



Published in final edited form as:

Nat Chem Biol. 2009 December ; 5(12): 879–881. doi:10.1038/nchembio.235.

LC/MS analysis of cellular RNA reveals NAD-linked RNA

Y. Grace Chen[†], Walter E. Kowtoniuk[†], Isha Agarwal, Yinghua Shen, and David R. Liu^{*}

Howard Hughes Medical Institute and the Department of Chemistry and Chemical Biology
Harvard University Cambridge, Massachusetts 02138

Abstract

We developed a general method to detect cellular small molecule-RNA conjugates that does not rely on the reactivity of the small molecule, revealing NAD-linked RNA in *E. coli* and *S. venezuelae*. Subsequent characterization shows NAD is a 5' modification of RNA, cannot be installed *in vitro* through aberrant transcriptional initiation, is only found among smaller cellular RNAs, and is present at a surprisingly high abundance of ~3000 copies per cell.

The known chemical diversity of natural RNA has remained limited primarily to canonical polyribonucleotides, 3'-aminoacylated tRNAs, and modified nucleobases in tRNA and rRNA1 in contrast with a growing number of elucidated biological roles for RNA that include catalysis, gene regulation, and defense against viral infection^{2,3}. Researchers have previously speculated that early biotic systems carried out biochemical reactions using small molecule-RNA conjugates as participants in RNA-templated chemistries⁴⁻⁶. Recent studies of non-natural DNA-templated chemistries^{7,8} also highlight the unusual functional capabilities of small molecule-nucleic acid conjugates. These observations collectively led us to speculate that small molecule-RNA conjugates beyond those previously described may exist in modern cells⁹.

In the six decades since their initial discovery¹⁰, researchers have identified over 100 modified nucleosides, almost all in RNA^{11,12}. Virtually all past efforts to characterize modified nucleotides, however, have examined modifications in a targeted manner by studying a specific cellular RNA or class of RNAs^{13,14}. We recently described the development of a broad approach to the discovery of biological small molecule-RNA conjugates that uses size-exclusion chromatography and mass spectrometry to detect base- or nucleophile-labile small molecules that are cleaved from any cellular RNA⁹. While fruitful, we sought 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.

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

^{*}To whom correspondence should be addressed: David R. Liu Oxford Street Cambridge, Massachusetts 02138
drliu@fas.harvard.edu Tel: (617) 496-1067 Fax: (617) 496-5688.

[†]These authors contributed equally to this work.

Author contributions:

Y.G.C., W.E.K., and D.R.L. designed research; Y.G.C., W.E.K., I.A., and Y.S. performed research; Y.G.C., W.E.K., and I.A. contributed new reagents/analytic tools; Y.G.C., W.E.K., I.A., Y.S., and D.R.L. analyzed data; and Y.G.C., W.E.K., Y.S., and D.R.L. wrote the paper.

The authors declare no competing financial interests.

Here we report the development and application of a method that in principle enables the detection of any small molecule-RNA conjugate independent of its chemical structure (Fig. 1a and Supplementary Results). Whole cellular RNA is subjected to size-exclusion chromatography and the macromolecular fraction ($> \sim 2,500$ Da) is divided into two halves (Supplementary Methods). One half is treated with nuclease P115, while the second half is treated with heat-inactivated nuclease P1 under otherwise identical conditions. Both samples are subjected to size-exclusion chromatography again and the small-molecule fraction from each is subjected to comparative high-resolution LC/MS9. Non-canonical species more abundant in the active nuclease-treated sample compared with the inactive nuclease-treated are considered possible novel small molecule-RNA conjugates (Supplementary Fig. 1). After validating this method by observing 16 of the 20 major 3'-aminoacyl adenosine monophosphates as enriched 2-fold in samples treated with active nuclease compared with heat-inactivated nuclease (Supplementary Figs. 2 and 3), we applied the general approach to small molecule-RNA discovery to *E. coli* and *S. venezuelae* RNA, revealing 24 and 28 unknown species, respectively, that were enriched 2-fold (Supplementary Fig. 4).

