to target diphosphorylated RNAs for degradation. A second possibility is that bacteria have evolved an additional point to control RNA degradation rate by modulating the activity or concentration of the enzyme(s) responsible for γ phosphate removal. A third possibility is that the diphosphorylated RNA ends may govern additional cellular processes, perhaps by enabling binding by proteins other than RppH or by preventing RNA ends from being mistaken for nucleoside triphosphates. For example, reduced number of 5' phosphates can increase translation efficiency of leaderless mRNAs as it was shown that the ribosome binds stronger to the 5'-monophosphorylated RNA than to the 5'-triphosphorylated RNA (Giliberti et al., 2012).

A metabolic enzyme modulates RppH activity

A recent study identified the first cellular factor that accelerates the rate of RNA decay through stimulating the activity of *E. coli* RppH (C. R. Lee, Kim, Park, Kim, & Seok, 2014). This factor is the metabolic enzyme diaminopimelate epimerase, DapF, that catalyzes the stereoconversion of an intermediate in lysine and peptidoglycan biosynthesis (Mengin-Lecreulx, Michaud, Richaud, Blanot, & van Heijenoort, 1988; Richaud, Higgins, Mengin-Lecreulx, & Stragier, 1987; Richaud & Printz, 1988). DapF forms a tight complex with RppH and does not require catalytic activity to stimulate RppH (C. R. Lee et al., 2014).

Biochemical and structural studies discovered a 2:2 stoichiometry of the complex, consistent with dimerization of an RppH-DapF heterodimer through DapF-DapF interactions (Fig. 3(d)) (Gao et al., 2018; C. R. Lee et al., 2014; Q. Wang et al., 2018). The structures showed that the RppH-DapF interface is located away from both the catalytic site and RNA-binding

cleft of RppH, therefore, DapF cannot directly participate in 5' end binding and catalysis (Gao et al., 2018; Q. Wang et al., 2018). However, the structures of the individual enzymes and RppH-DapF complex in the *apo* and RNA-binding states did not reveal pronounced allosteric changes in RppH that would explain its higher reactivity upon DapF binding (Gao et al., 2018). Allosteric modulation was expected at least for short triphosphorylated substrates of RppH, which cannot reach DapF in the complex but which are processed by DapF-bound RppH at greater rate than by RppH alone (Gao et al., 2018; C. R. Lee et al., 2014). Remarkably, RppH reactivity on short diphosphorylated RNAs was not affected while a rate acceleration was observed on substrates ~8 nucleotides long, a length sufficient to reach DapF in the complex (Gao et al., 2018). Likewise, RppH showed additional DapF-mediated acceleration of reactivity on long triphosphorylated RNAs. Cumulatively, acceleration appeared to be greater for triphosphorylated RNAs than for diphosphorylated substrates, but the catalytic rate for triphosphorylated RNAs did not reach the levels observed for diphosphorylated species. Although the molecular mechanisms of DapF-mediated RppH stimulation are not understood, DapF obviously exploits two different means for potentiating RppH activity, depending on the nature of the substrate. While the enhanced reactivity of triphosphorylated RNAs may involve DapF-induced changes in the structure and/or dynamics of RppH, the stimulatory effects on long di- and triphosphorylated RNAs may be associated with either additional interactions of the substrate RNA with DapF or accelerated product release caused by reducing the binding affinity for RppH through shielding part of the RNA-binding surface of RppH.

Although DapF stimulates the decay of representative transcripts whose degradation is triggered by RppH (Gao et al., 2018; C. R. Lee et al., 2014) (Fig. 2(c)), it remains unclear why a metabolic enzyme affects RNA degradation and how such modulation would benefit bacteria. One could hypothesize that DapF contributes to relieving stresses that affect the integrity of the cell wall, such as osmotic stress, or impact cellular metabolism, such as starvation for amino acids (C. R. Lee et al., 2014). It also remains unclear why the reactivity of RppH is modulated differently on tri- and diphosphorylated RNAs and whether such selectivity of RppH stimulation can change the balance of tri- and diphosphorylated RNAs in cells. Regardless of the number and identity of transcripts affected by DapF, the influence of this protein on RNA decay offers an exciting opportunity to acquire a deeper understanding of the intricate relationships between metabolism and RNA degradation and elucidate the impact of RNA modification on RNA stability.

NON-CANONICAL 5' ENDS OF BACTERIAL RNAS

5' terminus of some bacterial RNAs contains coenzyme A

The disparity between limited chemical diversity of RNA nucleotides and plethora of natural RNA functions, coupled with the powerful functional properties of synthetic small molecule-nucleic acid conjugates (Gartner & Liu, 2001; Gartner et al., 2004; Kanan, Rozenman, Sakurai, Snyder, & Liu, 2004; X. Li & Liu, 2004), prompted searches for novel small molecule-RNA conjugates that may exist in modern organisms as evolutionary fossils or even as novel RNAs with functions enabled by their modifications (Chen, Kowtoniuk, Agarwal, Shen, & Liu, 2009; Kowtoniuk, Shen, Heemstra, Agarwal, & Liu, 2009). To

discover new natural small molecule–RNA conjugates, two studies from the David Lui laboratory applied a broad approach that used size-exclusion chromatography and mass spectrometry to detect small molecules covalently attached to cellular RNAs (Chen et al., 2009; Kowtoniuk et al., 2009).

The first study focused on detecting and characterizing base- or nucleophile-labile small molecules that are cleaved from cellular RNA by a treatment with base or a nucleophile (Kowtoniuk et al., 2009). To identify small molecule–RNA conjugates, total cellular RNA of 2,500 Da in size from two distant bacterial species, *Escherichia coli* and *Streptomyces venezuelae*, was treated with ammonium bicarbonate or n-butylamine to cleave base-labile and nucleophile-labile small molecules, respectively, and the liberated fraction 2,500 Da in size was analyzed by high-resolution liquid chromatography and tandem mass spectrometry (LC-MS/MS), isotope labeling, and comparison with authentic standards to elucidate the structures of small molecules. These methods revealed a number of new putative small molecule–RNA conjugates.

To directly identify nucleotides conjugated to base-labile or nucleophile-labile small molecules, the macromolecular fraction of whole cellular RNA was treated with nuclease P1, an endonuclease that cleaves RNA to generate mononucleotides with a 3′-hydroxyl group and a 5′-phosphate (Kowtoniuk et al., 2009). The digest was then treated with a base or a nucleophile. LC-MS analysis of samples prepared from bacteria grown in normal or isotopically-labelled media revealed over a dozen unknown modified species and found that several of them correspond to 3′-dephospho-coenzyme A (dpCoA) and its succinyl-, acetyl-, and methylmalonyl-thioester derivatives (Fig. 4(a)). These modifications are positioned on the 5′ ends of RNAs and present in ~80–120 RNA molecules per *E. coli* or *S. venezuelae* cell.

