The metabolome alters the thermodynamic and chemical stability of RNA

ABSTRACT: Herein, we examine the complex network of interactions among RNA, the metabolome, and the metalome in conditions that mimic the *E. coli* cytoplasm. We determined Mg2+ binding constants for the top 15 *E. coli* metabolites, comprising 80% of the total metabolome, at physiological pH and monovalent ion concentrations. This information was used to inform development of artificial cytoplasms that mimic *in vivo E. coli* conditions, termed “Eco80”. We empirically determined that the mixture of *E. coli* metabolites in Eco80 approximates single site binding behavior towards Mg2+ in the biologically relevant free Mg2+ range of ~0.5 to 3 mM Mg2+, using a Mg2+ binding fluorescent dye . Effects of Eco80 conditions on the thermodynamic stability, chemical stability, structure, and catalysis of RNA were examined. We find that Eco80conditions lead to opposing effects on the thermodynamics and chemical stability of RNA. The thermodynamic stability of RNA helices were weakened but chemical stability, compactness, and catalysis of RNA were enhanced, dependent on the speciation of Mg2+ between weak and strong Mg2+-metabolite complexes in Eco80. Thus, the effects of Eco80 enhance RNA function and increase the biological relevance of mechanistic studies of RNA *in-vitro*.

**Introduction**

Summary of progress on mimicking in vivo-like conditions (Ref Kate’s QRB).

Set up *aufbau* approach in paragraph 1. Start with top down. Pielak, Gruebele, N. Walters. Neet to simplify and work of Znosko and Sugimoto.

In contrast, we take a bottom up, *aufbau*, approach that builds up complexity, to make an artificial cytoplasm that contains 80% of *E. coli* metabolites with biologically relevant concentrations of monovalent ions and free Mg2+ ions.This *aufbau* approach allows us to understand the effects of 80% of metabolite and metal ion species that compose a part of the network of interactions that RNA experiences in *E. coli* cells.

**Results**

**Eco80: An artificial cytoplasm containing 80% of *E. coli* metabolites**

*E. coli* cells contain hundreds of metabolites (~240 mM total), which is too many metabolites to test systematically. However, the 15 most abundant metabolites, an experimentally manageable number, comprise 80% (195 mM) of total metabolites (Figure 1A). Thus, we sought to prepare Eco80*,* an artificial cytoplasm containing the biological concentrations of the 15 most abundant metabolites in *E. coli* (Table 1).

Eco80 was prepared at a 2x final concentration so that it could be diluted into other reagents for experiments and contain physiological concentrations of monovalent metal ions at pH 7.0 (see methods for details). Briefly, all of the metabolites in Eco80 are zwitterions or negatively charged near physiological pH 7, which requires electrostatic neutralization with metal ions. Metabolite salts and free acids were prepared to a final 2xconcentration, and the amount of Na+ and K+ added with each metabolite was recorded (Supplementary information (SI) Table 1). Next, the pH of the 2xstock was adjusted to pH 7.0 using NaOH, and the amount of Na+ was recorded. Lastly, NaCl and KCl were added to a final 480 mM Na+ and 280 mM K+, twice the physiological value of 240 mM Na+ and 140 mM K+.

Next, we considered how metabolites affect the speciation of free and chelated Mg2+. All 15 Eco80 metabolites have functional groups, carboxylates and phosphates, that drive chelating interactions with Mg2+ ions (Table 1), and we have previously estimated thatthe metabolite pool in *E. coli* has potential to chelate 51 mM Mg2+, assuming 2 mM free Mg2+, at an ionic strength of 0.15 M and a pH of 7.5. While extensive literature exists on chelating interactions between Mg2+ and small molecules, our previous estimates are approximate as Mg2+ binding affinity is dependent on environmental factors such as pH, ionic strength, composition of background ions, and temperature. Thus, we sought to better characterize Mg2+ chelation by the metabolites inEco80,at the physiological background.