One unknown species from both *E. coli* and *S. venezuelae* that was enriched 8-fold in the nuclease versus heat-inactivated nuclease samples was $[M-H]^- m/z = 540.0533$ (Figs. 1b and 1c). Culturing *S. venezuelae* in media containing ^{13}C -glucose (**1**) as the sole carbon source, or containing ^{15}N -ammonium sulfate as the sole nitrogen source, resulted in mass increases of this species of +15 Da and +5 Da, respectively (Fig. 1d). These results enabled us to deduce a molecular formula of $C_{15}H_{20}N_5O_{13}P_2$ (expected $[M-H]^- m/z = 540.0538$). The MS/MS spectrum of this ion revealed ADP (expected $[M-H]^- m/z = 426.0221$) as a major fragment (Supplementary Fig. 5). We therefore reasoned that the 540.0533 Da species likely consists of a 115.0395 Da group ($C_5H_7O_3$) attached to ADP.

An elimination reaction on an adenosine-containing cofactor could generate an unsaturated ribose group. We therefore speculated that the $[M-H]^- m/z = 540.0533$ ion is a breakdown product of a larger small molecule-RNA conjugate that undergoes elimination during MS ionization. Indeed, when MS analysis was performed using milder ionization conditions, a $[M-H]^- m/z = 662.1032$ species appeared at the same retention time as the $[M-H]^- m/z = 540.0533$ species (Figs. 1e and 1f). Collectively, these observations led us to propose that the $[M-H]^- m/z = 540.0533$ species is a fragment of nicotinamide adenine dinucleotide (NAD) (**2**) (expected $[M-H]^- m/z = 662.1018$). This hypothesis was confirmed by LC/MS and MS/MS comparisons of the cellular species with authentic NAD (Supplementary Fig. 5 and Supplementary Fig. 6). The analysis of whole *E. coli* RNA using the milder ionization conditions revealed that $[M-H]^- m/z = 662.1032$ was also significantly enriched (8-fold) in the nuclease versus heat-inactivated nuclease samples, consistent with these conclusions.

Since NAD is prevalent as a cellular metabolite and cofactor, we sought to confirm that the detected NAD species was a bona fide small molecule-RNA conjugate such as the structure shown in Fig. 1g rather than an intracellular contaminant that was unexpectedly carried through RNA purification and size exclusion. We spiked varying quantities of authentic NAD or NADH into *E. coli* and *S. venezuelae* cell lysates, and repeated the RNA isolation, nuclease P1 digestion, and LC/MS analysis. The addition of up to 10,000-fold more NAD and NADH than was observed in the unspiked samples did not significantly change the

abundance of the corresponding NAD species (Fig. 2a). These results demonstrate that the NAD species observed in our experiments on *E. coli* and *S. venezuelae* RNA cannot be accounted for by endogenous NAD (or by the oxidation of endogenous NADH), and further support the conclusion that these species arise from cellular small molecule-RNA conjugates.

The structure of NAD led us to hypothesize that NAD is a 5' RNA modification. Because nuclease P1 catalyzes the attack of a water molecule on RNA to generate 5'-phosphonucleotides¹⁵, all nuclease P1 digestion products other than the nucleotides at the 5' termini undergo a mass shift when the digestion is performed in the presence of isotopically labeled water, compared with digestion in unlabeled water. We carried out nuclease P1 digestion of total RNA from both *E. coli* and *S. venezuelae* in ¹⁸O-enriched water. As expected, Phe-AMP (3) exhibited a +2 Da mass shift when the digestion was carried out in ¹⁸O water compared to in ¹⁶O water (Supplementary Fig. 7a). In contrast, there was no mass shift for NAD in both isotopically-labeled and unlabeled water (Supplementary Fig. 7b). This finding is consistent with a model in which NAD is present at the 5' termini of RNA (Fig. 1g).

We quantified the amount of NAD-linked RNA in *E. coli* cells using two methods. First, we generated a standard curve that relates known quantities of authentic NAD to observed ion counts under both the original (harsher) and the milder ionization conditions. Ion counts from *E. coli* RNA when plotted on the resulting curve result in the estimate of ~3000 copies of NAD-RNA per *E. coli* cell for both ionization conditions (Supplementary Figs. 8a and b). In the second approach, we added a known quantity of authentic NAD to the nuclease P1-digested cellular RNA sample before LC/MS analysis, and used the resulting increase in the NAD ion count to relate added NAD concentration with observed ion counts. This second method resulted in a similar estimate of ~3300 copies of NAD-RNA per *E. coli* cell (Supplementary Fig. 8b). This abundance level is comparable to that of Phe-linked tRNA in *E. coli*¹⁶, ~4-fold higher than that of *E. coli* tmRNA¹⁷, and ~300-fold higher than that of short RNAs generated by DNA primase during DNA synthesis¹⁸.