Intriguingly, the dpCoA-linked RNA(s) are not widely distributed in their size but instead are observed in RNAs <200 nt in length (Kowtoniuk et al., 2009). Since the identity of the dpCoA-linked RNA(s) have not been determined yet, function(s) of the CoA modifications are not known.

NAD as a cap-like structure for bacterial RNAs

The second study aimed to develop a more general method to detect small molecule–RNA conjugates that can be applied to any such conjugate regardless of its chemical reactivity (Chen et al., 2009). Low-molecular weight contaminations were removed from the cellular RNA sample by chromatography and the macromolecular fraction (>2,500 Da) was divided into two halves. The first half was treated with nuclease P1, and the second half was treated with heat-inactivated nuclease P1, followed by size-exclusion chromatography and comparative high-resolution LC-MS analysis of the small-molecule fraction. Application of this method to *E. coli* or *S. venezuelae* RNA revealed 24 or 48 non-canonical species that were enriched in the sample.

Follow-up experiments showed that one of the highly enriched species corresponds to nicotinamide adenine dinucleotide (NAD) linked to the 5′ end of RNA (Fig. 4(a)) (Chen et al., 2009). The NAD-RNAs are much more abundant than CoA-RNA and are estimated at

~3,000 copies per *E. coli* cell. As in the case of CoA, NAD-RNAs are predominantly less than ~200 nt in length.

To identify NAD-modified RNAs, the Andres Jäschke laboratory developed a clever method to capture NAD-linked RNAs and determine their sequences by next-generation sequencing (Cahova, Winz, Hofer, Nubel, & Jäschke, 2015). The method is based on the ability of adenosine diphosphate-ribosylcyclase from *Aplysia californica* to replace nicotinamide with alkynyl alcohols. This reaction creates a 'clickable' transglycosylation product that was attached to biotin and captured on streptavidin beads. Therefore, all RNAs that are linked to NAD and present in sufficient quantity can be captured on beads and subsequently identified. Sequencing of the captured RNAs from *E. coli* revealed many sRNAs reported to act in different pathways, for example, GadY, GcvB, McaS, CopA and RNAI. The latter two were observed in the plasmid-bearing strains because they are encoded on plasmids and control plasmid replication. Other NAD-linked species were short 5'-terminal fragments of ~60 different mRNAs, including *gatY*, *pgk*, *hdeD*, *ilvL* and *hisL*. These mRNAs encode metabolic enzymes and leader peptides. Other enzymes and several enriched sRNAs act in stress responses induced by, for example, exposure to acid and ultraviolet light.

Aberrant transcription initiation installs 5' caps on bacterial RNAs

Given that dpCoA, NAD, and NADH contain adenosine diphosphate moieties, these cofactors can in principle be used by *E. coli* RNA polymerase (*Eco* RNAP) for aberrant transcription initiation instead of ATP. A few decades ago, a study hinted at the ability of *E. coli* RNAP to start RNA synthesis with the cofactors NAD, NADH and flavin adenine dinucleotide (FAD) (Malygin & Shemyakin, 1979). Subsequent *in vitro* experiments questioned co-transcriptional incorporation of dpCoA and NAD(H) into RNA, possibly because ATP outcompeted cofactors during transcription initiation (Chen et al., 2009; Kowtoniuk et al., 2009). Recent biochemical and structural studies clarified this issue and demonstrated that NAD, NADH and dpCoA are incorporated into RNA during transcription initiation by serving as non-canonical initiating nucleotides (NCINs) for *de novo* transcription initiation by both *E. coli* and eukaryotic RNAPs *in vitro* and *in vivo* (Bird et al., 2016; Julius & Yuzenkova, 2017). The efficiency of transcription initiation with NCINs is however lower than with ATP (Julius & Yuzenkova, 2017). Because of the ~4, 7, and 120 fold lower intracellular concentration of NAD, dpCoA and NADH than ATP (Bennett et al., 2009) in *E. coli*, the low abundance of NCIN-RNAs (Chen et al., 2009; Kowtoniuk et al., 2009) is not surprising. NCIN capping requires +1A promoters, and its efficiency depends on the promoter DNA sequence at and upstream of the transcription start site (TSS) and varies by as much as two orders of magnitude (Bird et al., 2016). Specifically, position -1 was reported as a key determinant of NAD-mediated initiation, possibly because of additional interactions formed by NAD with the template DNA (Bird et al., 2016). This conclusion was recently challenged by a kinetic study that revealed similar changes in the efficiency of the initiation step for NAD and ATP on templates with different -1 nucleotides (Julius & Yuzenkova, 2017). Therefore, the -1 position of the promoter may affect the properties of other steps(s) of initiation and consequently affect the efficiency of RNA NADylation (Julius & Yuzenkova, 2017). The specific composition of the promoter could restrict the number of promoters for efficient NAD-capping. Indeed, almost half of the *E.*

coli promoters are less likely to use NAD for initiation because they contain cytosine at the –1 position (Kim et al., 2012), which was found to be unfavorable for NADylation (Bird et al., 2016). The need for a specific composition of a promoter for efficient incorporation of NCINs has been recently corroborated by the CapZyme-seq technique (Vvedenskaya et al., 2018). Using synthetic library of DNA promoters, the study identified the consensus TSS and the sequence HRRASWW (TSS underlined; H: A,C, or T; R: A or G; S: G or C; W: A or T) around it, which favor incorporation of NAD⁺ as a 5'-terminal nucleotide into nascent RNA. When applied to the RNA isolated from bacterial culture, CapZyme-seq mostly confirmed *in vitro* results and revealed that incorporation of NCINs depends on positions –3 to +2. The strongest dependence of capping efficiency on nucleotide identity was observed at TSS because of pairing with the adenosine moiety of adenosine-containing NCIN, and –1 position, which displayed a clear bias towards A or G possibly because of favorable stacking interactions with NCIN moiety.