We determined apparent disassociation constants (KD) for Eco80 metabolites in 240 mM NaCl, 140 mM KCl, pH 7.0 buffer at 37 °C (Table 1). Isothermal titration calorimetry (ITC) was used to measure KDs for phosphorylated metabolites (SI figure 1, SI Table 2). A fluorescence assay, that measures the free Mg2+ concentration in a sample using the metal ion binding dye 8-Hydroxy-5-quinolinesulfonic (HQS) acid, was used to estimate the KD for Mg2+ of metabolites that did not produce enough heat on binding Mg2+ to measure with ITC (SI figure 2, SI Table 3). For the HQS assay, Mg2+ is titrated into HQS solutions in the absence and presence of chelators. HQS emission as a function of the total Mg2+ in the absence of chelators is then fit to a binding model (SI figure 2A, top blue data and black fit). The free Mg2+ concentration is then then calculated from the fluorescence emission for each data point using the binding model, providing the free Mg2+ concentration as a function of the total Mg2+ concentration (SI figure 2A, bottom). Mg2+ binding by metabolites is thus obtained by fitting the free Mg2+ concentration as a function of the total Mg2+ concentration, which is shifted to the right as Mg2+ is sequestered by metabolites.

The binding affinity for Eco80 metabolites and Mg2+ range from extremely strong to negligible. The four nucleotide triphosphates, ATP, UTP, GTP, and dTTP, were classified as strong Mg2+ binders, with KD values less than the approximate free Mg2+ concentration in *E. coli* of 2 mM (Table 1). Conversely, 8 other metabolites, L-glutamic acid, fructose 1,6-BP, UDP-N-acetylglucosamine, glucose 6-phosphate, L-aspartic acid, 6-phospho-gluconic acid, dihydroxyacetone phosphate, and pyruvic acid were classified as weak Mg2+ binders, with a KD value greater than 2 mM (Table 1). Three metabolites, glutathione, L-valine, and L-glutamine had negligible Mg2+ binding properties, as measured with HQS (SI figure 2). In an effort to understand the the effects of Eco80 on RNA mechanistically, we created two sub-artificial cytoplasm: NTP chelated Mg2+ (NTPCM), and weak metabolite chelated Mg2+ (WMCM), composed of the strong Mg2+ chelators (NTPs), and weak Mg2+ chelators, respectively (Table 1).

We used two methods to estimate how Eco80 metabolites affect the speciation of free and chelated Mg2+ as a mixture. The first method was the same HQS assay that we used to estimate binding constants for metabolites that did not produce enough heat on Mg2+ binding to measure with ITC, based on calculating the free Mg2+ concentration in the presence of metabolites using HQS fluorescence emission (Figure 1B-C, SI Table 3). The second method to determine Mg2+ speciation was a statistical model that accounts for experimental uncertainty in metabolite concentrations and uncertainty in KD determination, and estimates Mg2+ speciation assumine single-site binding (meaning that one metabolite associates one Mg2+). This model is described in detail in the Supplementary Methods. Briefly, concentration errors were propagated from uncertainties in reagent masses and volumes used during sample preparation, and KD uncertainties were obtained from the fits (Table 1). Both uncertainties were then randomly seeded into Equation 1 to create 1000 virtual artificial cytoplasms, where [Mg]T is the total Mg2+ concentration, [Mg] is the free Mg2+ concentration, i is an integer representing each metabolite in a mixture, N is the total number of metabolites in a mixture, [Li]T is the concentration of the “*i’th*” metabolite in a mixture, and KD is the disassociation constant.

Then, equation 1 is solved numerically to determine the free Mg2+ concentration produced at a given total Mg2+ concentration in a virtual artificial cytoplasm.