Structural similarities between NAD and ATP led us to speculate that NAD might be incorporated into RNA at the 5'-terminus through aberrant transcriptional initiation with NAD or NADH instead of ATP. Indeed, NAD has been incorporated into the 5'-terminus of RNA transcripts *in vitro* using T7 RNA polymerase¹⁹. To test if transcriptional initiation was responsible for incorporating NAD into RNA transcripts, we used *E. coli* RNA polymerase to carry out *in vitro* transcription in the presence of high concentrations of NADH using a modified pUC19 plasmid or *E. coli* genomic DNA as templates. When the RNA was purified, digested with nuclease P1, and analyzed by LC/MS, no NAD or NADH was detected (Supplementary Fig. 9). In contrast, when an authentic 5'-NAD-linked transcript (generated using T7 RNA polymerase¹⁹) was spiked into an *in vitro* transcription reaction and processed in the same way, NAD-linked RNA was readily detected (Supplementary Fig. 9). If one assumes that the inability of *E. coli* RNA polymerase to incorporate these levels of NAD *in vitro* parallels an inability to do so in cells, these results suggest that NAD groups are installed following transcriptional initiation.

Prior to nuclease P1 digestion, size-exclusion chromatography retains molecules of molecular weight greater than ~2,500 Da. To narrow the size range of possible NAD-linked RNAs, we used silica-based RNA purification columns (Qiagen RNeasy columns) to further fractionate RNA molecules into two fractions that are less than or greater than ~200 nucleotides (Supplementary Fig. 10). Each of the two fractions was then subjected to nuclease P1 digestion and LC/MS analysis. The rRNA nucleoside modification N^6,N^6 -dimethyladenine (**4**) (conjugated to 1.5 kB-2.9 kB rRNAs) was detected in the > 200 base fraction and 3'-aminoacyl adenosine monophosphates conjugated to tRNAs (~76 nucleotides) were present predominantly in the <200 nucleotide flow-through fraction, as expected (Fig. 2b)²⁰. The NAD-linked nucleotides were predominantly detected in the flow-through fraction. This result suggests that the NAD-linked RNA(s) from *E. coli* and *S. venezuelae* are not widely distributed in size but instead are predominantly below ~200 nucleotides in length, as was also observed in the case of CoA-linked RNA⁹. In addition, this finding further supports the hypothesis that the NAD modifications arise through a mechanism other than non-specific transcriptional initiation, which would be expected to generate a broader size distribution of NAD-linked RNAs.

The method developed in this work enables the detection, in principle, of any small molecule-RNA conjugate. The application of this method resulted in the discovery of NAD-RNA in *E. coli* and *S. venezuelae*, as well as in the detection of a variety of additional unknown, non-canonical small molecule-RNA conjugates. Following the recent discovery of CoA-linked RNA, these results further indicate that the chemical diversity of biological RNA is greater than previously understood.

The biological role, if any, of the NAD-RNA conjugates remains unknown at this point. Perhaps the simplest such possible role is one in which the NAD group serves to localize NAD-linked RNA(s) to NAD-binding proteins. It is also tempting to speculate that NAD-linked RNA (or NADH-linked RNA) might mediate redox reactions. In such a scenario, the RNA component of the NAD-linked RNA might assist in localizing or positioning the NAD for reaction with the appropriate substrates. Indeed, laboratory-evolved ribozymes that use cofactors including NAD and CoA have been previously reported although no natural ribozyme has yet been discovered that requires either NAD or CoA^{21,22}. Alternatively, based on previous examples of RNAs that are either conjugated to small molecules²³ or that bind small molecules²⁴, it is also possible that the NAD-linked RNAs might play a role in RNA stability or even in gene regulation. Indeed, the state of the 5' terminus of RNA can serve as determinant of mRNA half-life²⁵, and it is possible that NAD (or CoA) at the 5' terminus of a cellular RNA protects an RNA from degradation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Howard Hughes Medical Institute and the NIH/NIGMS (R01GM065865). Y.G.C and W.E.K gratefully acknowledge NSF Graduate Research Fellowships.