Kinetic studies showed that NCINs increase the efficiency of promoter escape, i.e. *Eco* RNAP extends the NCIN-capped 2-nt RNAs to a 9-nt length better than the ATP-capped RNAs (Julius & Yuzenkova, 2017). These data support the idea of specific interactions between NCINs and RNAP during transcription initiation. To determine the molecular mechanism of transcription initiation starting with an NCIN, Bird *et al.* solved the X-ray crystal structures of the *Thermus thermophilus* RNAP (*Tth* RNAP) open promoter complex with a dinucleotide bearing either NAD or dpCoA at the 5' end (Bird et al., 2016). The structures showed that RNAP has sufficient space to accommodate a bulky NMN moiety in the so-called rifampicin-binding pocket of the β -subunit (Fig. 5(a)). This region is similar in *Tth* and *Eco* RNAPs; therefore, the *Tth* RNAP structures can provide structural insights into biochemical results obtained with *Eco* RNAP.

The structure with the dpCoA cap revealed a disordered pantetheine tail of dpCoA and no specific recognition by RNAP (Bird et al., 2016). In contrast, the NAD-containing dinucleotide was well defined, although its conformation also did not reveal base-specific recognition by RNAP (Fig. 5(b)). The only direct hydrogen bond observed was between *Tth* RNAP β Thr398 (equivalent to *Eco* RNAP β Asn518) and the 2' hydroxyl of the nicotinamide riboside. The nucleobase of NAD forms a water-mediated contact with the main chain of β Phe394 (*Eco* β Phe514) and van der Waals interactions with β Asp396 (*Eco* β Asp516). The latter contact appears to be most important for the ability of RNAP to initiate transcription with NAD since mutation of the *Eco* β Asp516 to valine or tyrosine reduces NAD incorporation efficiency $\approx$ 10-fold compared to that of ATP (Julius & Yuzenkova, 2017). Given the dynamic nature of transcription initiation, other contacts may form and disappear during RNA chain growth, as suggested by mutations of amino acids constituting rifampicin-binding pocket. These mutations showed a diminished ability of mutant RNAPs to start transcription with NAD even if the mutated amino acids are not located in the vicinity of NAD (Julius & Yuzenkova, 2017).

Can transcription initiation further diversify 5' RNA ends?

Although no cap-like structures in addition to NAD(H), dpCoA and its derivatives have been identified in bacterial sRNAs and mRNAs, a large number of new uncharacterized

modifications obtained in the MS studies (Kowtoniuk et al., 2009) suggest that methodological advancements will likely expand the repertoire of 5' end modifications. What could such modifications be? The primary candidates are small cellular molecules that contain nucleotide moieties which can be used by RNA polymerase for transcription initiation.

Following a pioneer work on incorporating FAD into RNA (Malygin & Shemyakin, 1979), a recent study demonstrated efficient promoter-dependent transcription initiation with FAD, which was only ~2 and 1.5 fold lower than incorporation of NAD (Julius & Yuzenkova, 2017). These data suggest that FAD-capped RNAs may exist in cells (Fig. 4(b)). However, the ~60 fold lower intracellular concentration of FAD in comparison to ATP (Bennett et al., 2009) suggests that FAD-capped RNAs, if they exist, are not abundant and can easily escape detection.

In contrast to FAD, concentration of the nucleotide analog UDP-N-acetylglucosamine (UDP-GlcNAc) (Fig. 4(b)), which participates in making cell wall polymers, is second only to ATP in *E. coli* cells and exceeds concentrations of other NTPs (Bennett et al., 2009). *E. coli* RNAP efficiently initiates transcription with UDP-GlcNAc and related precursor of cell wall biosynthesis, uridine 5'-diphosphoglucose (UDP-Glc) (Fig. 4(b)), at the rate comparable with UTP and with K_m values only 2-fold higher than that of UTP and >10-fold lower than their intracellular concentrations (Bennett et al., 2009). Since the UDP-GlcNAc- and UDP-Glc-capped initial transcripts elongate with higher efficiency than the transcripts initiated with UTP, UDP-Glc and UDP-GlcNAc may serve as RNA caps *in vivo* (Julius & Yuzenkova, 2017). However, the number of transcripts initiated with U (~6%) is much lower than the number starting with A or G (79% together) (Kim et al., 2012) in *E. coli* cells; therefore, UDP-Glc- and UDP-GlcNAc-capped RNAs, if they exist in cells, remain undetected.

Interestingly, UDP-MurNAc pentapeptide (Fig. 4(b)), a more complex compound that contains a pentapeptide attached to the GlcNAc moiety, does not support transcription (Julius & Yuzenkova, 2017). Since this compound is the precursor of the last step before the formation of a building block of the cell wall, its incorporation into RNA may lead to undesired targeting of the capped RNA to the membranes, where it may interfere with cell wall biosynthesis. RNAP appear to possess a mechanism that guards against incorporation of cellular nucleotide analogues with long side chains, possibly through clashing with region 3.2 of $\sigma 70$ subunit (Julius & Yuzenkova, 2017).

RNAP also poorly incorporates NADP, a derivative of NAD with a phosphate group at the 2' position of adenosine (Julius & Yuzenkova, 2017). This extra phosphate moiety likely interferes with phosphodiester bond formation by sterical hindrance. Since intracellular concentrations of NADPH and NADP are >80 fold lower than ATP (Bennett et al., 2009), it is highly unlikely that cellular RNAs contain NADP caps.

NCIN caps are removed by NudC

The NCIN caps resist the activity of RppH, therefore RppH cannot convert capped *E. coli* RNAs to the 5'-monophosphorylated species for subsequent efficient degradation by RNase

E (Bird et al., 2016; Cahova et al., 2015). Instead of RppH, the NAD cap was found to be removed by another Nudix hydrolase NudC, which hydrolyses NAD(H) into nicotinamide mononucleotide (NMN) and adenosine 5'-monophosphate (AMP). NudC is a genuine RNA-binding protein originally reported to co-purify with nucleic acids (Frick & Bessman, 1995). NudC has higher activity on the NAD-RNA substrates than on NAD or NADH and requires at least three non-paired nucleotides on the 5'-end and a purine at the first position of RNA for efficient catalysis (Hofer et al., 2016). NudC can also remove dpCoA caps *in vitro*, suggesting that NudC does not restrict its activity to the NAD caps (Bird et al., 2016).