The two methods indicate that Mg2+ speciates in artificial cytoplasms according to a single-site model within or below the biological free Mg2+ range of 0.5 to 3 mM Mg2+. However, Mg2+ does not speciate according to a single-site model at higher free Mg2+ concentrations (Figure 1 E-F). For example, in Eco80, the statistical model suggests that the metabolites should buffer the free Mg2+ concentration in the biological Mg2+ range, where a 20 mM increase in the total Mg2+ from 20 to 40 mM leads to a 2.5 mM increase in free Mg2+ from 0.5 to 3 mM (Figure 1E, hex bins). Free Mg2+ concentrations measured in this range with HQS emission are consistent with this single-site behavior (Figure 1E, data points). At higher free Mg2+ concentrations, Eco80 should loose its free Mg2+ buffering capacity as chelators become saturated, and the free Mg2+ should increase with the total Mg2+ (Figure 1E, hex bins). However, the free Mg2+ concentration measured with HQS does not increase as fast as the statistical model predicts above 3 mM free Mg2+ (Figure 1E, data points). For example, measure free Mg2+ in Eco80 is expected to increase from 3 mM to ~100 mM as the total Mg2+ concentration is increased from 40 mM to 200 mM (Figure 1E, hex bins). However, the the free Mg2+ concentration measured with HQS only increases from 3 mM to ~10 mM (Figure 1E, data points). One possibility is that multivalent interactions, which the statistical model does not account for, where several Mg2+-saturated metabolites interact with additional Mg2+ molecules, dominate the equilibrium. Such non-single-site behavior above 3 mM free Mg2+ is also observed in the NTPCM and WMCM artificial cytoplasms (Figure 1 F & G), and was observed previously (Ryota Nat Com.).

Lastly, we sought to empirically determine how much total Mg2+ is required to attain a free Mg2+ concentration of 2 mM in Eco80, NTPCM, and WMCM. The relationship between the free Mg2+ calculated from HQS emission and the total Mg2+ concentration in each artificial cytoplasms was fit to a polynomial to empirically approximate the data (Figure 1 E-G, colored lines), and the total Mg2+ concentration required to produce 2 mM Free Mg2+ was calculated from the polynomial fit (see methods for details). This resulted in predicted 31.6, 25.0, and 6.5 mM total Mg2+ concentration to produce 2 mM free Mg2+ in Eco80, NTPCM, and WMCM, respectively (Table 2).

**Thermodynamic analysis of RNA helices in Eco80**

We sought to understand how Eco80 affects the thermodynamic stability of RNA. We began on RNA helices composed of Watson-Crick base pairs. RNA helix stability is traditionally measured with UV-absorbance melting curves, typically monitored at 260 or 280 nm. However, such absorbance melting curves were not appropriate for measuring helix stability in Eco80 because of the high absorbtivity of ATP, UTP, GTP, dTTP, and UDP-N-acetylglucosamine, which are major components of Eco80. Thus, we designed a fluorescence binding isotherm assay, which is optically orthogonal to Eco80.

Helix stability was measured using the emission of a 5’-fluorophore labeled RNA strand (FAM-RNA) in equilibrium with a complementary 3’-quencher labeled RNA strand (RNA-BHQ1), (Figure 2A). High emission indicates that the FAM-RNA is single-stranded, while low emission indicates that the FAM-RNA is bound in duplex with a RNA-BHQ1 strand. We used a binding isotherm method, where RNA-BHQ1 is titrated into a constant concentration of FAM-RNA at a single temperature (SI figure 3), resulting in an apparent binding isotherm (Figure 2B). Fluorescence binding isotherms were favored over fluorescence melts because of the unpredictable dependence of FAM emission on temperature. Emission of FAM was monitored in a Real-Time PCR instrument at different temperatures, resulting in a isotherm every 0.5 °C from 20 to 80 °C (Figure 2B).

Raw fluorescence was fit with a new program called MeltR, created by the authors, to determine folding energies. MeltR is a package of functions in the popular R programming language, that allows facile conversion of raw data to folding energies (see methods for details). Importantly, MeltR handles two sources of experimental error, uncertainty in RNA concentration and inaccurate KDs collected at low and high temperatures. MeltR handles uncertainties in RNA concentration by using a Job plot obtained from fluorescence isotherms at low temperatures. Likewise, MeltR allows the user to select isotherms that provide the most accurate KDs, which are then passed to the non-linear regression algorithms that determine helix folding energies. MeltR then calculates folding energies using two Van’t Hoff methods, directly fitting a Van’t Hoff plot (Figure 3C) and globally fitting raw fluorescence emission to SI equation X. Errors in the main text are reported 1.5% in terms of the ΔG37°C and a detailed error analysis is in the methods.