References

1. Hoagland MB, Stephenson ML, Scott JF, Hecht LI, Zamecnik PC. A soluble ribonucleic acid intermediate in protein synthesis. *J Biol Chem.* 1958; 231:241–57. [PubMed: 13538965]
2. Waters LS, Storz G. Regulatory RNAs in bacteria. *Cell.* 2009; 136:615–28. [PubMed: 19239884]
3. Repoila F, Darfeuille F. Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell.* 2009; 101:117–31. [PubMed: 19076068]
4. Szostak JW, Bartel DP, Luisi PL. Synthesizing life. *Nature.* 2001; 409:387–90. [PubMed: 11201752]
5. Visser CM, Kellogg RM. Bioorganic chemistry and the origin of life. *J Mol Evol.* 1978; 11:163–8. [PubMed: 671563]
6. White HB 3rd. Coenzymes as fossils of an earlier metabolic state. *J Mol Evol.* 1976; 7:101–4. [PubMed: 1263263]
7. Gartner ZJ, et al. DNA-templated organic synthesis and selection of a library of macrocycles. *Science.* 2004; 305:1601–5. [PubMed: 15319493]
8. Tse BN, Snyder TM, Shen Y, Liu DR. Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J Am Chem Soc.* 2008; 130:15611–26. [PubMed: 18956864]
9. Kowtoniuk WE, Shen Y, Heemstra JM, Agarwal I, Liu DR. A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. *Proc Natl Acad Sci U S A.* 2009; 106:7768–73. [PubMed: 19416889]
10. Wyatt GR. Occurrence of 5-methylcytosine in nucleic acids. *Nature.* 1950; 166:237–8. [PubMed: 15439258]
11. Limbach PA, Crain PF, McCloskey JA. Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* 1994; 22:2183–96. [PubMed: 7518580]
12. Pomerantz SC, McCloskey JA. Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry. *Methods Enzymol.* 1990; 193:796–824. [PubMed: 1706064]
13. Grosjean, H. Topics in current genetics. Vol. 12. New York: 2005. Fine-tuning of RNA functions by modification and editing..
14. Hiley SL, et al. Detection and discovery of RNA modifications using microarrays. *Nucleic Acids Res.* 2005; 33:e2. [PubMed: 15640439]
15. Romier C, Dominguez R, Lahm A, Dahl O, Suck D. Recognition of single-stranded DNA by nuclease P1: high resolution crystal structures of complexes with substrate analogs. *Proteins.* 1998; 32:414–24. [PubMed: 9726413]
16. Jakubowski H, Goldman E. Quantities of individual aminoacyl-tRNA families and their turnover in *Escherichia coli*. *J Bacteriol.* 1984; 158:769–76. [PubMed: 6373741]
17. Moore SD, Sauer RT. Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*. *Mol Microbiol.* 2005; 58:456–66. [PubMed: 16194232]
18. Frick DN, Richardson CC. DNA primases. *Annu Rev Biochem.* 2001; 70:39–80. [PubMed: 11395402]
19. Huang F. Efficient incorporation of CoA, NAD and FAD into RNA by in vitro transcription. *Nucleic Acids Res.* 2003; 31:e8. [PubMed: 12560511]
20. Brosius J, Dull TJ, Noller HF. Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli*. *Proc Natl Acad Sci U S A.* 1980; 77:201–4. [PubMed: 6153795]
21. Tsukiji S, Pattnaik SB, Suga H. Reduction of an aldehyde by a NADH/Zn²⁺-dependent redox active ribozyme. *J Am Chem Soc.* 2004; 126:5044–5. [PubMed: 15099068]
22. Breaker RR, Joyce GF. Self-incorporation of coenzymes by ribozymes. *J Mol Evol.* 1995; 40:551–8. [PubMed: 7643406]
23. Furuichi Y, Miura K. A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. *Nature.* 1975; 253:374–5. [PubMed: 163011]
24. Mandal M, Breaker RR. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol.* 2004; 5:451–63. [PubMed: 15173824]

25. Celesnik H, Deana A, Belasco JG. Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol Cell*. 2007; 27:79–90. [PubMed: 17612492]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