The three-dimensional structures of *E. coli* NudC were determined in the *apo* form (PDB 1VK6 and 2GB5) and bound to the substrate mimic NAD (Hofer et al., 2016; Zhang et al., 2016) and the product NMN (Hofer et al., 2016). NudC consists of two domains and forms a tight dimer both in solution and crystals (Fig. 6(a)). NAD binds in the preorganized pocket predominately located in the C-terminal domain (CTD). The nicotinamide moiety (NA) of NAD is trapped in a cavity mainly comprised of bulky hydrophobic residues and is specifically recognized by hydrogen bonding (Fig. 6(b)). The adenosine moiety is located in a separate cavity formed at the interface of the dimer. The nucleobase is sandwiched between aromatic residues that belong to both protomers thereby suggesting essential contribution of dimerization to the RNA recognition. Although two structural studies have placed adenine in the same position, they disagreed on the glycosidic bond angle and reported anti- and syn- orientation of the nucleobase as well as slightly different conformation of the sugars (Fig. 6(b)) (Hofer et al., 2016; Zhang et al., 2016). Nevertheless, phosphates of NAD were located in the same places near several glutamates of the Nudix helix involved in the cleavage of the interphosphate linkage in other Nudix hydrolases. Despite involvement of metal cations in the activation of a water molecule for catalysis in other Nudix enzymes, neither NudC structure identified metal cations in the catalytic site. Superposition with the RNA- and Mg²⁺-bound RppH structure showed that while the phosphate of the NMN moiety coincides with the α -phosphate of the 5' RNA end bound to RppH, the adenosine-linked phosphate of NAD is located away from the catalytically-relevant position occupied by the β -phosphate of the RppH-bound RNA (Fig. 6(c)). Given conservation of the catalytic residues of the Nudix motif in RppH and NudC, the structure of the NAD-bound NudC appears to capture the pre-catalytic state of the enzyme, awaiting cation binding and re-location of the adenosine-linked phosphate into the catalytically relevant site. Interestingly, the conformation and position of NMN in the NMN-bound NudC structure (Hofer et al., 2016) differ from ones of the nicotinamide moiety of NAD (Fig. 6(d)). These differences indicate that the catalytic step and product release likely involve conformational adjustments in the enzyme.

Surface representation of the NAD-bound NudC structure shows that the NA-binding cavity does not form a tight pocket for the NA moiety and leaves empty space in the bottom. Therefore, the cavity has potential to accommodate other moieties, including dpCoA (Fig. 6(e)), which is less bulky than NAD and can be removed by NudC (Bird et al., 2016).

What is the role of bacterial caps?

Decapping by NudC yields 5'-monophosphorylated RNAs, which are optimal substrates for the 5'-end dependent degradation by RNase E. Therefore, the most appealing hypothesis that explains the existence of NAD and probably other caps is a selective mechanism for initiating the degradation of a subset of cellular RNAs that is distinct from the RNA pool directly targeted for degradation through the RppH-dependent pathway (Fig. 2(c)) (McLennan, 2006; Mildvan et al., 2005). Of some concern is the observation that only a fraction of the full-length RNAI contains NAD cap and its effect on RNAI degradation *in vivo* is negligible (Cahova et al., 2015). However, another study observed a three- to fourfold higher stability of NAD-RNAs in *nudC* strain, demonstrating a functional consequence of NCIN capping *in vivo* (Bird et al., 2016). It should be noted that some RNAs with high NAD cap content are inherently very stable. For example, SroB has a half-life of more than 30 minutes (Vogel et al., 2003). All these data suggest possible additional roles for prokaryotic capping apart from RNA stability.

It is conceivable that the impact of NAD-capping to physiological processes is maximal under certain conditions, perhaps associated with the redox state of the cell, changes in ATP/cofactor concentration ratio, or the production of certain proteins. Indeed, the levels of NCIN capping were approximately twofold higher in stationary phase than in exponential phase, demonstrating the growth dependence of capping (Bird et al., 2016). Since a number of proteins discriminate between NAD and NADH, an interesting possibility is that currently unknown players, perhaps specific for a certain state of the cofactor, participate in 5'-NAD-RNA function and processing.

Conclusion

The exciting discovery of modified 5' ends on bacterial RNAs poses an important question about the physiological significance of these modifications. Although cap-like structures appear to protect RNA from degradation, their low abundance makes it unclear if their impact on RNA stability is sufficiently strong to affect well-being of bacteria and if a larger percentage of *E. coli* transcripts become capped under certain growth conditions, resulting in a greater effect on cells. Since NAD capping was observed for sRNAs, there is a possibility that modulation of sRNA stability would affect the degradation rate of their mRNA targets. Alternatively, NAD and other caps may influence RNA function in unforeseen ways not related to degradation, perhaps by mediating binding to cellular enzymes that use NAD or other caps as a substrate. However, one cannot completely exclude the possibility that capped RNAs are merely the products of spurious transcription initiation and do not carry any physiological relevance. In this case, many transcripts starting with adenosine would have a chance to get capped and detected. Experimental data do not support this idea, and the question why only a small subset of RNAs becomes capped awaits further investigation.

In contrast to capped RNAs, 5'-diphosphorylated RNAs appear to have their place in the regular 5'-end-dependent RNA degradation pathway as the preferred substrates for RppH-mediated production of monophosphorylated RNAs. Intriguingly, *E. coli* appears to have an enzyme(s) that convert triphosphorylated RNAs to diphosphorylated species despite the ability of RppH to directly convert triphosphorylated substrates into monophosphorylated

products for degradation. If RppH evolved to degrade diphosphorylated RNAs that were generated as byproducts of unspecific enzymatic activities, it is unclear why RppH has greater reactivity on diphosphorylated than triphosphorylated RNAs and does not retain high reactivity with the triphosphorylated RNA that remain in cells. These facts argue against the conclusion that diphosphorylated RNAs are merely by-products of nonspecific enzymatic reactions or intermediates in RNA degradation pathway and raise the possibility of other function(s) of this modification.

However diverse their biological functions might be, the discovery of capped and diphosphorylated RNA in *E. coli* demonstrates that bacteria are capable of modifying the 5' ends of mRNAs and sRNAs and suggests that the 5' end of bacterial transcripts may be a busier and more influential place than heretofore imagined.

Acknowledgements

This research was supported by the NIH grant R01GM112940 to A.S. We thank Joel Belasco (New York University School of Medicine) for critical reading of the manuscript and discussions.