**Eco80 destabilizes RNA helices**

We used fluorescence binding isotherms to determine helix folding energies in background monovalent metal ion control (240 mM NaCl 140 mM KCl), Eco80, NTPCM, and WMCM, for a data set of five representative eight base-pair RNA helices. The helix set contains representatives of all 10 Watson-Crick nearest neighbor parameters and vary in AU content from 25% to 75% and thus represent a good range of helical features. All solutions contain 2 mM Mg2+, as per Table 2. We used both methods to determine folding energies in MeltR, where method 1 was fitting the Van’t Hoff relationship between the disassociation constant and temperature and method 2 was directly globally fitting the raw fluorescence binding isotherms (SI table 4). Results are summarized in Table 3.

All five representative helices were significantly destabilized in Eco80, meaning the the ΔΔG37°C between the background monovalent condition and Eco80 was larger than the propagated uncertainty in ΔΔG37°C, by 0.44 to 1.12 kcal/mol (Table 3, Figure 2D). We did not observe a clear relationship between AU content and destabilization. For example, Helix 2 has a relatively low AU content of 38% and was strongly destabilized by 1.12 kcal/mol in Eco80. However, helix 1 and 5. the sequences with the lowest (25%) and highest (75%) AU content were destabilized by the same amount, 0.60 and 0.61 kcal/mol, respectively. Likewise, helices 3 and 4 both have an AU content of 63%, but they were destabilized by different amounts, 0.44 and 0.64 kcal/mol respectively. Thus, Eco80 destabilizes RNA helices but the underlying sequence dependence is unclear.

To better understand how components of Eco80 destabilize RNA helices, we then analyzed the effects of strong Mg2+ chelating metabolites and weak Mg2+ chelating metabolites, separately. NTPCM, which is composed of strong Mg2+ chelating metabolites, consistently destabilized RNA helices (Figure 2D). Helices 2, 3, 4, and 5 were significantly destabilized by 0.41 to 0.60 kcal/mol (Table 3). NTPCM did not significantly destabilize Helix 1, however, the the 0.32 kcal/mol destabilization effect is consistent with the other 5 helices. In fact, as the AU content of helices increases the destabilization increases steadily from 0.32 kcal/mol at a 25% to 0.60 kcal/mol at 75% with a slope of 5.5 cal/mol/percent AU content (R2 = 0.99, SI figure 6).

In contrast to NTPCM, WMCM, which is composed of strong Mg2+ chelating metabolites, destabilized, had no effect, or stabilized RNA helices depending on the helix identity (Figure 2D). Similar to Eco80, the sequence dependence of stabilization or destabilization is not clear. No significant change in stability was observed for helices 3 and 4 in WMCM and both have a AU content of 63%. Helix 1 and 5, with the lowest and highest AU content of 25% and 75% respectively, were destabilized by 0.40 and 0.82 kcal/mol in WMCM. In contrast, Helix 2, with an intermediate AU content of 38% was stabilized in WMCM.

Thus, the net effect of Eco80 on RNA helices is destabilization, with destabilizing interactions dominating for strong Mg2+ chelating metabolites, and a mixture of stabilizing and destabilizing interactions for weak Mg2+ chelating metabolites.

**Eco80 protects RNA from degradation**

Several studies indicate that weak and strong Mg2+ chelating metabolites reduce Mg2+ mediated RNA degradation and we hypothesized that Eco80 could also protect RNA from degradation. We used an in-line degradation (in-line probing) assay to determine if Eco80 stabilizes the chemical structure of RNA. In-line probing takes advantage of the natural susceptibility of the RNA phosphodiester backbone to cleavage. In this mechanism, the 2’-hydroxyl is deprotonated by a hydroxide anion (often facilitated by an Mg2+), and the 2’-hydroxyl serves as a nucleophile to replace the 5’-hydroxyl of the next nucleotide, in an SN2-like mechanism (Figure 3A). Unstructured nucleotides are more susceptible to cleavage because they are more likely to adopt an in-line conformation that favors cleavage. For this assay, 5’-P32 RNA are incubated at 37 °C for about 90 hours to facilitate in-line cleavage. RNA fragments are then fractional on a denaturing PAGE gel, providing single nucleotide resolution of cleavage rates, measured by the increase in counts with time for a given band (SI figure 7).