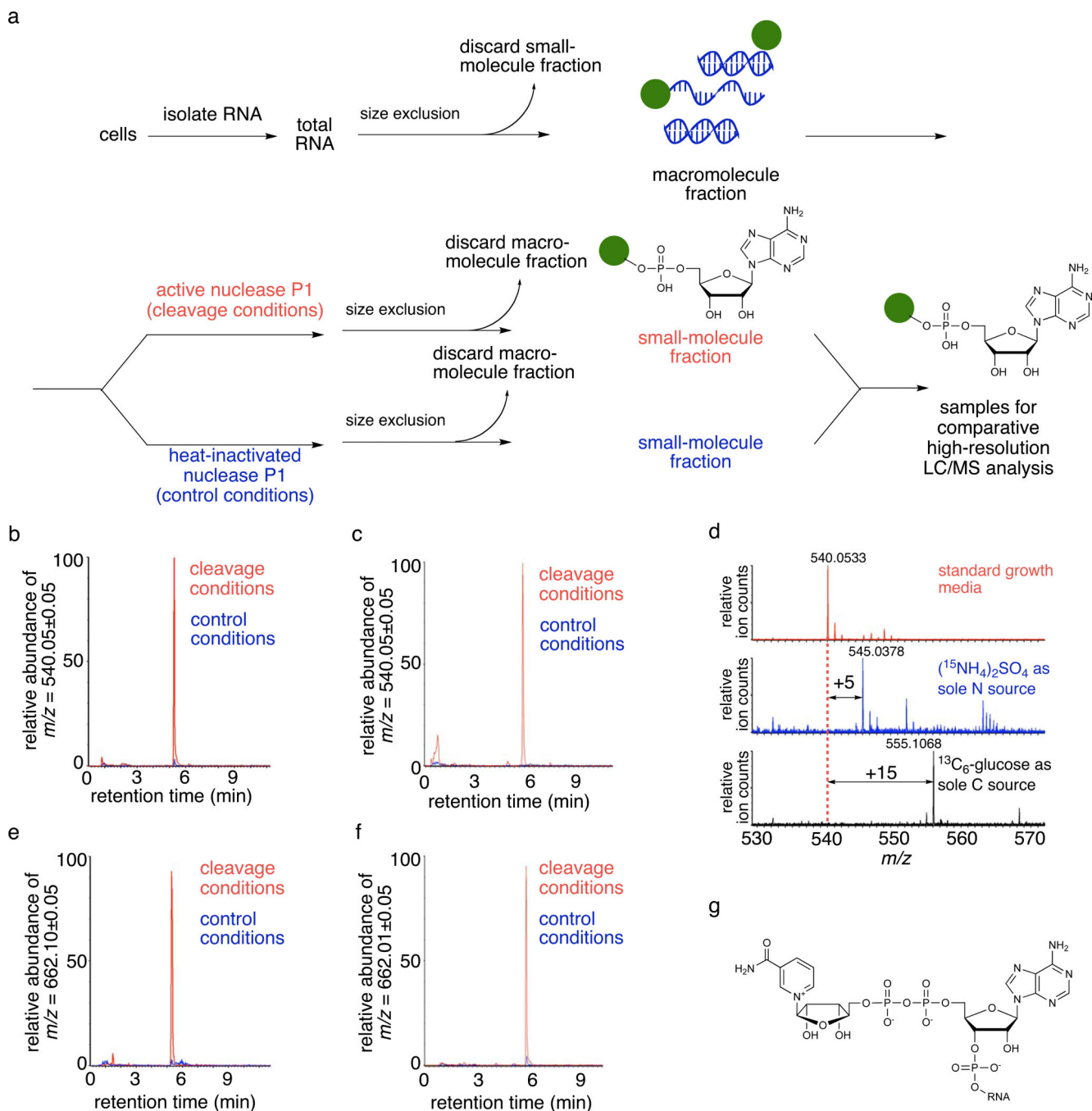


Figure 1.