References

- Adhikari S, Xiao W, Zhao YL, & Yang YG (2016). m(6)A: Signaling for mRNA splicing. *RNA Biol*, 13(9), 756–759. doi:10.1080/15476286.2016.1201628 [PubMed: 27351695]
- Apirion D (1973). Degradation of RNA in *Escherichia coli*. A hypothesis. *Mol Gen Genet*, 122(4), 313–322. [PubMed: 4577538]
- Apirion D (1978). Isolation, genetic mapping and some characterization of a mutation in *Escherichia coli* that affects the processing of ribonucleic acid. *Genetics*, 90(4), 659–671. [PubMed: 369943]
- Arnold TE, Yu J, & Belasco JG (1998). mRNA stabilization by the ompA 5' untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA*, 4(3), 319–330. [PubMed: 9510333]
- Babitzke P, & Kushner SR (1991). The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 88(1), 1–5. [PubMed: 1846032]
- Belasco JG (2010). All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol*, 11(7), 467–478. doi:10.1038/nrm2917 [PubMed: 20520623]
- Belasco JG, & Higgins CF (1988). Mechanisms of mRNA decay in bacteria: a perspective. *Gene*, 72(1–2), 15–23. [PubMed: 3072246]
- Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, & Rabinowitz JD (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol*, 5(8), 593–599. doi:10.1038/nchembio.186 [PubMed: 19561621]
- Bi Y, Li H, Fan S, Xia B, & Jin C (2009). 1H, 13C and 15N resonance assignments of RNA pyrophosphohydrolase RppH from *Escherichia coli*. *Biomol NMR Assign*, 3(1), 149–151. doi: 10.1007/s12104-009-9162-8 [PubMed: 19636968]
- Bird JG, Zhang Y, Tian Y, Panova N, Barvik I, Greene L, ... Nickels BE (2016). The mechanism of RNA 5' capping with NAD⁺, NADH and desphospho-CoA. *Nature*, 535(7612), 444–447. doi: 10.1038/nature18622 [PubMed: 27383794]
- Bouvet P, & Belasco JG (1992). Control of RNase E-mediated RNA degradation by 5'-terminal base pairing in *E. coli*. *Nature*, 360(6403), 488–491. doi:10.1038/360488a0 [PubMed: 1280335]
- Bremer H, Konrad MW, Gaines K, & Stent GS (1965). Direction of chain growth in enzymic RNA synthesis. *J Mol Biol*, 13(2), 540–553. [PubMed: 5867035]
- Brook M, & Gray NK (2012). The role of mammalian poly(A)-binding proteins in co-ordinating mRNA turnover. *Biochem Soc Trans*, 40(4), 856–864. doi:10.1042/BST20120100 [PubMed: 22817748]

- Cahova H, Winz ML, Hofer K, Nubel G, & Jaschke A (2015). NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature*, 519(7543), 374–377. doi:10.1038/nature14020 [PubMed: 25533955]
- Celesnik H, Deana A, & Belasco JG (2007). Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol Cell*, 27(1), 79–90. doi:10.1016/j.molcel.2007.05.038 [PubMed: 17612492]
- Celesnik H, Deana A, & Belasco JG (2008). PABLO analysis of RNA: 5'-phosphorylation state and 5'-end mapping. *Methods Enzymol*, 447, 83–98. doi:10.1016/S0076-6879(08)02205-2 [PubMed: 19161839]
- Chandler DS (2011). Pre-mRNA Splicing In Schwab M (Ed.), *Encyclopedia of Cancer* (pp. 2972–2977). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Chao Y, Li L, Girodat D, Forstner KU, Said N, Corcoran C, ... Vogel J (2017). In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol Cell*, 65(1), 39–51. doi:10.1016/j.molcel.2016.11.002 [PubMed: 28061332]
- Chen YG, Kowtoniuk WE, Agarwal I, Shen Y, & Liu DR (2009). LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nat Chem Biol*, 5(12), 879–881. doi:10.1038/nchembio.235 [PubMed: 19820715]
- Condon C (2007). Maturation and degradation of RNA in bacteria. *Curr Opin Microbiol*, 10(3), 271–278. doi:10.1016/j.mib.2007.05.008 [PubMed: 17560162]
- Deana A, Celesnik H, & Belasco JG (2008). The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature*, 451(7176), 355–358. doi:10.1038/nature06475 [PubMed: 18202662]
- Decatur WA, & Fournier MJ (2002). rRNA modifications and ribosome function. *Trends Biochem Sci*, 27(7), 344–351. doi:10.1016/S0968-0004(02)02109-6 [PubMed: 12114023]
- Deutscher MP (2009). Chapter 9 Maturation and Degradation of Ribosomal RNA in Bacteria *Molecular Biology of RNA Processing and Decay in Prokaryotes* (Vol. 85, pp. 369–391): Academic Press.
- Emory SA, Bouvet P, & Belasco JG (1992). A 5'-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev*, 6(1), 135–148. [PubMed: 1370426]
- Even S, Pellegrini O, Zig L, Labas V, Vinh J, Brechemmier-Baey D, & Putzer H (2005). Ribonucleases J1 and J2: two novel endoribonucleases in *B. subtilis* with functional homology to *E. coli* RNase E. *Nucleic Acids Res*, 33(7), 2141–2152. doi:10.1093/nar/gki505 [PubMed: 15831787]
- Foley PL, Hsieh PK, Luciano DJ, & Belasco JG (2015). Specificity and evolutionary conservation of the *Escherichia coli* RNA pyrophosphohydrolase RppH. *J Biol Chem*, 290(15), 9478–9486. doi:10.1074/jbc.M114.634659 [PubMed: 25657006]
- Frick DN, & Bessman MJ (1995). Cloning, purification, and properties of a novel NADH pyrophosphatase. Evidence for a nucleotide pyrophosphatase catalytic domain in MutT-like enzymes. *J Biol Chem*, 270(4), 1529–1534. [PubMed: 7829480]
- Gao A, Vasilyev N, Luciano DJ, Levenson-Palmer R, Richards J, Marsiglia WM, ... Serganov A (2018). Structural and kinetic insights into stimulation of RppH-dependent RNA degradation by the metabolic enzyme DapF. *Nucleic Acids Res*. doi:10.1093/nar/gky327
- Gartner ZJ, & Liu DR (2001). The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *J Am Chem Soc*, 123(28), 6961–6963. [PubMed: 11448217]
- Gartner ZJ, Tse BN, Grubina R, Doyon JB, Snyder TM, & Liu DR (2004). DNA-templated organic synthesis and selection of a library of macrocycles. *Science*, 305(5690), 1601–1605. doi:10.1126/science.1102629 [PubMed: 15319493]
- Ghosh A, & Lima CD (2010). Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA*, 1(1), 152–172. doi:10.1002/wrna.19 [PubMed: 21956912]
- Giliberti J, O'Donnell S, Etten WJ, & Janssen GR (2012). A 5'-terminal phosphate is required for stable ternary complex formation and translation of leaderless mRNA in *Escherichia coli*. *RNA*, 18(3), 508–518. doi:10.1261/rna.027698.111 [PubMed: 22291205]
- Gopalakrishna Y, Langley D, & Sarkar N (1981). Detection of high levels of polyadenylate-containing RNA in bacteria by the use of a single-step RNA isolation procedure. *Nucleic Acids Res*, 9(14), 3545–3554. [PubMed: 6169008]