We first determined in-line degradation rates for the guanine riboswitch aptamer (Figure 3B). The guanine riboswitch aptamer is a highly structured RNA aptamer that binds the guanine nucleobase, and has been studied extensively, thus providing structural and mechanistic information. Moreover, the apo (guanine ligand un-bound) state and the guanine ligand bound state are similar, showing structural changes only in the nucleotides that directly mediate the guanine binding site.

Degradation rates for nucleotides 29 to 63 in Eco80, NTPCM, and WMCM with enough total Mg2+ to maintain 2 mM free Mg2+ were compared to degradation rates in a 2 mM free Mg2+ and 25 mM free Mg2+. The 25 mM free Mg2+ condition was chosen because it is a common free Mg2+ condition *in-vivo* and is similar to the 25 and 31.6 total Mg2+ condition used for NTPCM and Eco80, respectively (Table 2). Degradation fragments greater than 63 could not be resolved on the gel and degradation fragments shorter than 29 were smeared out because of the high concentrations of negatively charged ions in the samples (SI figure 6). Overall, the guanine riboswitch aptamer exhibited similar degradation patterns between the 2 mM free, Eco80, NTPCM, and WMCM conditions, with high degradation in the 5’-region of the P2 stem and high reactivity in the P3-stem loop region. This indicates that the guanine riboswitch aptamer adopts a similar structure in these conditions (Figure 3 C, SI figure 8). The 25 mM free condition, which exhibited higher degradation rates in the P2-P3 junction (Figure 3 C). This pattern is similar to ILP data published for another guanine riboswitch aptamer, but in a high Mg2+ condition and pH 8.0, indicating that the increase in degradation rates in the 25 mM free Mg2+ condition is dependence on the presence of Mg2+-OH- complexes (SI figure 8). To confirm that the guanine aptamer adopts similar structures between all conditions, we collected small angle X-ray scattering (SAXS) data in every artificial cytoplasm. Raw scattering overlay nicely, indicating that the structure of the guanine aptamer is similar inbetween conditions (SI figure 9A). P(r) plots, where the maximum is the radius of gyration, indicate that the radius of gyration is similar between solution conditions (SI figure 9B). Lastly, electron density reconstructions are consistent with the known structure of the guanine riboswitch aptamer in every condition (Figure 3D, Si Figure 9C-E). Thus, decreased degradation rates in the 2 mM free, Eco80, NTPCM, and WMCM conditions in comparison to the 25 mM free condition is likely due to a reduction in the availability Mg2+-OH- complexes caused by chelation of Mg2+ by metabolites rather than changes in RNA structure.

We sought to better characterize the structural dependence of guanine aptamer degradation in different conditions. We thus examined the crystal structure for the guanine riboswitch aptamer and classified each residue as single stranded, meaning that the base was not participating in hydrogen bonding interactions with other residues, non-canonical, meaning that the base was forming non-canonical hydrogen bonding interactions with other residues, Watson-Crick (WC), meaning that the base was in a traditional Watson-Crick base pair with another residues, Watson-Crick + Non-canonical (WC + NC), meaning that the base was forming a traditional Watson-Crick base pair with another residues and participating in a non-canonical contact.

We observed decreased degradation rates at the single stranded nucleotide U39 in 2 mM free Mg2+, Eco80, NTPCM, and WMCM, in comparison to the 25 mM free Mg2+ condition. Likewise, we observed an overall decrease in reactivity in 2 mM free Mg2+, Eco80, NTPCM, and WMCM in comparison to the 25 mM free Mg2+ condition. In contrast, degradation rates for nucleotides participating in Watson-Crick base pairing interactions were independent across solution conditions. Interestingly, the same overall decrease in degradation rates in artificial cytoplasm, in comparison to the 25 mM free Mg2+ condition, was observed for nucleotides that participate in Watson-Crick base pairs and participate in a second hydrogen bonding interaction with the base of another residue. Thus, we observed a trend of protection from Mg2+ mediated degradation in artificial cytoplasm composed of Mg2+ chelating metabolites in comparison to the 25 mM free Mg2+ condition.