Discovery of a small molecule-linked nucleotide of $[\text{M-H}]^-$ $m/z = 540.0533$ from *E. coli* and *S. venezuelae* RNA. (a) The general method for biological small molecule-RNA conjugate discovery developed in this work. Paired samples for comparative LC/MS analysis are generated by treatment of cellular RNA with active nuclease P1 versus treatment with heat-inactivated nuclease P1 under otherwise identical conditions. (b) The extracted ion chromatograms (EICs) for $[\text{M-H}]^-$ $m/z = 540.0533$ from *E. coli* RNA exposed to active nuclease P1 (cleavage conditions) or to heat-inactivated nuclease P1 (control conditions). (c)

Same as **(b)**, but using *S. venezuelae* RNA. **(d)** Isotope-labeled *S. venezuelae* total RNA reveals that the $[M-H]^-$ $m/z = 540.0533$ ion contains five nitrogen atoms and 15 carbon atoms. **(e)** The EICs for $[M-H]^-$ $m/z = 662.1018$ from *E. coli* RNA or **(f)** from *S. venezuelae* RNA digested with nuclease P1 or control conditions (heat-inactivated nuclease incubation). **(g)** One possible structure of NAD-linked cellular RNA. We note that our data is also consistent with RNA attachment to the 2' hydroxyl of NAD, or to either of the nicotinamide-linked ribose hydroxyl groups.

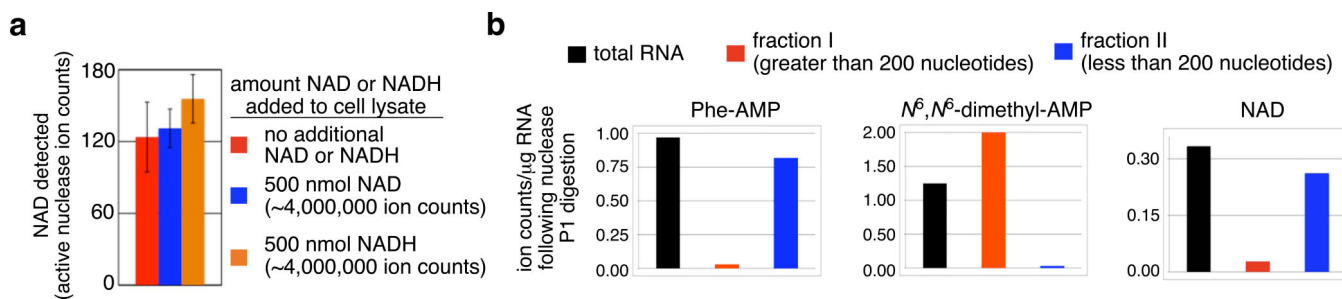


Figure 2.

Characterization of NAD-RNA. **(a)** Spiking large quantities of NAD or NADH into *E. coli* cell lysate before RNA isolation and treatment with active nuclease P1 does not change the observed ion counts of these species. Error bars represent the standard deviation of three independent trials. **(b)** Total *E. coli* RNA was separated into RNAs of length greater than 200 nucleotides (fraction I) and RNAs of length less than 200 nucleotides (fraction II) using a silica column (Qiagen RNeasy). Each fraction was subjected to nuclease P1 digestion and analyzed by LC/MS. The presence of NAD in fraction II, similar to that of Phe-AMP, suggests that the NAD-linked RNA(s) are primarily less than 200 nucleotides in length.

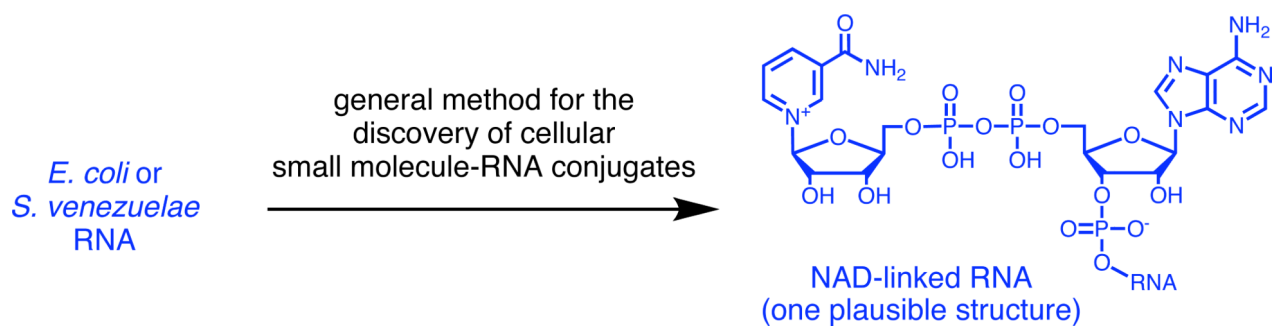


Figure 3.