- Hausner G, Hafez M, & Edgell DR (2014). Bacterial group I introns: mobile RNA catalysts. *Mob DNA*, 5(1), 8. doi:10.1186/1759-8753-5-8 [PubMed: 24612670]
- Hofer K, Li S, Abele F, Frindert J, Schlotthauer J, Grawenhoff J, ... Jäschke A (2016). Structure and function of the bacterial decapping enzyme NudC. *Nat Chem Biol*, 12(9), 730–734. doi:10.1038/nchembio.2132 [PubMed: 27428510]
- Hor J, Gorski SA, & Vogel J (2018). Bacterial RNA biology on a genome scale. *Mol Cell*, 70(5), 785–799. doi:10.1016/j.molcel.2017.12.023 [PubMed: 29358079]
- Hsieh PK, Richards J, Liu Q, & Belasco JG (2013). Specificity of RppH-dependent RNA degradation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A*, 110(22), 8864–8869. doi:10.1073/pnas.1222670110 [PubMed: 23610425]
- Jorgensen SE, Buch LB, & Nierlich DP (1969). Nucleoside triphosphate termini from RNA synthesized in vivo by *Escherichia coli*. *Science*, 164(3883), 1067–1070. [PubMed: 4890175]
- Julius C, & Yuzenkova Y (2017). Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors. *Nucleic Acids Res*, 45(14), 8282–8290. doi:10.1093/nar/gkx452 [PubMed: 28531287]
- Kanan MW, Rozenman MM, Sakurai K, Snyder TM, & Liu DR (2004). Reaction discovery enabled by DNA-templated synthesis and in vitro selection. *Nature*, 431(7008), 545–549. doi:10.1038/nature02920 [PubMed: 15457254]
- Katahira J (2015). Nuclear export of messenger RNA. *Genes (Basel)*, 6(2), 163–184. doi:10.3390/genes6020163 [PubMed: 25836925]
- Kim D, Hong JS, Qiu Y, Nagarajan H, Seo JH, Cho BK, ... Palsson BO (2012). Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet*, 8(8), e1002867. doi:10.1371/journal.pgen.1002867 [PubMed: 22912590]
- Konrad M, Toivonen JE, & Cook J (1976). The 5' ends of bacterial RNA. II. The triphosphate-terminated ends of primary gene transcripts. *Biochim Biophys Acta*, 425(1), 63–75. [PubMed: 764878]
- Kowtoniuk WE, Shen Y, Heemstra JM, Agarwal I, & Liu DR (2009). A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. *Proc Natl Acad Sci U S A*, 106(19), 7768–7773. doi:10.1073/pnas.0900528106 [PubMed: 19416889]
- Krishnan KM, Van Etten WJ 3rd, & Janssen GR (2010). Proximity of the start codon to a leaderless mRNA's 5' terminus is a strong positive determinant of ribosome binding and expression in *Escherichia coli*. *J Bacteriol*, 192(24), 6482–6485. doi:10.1128/JB.00756-10 [PubMed: 20971908]
- Kruger MK, Pedersen S, Hagervall TG, & Sorensen MA (1998). The modification of the wobble base of tRNA^{Glu} modulates the translation rate of glutamic acid codons in vivo. *J Mol Biol*, 284(3), 621–631. doi:10.1006/jmbi.1998.2196 [PubMed: 9826503]
- Kuhnel MG, Strickland R, & Palmer JD (1990). An ancient group I intron shared by eubacteria and chloroplasts. *Science*, 250(4987), 1570–1573. [PubMed: 2125748]
- Le Rhun A, Lecrivain AL, Reimegard J, Proux-Wera E, Broglia L, Della Beffa C, & Charpentier E (2017). Identification of endoribonuclease specific cleavage positions reveals novel targets of RNase III in *Streptococcus pyogenes*. *Nucleic Acids Res*, 45(5), 2329–2340. doi:10.1093/nar/gkw1316 [PubMed: 28082390]
- Lee CR, Kim M, Park YH, Kim YR, & Seok YJ (2014). RppH-dependent pyrophosphohydrolysis of mRNAs is regulated by direct interaction with DapF in *Escherichia coli*. *Nucleic Acids Res*, 42(20), 12746–12757. doi:10.1093/nar/gku926 [PubMed: 25313159]
- Lee ER, Baker JL, Weinberg Z, Sudarsan N, & Breaker RR (2010). An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science*, 329(5993), 845–848. doi:10.1126/science.1190713 [PubMed: 20705859]
- Li S, & Mason CE (2014). The pivotal regulatory landscape of RNA modifications. *Annu Rev Genomics Hum Genet*, 15(1), 127–150. doi:10.1146/annurev-genom-090413-025405 [PubMed: 24898039]
- Li X, & Liu DR (2004). DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew Chem Int Ed Engl*, 43(37), 4848–4870. doi:10.1002/anie.200400656 [PubMed: 15372570]

- Luciano DJ, Hui MP, Deana A, Foley PL, Belasco KJ, & Belasco JG (2012). Differential control of the rate of 5'-end-dependent mRNA degradation in *Escherichia coli*. *J Bacteriol*, 194(22), 6233–6239. doi:10.1128/JB.01223-12 [PubMed: 22984254]
- Luciano DJ, Vasilyev N, Richards J, Serganov A, & Belasco JG (2017). A Novel RNA Phosphorylation State Enables 5' End-Dependent Degradation in *Escherichia coli*. *Mol Cell*, 67(1), 44–54 e46. doi:10.1016/j.molcel.2017.05.035 [PubMed: 28673541]
- Luciano DJ, Vasilyev N, Richards J, Serganov A, & Belasco JG (2018). Importance of a diphosphorylated intermediate for RppH-dependent RNA degradation. *RNA Biol*, 1–15. doi: 10.1080/15476286.2018.1460995
- Mackie GA (1998). Ribonuclease E is a 5'-end-dependent endonuclease. *Nature*, 395(6703), 720–723. doi:10.1038/27246 [PubMed: 9790196]
- Madore E, Florentz C, Giegé R, Sekine S. i., Yokoyama S, & Lapointe J (1999). Effect of modified nucleotides on *Escherichia coli* tRNA structure and on its aminoacylation by glutamyl-tRNA synthetase. *European Journal of Biochemistry*, 266(3), 1128–1135. doi:10.1046/j.1432-1327.1999.00965.x [PubMed: 10583410]
- Malygin AG, & Shemyakin MF (1979). Adenosine, NAD and FAD can initiate template-dependent RNA synthesis catalyzed by *Escherichia coli* RNA polymerase. *FEBS Lett*, 102(1), 51–54. [PubMed: 222618]
- Marbaniang CN, & Vogel J (2016). Emerging roles of RNA modifications in bacteria. *Curr Opin Microbiol*, 30, 50–57. doi:10.1016/j.mib.2016.01.001 [PubMed: 26803287]
- Mathy N, Benard L, Pellegrini O, Daou R, Wen T, & Condon C (2007). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell*, 129(4), 681–692. doi:10.1016/j.cell.2007.02.051 [PubMed: 17512403]
- McLennan AG (2006). The Nudix hydrolase superfamily. *Cell Mol Life Sci*, 63(2), 123–143. doi: 10.1007/s00018-005-5386-7 [PubMed: 16378245]
- Melefors O, & von Gabain A (1991). Genetic studies of cleavage-initiated mRNA decay and processing of ribosomal 9S RNA show that the *Escherichia coli* *ams* and *rne* loci are the same. *Mol Microbiol*, 5(4), 857–864. [PubMed: 1713283]
- Mengin-Lecreulx D, Michaud C, Richaud C, Blanot D, & van Heijenoort J (1988). Incorporation of LL-diaminopimelic acid into peptidoglycan of *Escherichia coli* mutants lacking diaminopimelate epimerase encoded by *dapF*. *J Bacteriol*, 170(5), 2031–2039. [PubMed: 3283102]
- Mildvan AS, Xia Z, Azurmendi HF, Saraswat V, Legler PM, Massiah MA, ... Amzel LM (2005). Structures and mechanisms of Nudix hydrolases. *Arch Biochem Biophys*, 433(1), 129–143. doi: 10.1016/j.abb.2004.08.017 [PubMed: 15581572]
- Mohanty BK, & Kushner SR (2011). Bacterial/archaeal/organellar polyadenylation. *Wiley Interdiscip Rev RNA*, 2(2), 256–276. doi:10.1002/wrna.51 [PubMed: 21344039]
- Motorin Y, & Helm M (2011). RNA nucleotide methylation. *Wiley Interdiscip Rev RNA*, 2(5), 611–631. doi:10.1002/wrna.79 [PubMed: 21823225]
- Mudd EA, Krisch HM, & Higgins CF (1990). RNase E, an endoribonuclease, has a general role in the chemical decay of *Escherichia coli* mRNA: evidence that *rne* and *ams* are the same genetic locus. *Mol Microbiol*, 4(12), 2127–2135. [PubMed: 1708438]
- Muramatsu T, Nureki O, Kanno H, Niimi T, Tateno M, Kohno T, ... Yokoyama S (1990). Recognition of tRNA identity determinants by aminoacyl-tRNA synthetases. *Nucleic Acids Symp Ser*(22), 119–120. [PubMed: 2101890]
- Ono M, & Kuwano M (1979). A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of messenger RNA. *J Mol Biol*, 129(3), 343–357. [PubMed: 110942]
- Piton J, Larue V, Thillier Y, Dorleans A, Pellegrini O, Li de la Sierra-Gallay I, ... Condon C (2013). *Bacillus subtilis* RNA deprotection enzyme RppH recognizes guanosine in the second position of its substrates. *Proc Natl Acad Sci U S A*, 110(22), 8858–8863. doi:10.1073/pnas.1221510110 [PubMed: 23610407]
- Ramanathan A, Robb GB, & Chan SH (2016). mRNA capping: biological functions and applications. *Nucleic Acids Res*, 44(16), 7511–7526. doi:10.1093/nar/gkw551 [PubMed: 27317694]

- Richards J, Liu Q, Pellegrini O, Celesnik H, Yao S, Bechhofer DH, ... Belasco JG (2011). An RNA pyrophosphohydrolase triggers 5'-exonucleolytic degradation of mRNA in *Bacillus subtilis*. *Mol Cell*, 43(6), 940–949. doi:10.1016/j.molcel.2011.07.023 [PubMed: 21925382]
- Richaud C, Higgins W, Mengin-Lecreux D, & Stragier P (1987). Molecular cloning, characterization, and chromosomal localization of *dapF*, the *Escherichia coli* gene for diaminopimelate epimerase. *J Bacteriol*, 169(4), 1454–1459. [PubMed: 3031013]
- Richaud C, & Printz C (1988). Nucleotide sequence of the *dapF* gene and flanking regions from *Escherichia coli* K12. *Nucleic Acids Res*, 16(21), 10367. [PubMed: 3057443]
- Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, ... Vogel J (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature*, 464(7286), 250–255. doi:10.1038/nature08756 [PubMed: 20164839]
- Shell SS, Wang J, Lapierre P, Mir M, Chase MR, Pyle MM, ... Gray TA (2015). Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. *PLoS Genet*, 11(11), e1005641. doi:10.1371/journal.pgen.1005641 [PubMed: 26536359]
- Shine J, & Dalgarno L (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A*, 71(4), 1342–1346. [PubMed: 4598299]
- Sokolowski M, Klassen R, Bruch A, Schaffrath R, & Glatt S (2018). Cooperativity between different tRNA modifications and their modification pathways. *Biochim Biophys Acta*, 1861(4), 409–418. doi:10.1016/j.bbarm.2017.12.003
- Taraseviciene L, Miczak A, & Apirion D (1991). The gene specifying RNase E (*rne*) and a gene affecting mRNA stability (*ams*) are the same gene. *Mol Microbiol*, 5(4), 851–855. [PubMed: 1713282]
- Tian B, & Manley JL (2013). Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci*, 38(6), 312–320. doi:10.1016/j.tibs.2013.03.005 [PubMed: 23632313]
- Tian B, & Manley JL (2017). Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol*, 18(1), 18–30. doi:10.1038/nrm.2016.116 [PubMed: 27677860]
- Topisirovic I, Svitkin YV, Sonenberg N, & Shatkin AJ (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA*, 2(2), 277–298. doi:10.1002/wrna.52 [PubMed: 21957010]
- Vasilyev N, & Serganov A (2015). Structures of RNA complexes with the *Escherichia coli* RNA pyrophosphohydrolase RppH unveil the basis for specific 5'-end-dependent mRNA decay. *J Biol Chem*, 290(15), 9487–9499. doi:10.1074/jbc.M114.634824 [PubMed: 25657011]
- Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, Huttenhofer A, & Wagner EG (2003). RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res*, 31(22), 6435–6443. [PubMed: 14602901]
- Vvedenskaya IO, Bird JG, Zhang Y, Zhang Y, Jiao X, Barvik I, ... Nickels BE (2018). CapZyme-Seq comprehensively defines promoter-sequence determinants for RNA 5' capping with NAD⁺. *Mol Cell*, 70(3), 553–564 e559. doi:10.1016/j.molcel.2018.03.014 [PubMed: 29681497]
- Wang Q, Zhang D, Guan Z, Li D, Pei K, Liu J, ... Yin P (2018). DapF stabilizes the substrate-favoring conformation of RppH to stimulate its RNA-pyrophosphohydrolase activity in *Escherichia coli*. *Nucleic Acids Res*, 46(13), 6880–6892. doi:10.1093/nar/gky528 [PubMed: 29931175]
- Wang X, & He C (2014). Dynamic RNA modifications in posttranscriptional regulation. *Mol Cell*, 56(1), 5–12. doi:10.1016/j.molcel.2014.09.001 [PubMed: 25280100]
- Wilson RK, & Roe BA (1989). Presence of the hypermodified nucleotide N6-(delta 2-isopentenyl)-2-methylthioadenosine prevents codon misreading by *Escherichia coli* phenylalanyl-transfer RNA. *Proc Natl Acad Sci U S A*, 86(2), 409–413. [PubMed: 2643111]
- Wurtzel O, Sesto N, Mellin JR, Karunker I, Edelheit S, Becavin C, ... Sorek R (2012). Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol*, 8, 583. doi: 10.1038/msb.2012.11 [PubMed: 22617957]
- Xu MQ, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, & Shub DA (1990). Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science*, 250(4987), 1566–1570. [PubMed: 2125747]

- Zhang D, Liu Y, Wang Q, Guan Z, Wang J, Liu J, ... Yin P (2016). Structural basis of prokaryotic NAD-RNA decapping by NudC. *Cell Res*, 26(9), 1062–1066. doi:10.1038/cr.2016.98 [PubMed: 27561816]
- Zheng X, Hu GQ, She ZS, & Zhu H (2011). Leaderless genes in bacteria: clue to the evolution of translation initiation mechanisms in prokaryotes. *BMC Genomics*, 12, 361. doi: 10.1186/1471-2164-12-361 [PubMed: 21749696]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1 |.
Typical 5' ends of mRNA. (a), A cap structure on mRNA of eukaryotes. (b), A triphosphorylated end of primary bacterial transcripts.

Figure 2 |.

RNA degradation pathways in *E. coli*. (a) Direct access pathway. The pathway begins with the initial internal cleavage by RNase E. The resulting 5' fragment is then degraded by 3'-end-dependent exonucleases while the 3' fragment, typically protected from exonucleases by a stem-loop structure on the 3' end, is subjected to new rounds of RNase E/exonuclease cleavage. (b) 5' end-dependent pathway. In *E. coli*, this pathway begins with conversion of 5'-triphosphorylated RNAs to monophosphorylated species, a process that involves the activity of RppH. The monophosphorylated RNAs are optimal substrates for 5'-end-dependent RNase E cleavage, followed by rounds of RNase E/exonuclease digestion. (c) Recent modifications of the 5'-end-dependent RNA degradation pathway of *E. coli*. Conversion of the triphosphorylated 5'-ends of primary transcripts to monophosphorylated species requires the consecutive removal of phosphate by an unknown enzyme(s) and RppH. DapF stimulates RppH activity.

Figure 3 |.

Structures of *E. coli* RppH. (a) Structure of pppAGU-bound RppH (cyan sticks) shown with electrostatic surface representation of the protein (Vasilyev & Serganov, 2015). Mg²⁺ cations are depicted as magenta spheres. (b) A view of the RNA-binding cleft. Amino acids recognizing guanosine and adjacent moieties are in sticks. Intermolecular hydrogen bond and cation- π interactions are shown as black and red dashed lines, respectively. (c) Active site of RppH. Hydrogen bonds between RNA and RppH are shown as black dashed lines. Coordination bonds between RNA and Mg²⁺ cations are shown as red dashed lines. A water molecule likely involved in catalysis is shown as a red sphere. A black arrow shows in-line attack of the water molecule on the β phosphorus atom. (d) Structure of the ternary *E. coli* RppH-RNA-DapF complex with RNA (red sticks) bound to RppH (Gao et al., 2018).

Figure 4 |.

Non-canonical caps on bacterial mRNAs. (a), Cap structures identified experimentally in *E. coli* (Chen et al., 2009; Kowtoniuk et al., 2009). (b), Nucleotide derivatives which *Eco* RNAP can use for transcription *in vitro* (Julius & Yuzenkova, 2017).

Figure 5 |.

Structural basis of NCIN-mediated transcription initiation. (a), Crystal structures of *Tth* RNAP bound to pppApC (Bird et al., 2016). DNA is in grey. RNA is shown in atomic colors: cyan, carbon; orange, phosphorus; red, oxygen; and blue, nitrogen atoms. Protein is in atomic colors with carbon atoms in green. Hydrogen bonds are shown by dark blue dashed lines. The Mg²⁺ cation is depicted as a green sphere. NA, nicotinamide. (b), Crystal structures of *Tth* RNAP bound to NAD+pC (Bird et al., 2016). A red sphere represents a water molecule.

Figure 6 |.

Structures of the *E. coli* NudC enzyme. (a) Overall structure of the NudC dimer (Hofer et al., 2016). Bound NAD is shown in stick representation and atomic colors (carbon, yellow; oxygen, red; nitrogen, blue; and phosphorus, orange). (b) Zoomed-in view of the NAD-binding site. NADs from two available structures are superposed and shown in different colors of carbon atoms (yellow, ref. (Zhang et al., 2016); magenta, ref. (Hofer et al., 2016)). Intermolecular hydrogen bonds are depicted by black dashed lines. Amino acids shown in sticks are glutamates involved in catalysis and aromatic residues involved in stacking with the adenine base of NAD. (c) Superposition of NudC (green) (Zhang et al., 2016) and RppH (brown) (Vasilyev & Serganov, 2015) on the Nudix helix. RppH-bound RNA is in blue, and Mg^{2+} cations from the RppH structure are in magenta. (d) Superposition of the reaction substrate NAD (yellow) (Zhang et al., 2016) and reaction product NMN (cyan) (Hofer et al., 2016) in the NudC structures. (e) Electrostatic surface view of the catalytic pocket of NudC bound to NAD (yellow) (Zhang et al., 2016). The view shows that the pocket can accommodate dpCoA (violet), which was modelled based on the NudC-NAD structure (Zhang et al., 2016).