We repeated our in-line degradation assay with two other highly studied biological RNA, the cleaved- CPEB3 ribozyme and yeast tRNAphe, to confirm that a reduction of Mg2+ mediated degradation in artificial cytoplasm was a broadly applicable trend (SI figure 7, \* & 9). For the cleaved-CPEB3 ribozyme, degradation rates at singles stranded residues were reduced in Eco80 and WMCM in comparison to the 25 mM free Mg2+ condition (Figure 3F). The degradation rates for a residue that is predicted to participate in a non-canonical tertiary contact was reduced in Eco80 but not WMCM, in comparison to the 25 mM free Mg2+ condition. Degradation rates were constant for nucleotides participating in Watson-Crick base-pairs for the cleaved-CPEB3 ribozyme, between the 2 mM free Mg, the Eco80, and the 25 mM condition Mg conditions, like the guanine riboswitch aptamer. Interestingly, Degradation rates in the NTPCM condition were reduced for all structure classifications compared to all other conditions. Yeast tRNAphe exhibits almost no in-line degradation except for the single stranded nucleotides in the P3 stem loop, termed the anticodon loop (SI figure 9D), consistent with published data. Single stranded nucleotides, mostly from the anticodon loop showed reduced degradation rates in the in Eco80, NTPCM and WMC in comparison to the 25 mM free condition, like the guanine riboswitch aptamer and the CPEB3 ribozyme. Likewise, degradation rates were constant across conditions for nucleotides that form Watson-Crick base pairs, similar to the guanine riboswitch aptamer and the CPEB3 ribozyme. However, degradation rates were constant across conditions for nucleotides that form non-canonical contacts, which is different than the increased degradation observed for nucleotides that form non-canonical base pairs in the guanine riboswitch aptamer and the CPEB3 ribozyme. One possible explanation is that the tertiary structure of tRNAphe is less dynamic than the tertiary structure of the guanine riboswitch aptamer and the CPEB3 ribozyme.

Thus, overall, the in-line degradation assay indicated that Eco80, NTPCM, and WMCM protect RNA from Mg2+ mediated degradation, even though all three artificial cytoplasm’s have relatively high concentrations of Mg2+. This protective effect is enhanced in unstructured regions where RNA is most susceptible to Mg2+ mediated degradation.

**The weak Mg2+ chelating metabolites in Eco80 promote RNA catalysis**

Weak metabolite chelated Mg2+ is known to promote catalysis by ribozymes. For example, the CPEB3 ribozyme catalysis is enhanced by about 1.6-fold by 2 mM free Mg2+ in solution with an estimated 11.3 mM free glutamate chelated Mg2+, in comparison to catalysis in 2 mM free Mg2+ alone. Thus we also hypothesized that Eco80 metabolites would also promote CPEB3 catalysis.

We measured CPEB3 ribozyme cleavage rates in 2 mM free Mg2+, a total 25 mM free Mg2+, and Eco80, NTPCM, and WMCM containing enough total Mg2+ to produce 2 mM free Mg2+ (Table 2). Briefly, we purified full length CPEB3 ribozyme (Figure 4A), incubated CPEB3 in artificial cytoplasm’s, fractionated cleaved and un-cleaved CPEB3 from time points on a denaturing acrylamide gel, and calculated the fraction cleaved from the relative intensity of cleaved and un-cleaved RNA bands (SI Figure 10). Fraction cleaved as a function of time was fit to a single exponential equation to estimate the reaction rate (Figure 4B). We performed 4 reactions per condition.

CPEB3 ribozyme catalysis was reduced in all conditions in comparison to the 25 mM free Mg2+ control (Figure 4C). Surprisingly, CPEB3 catalysis was one half as fast in Eco80 in comparison to the 2 mM free Mg2+ control, despite the 31.6 mM total Mg2+ in Eco80. Likewise, CPEB3 catalysis was one third as fast in NTPCM in comparison to the 2 mM free Mg2+ control, an even stronger effect than Eco80. In contrast, CPEB3 catalysis was enhanced by 1.33 fold in WMCM, similar to the enhancement observed for glutamate chelated Mg2+.In summary, the artificial cytoplasm composed of weak Mg2+ chelating metabolites with 2 mM free Mg2+ enhanced CPEB3 catalysis, the artificial cytoplasm composed of strong Mg2+ chelating metabolites 2 mM free Mg2+ reduced CPEB3 catalysis, and the mixture of both types of metabolites had an intermediate effect.

**Discussion**

Making Eco80 artificial cytoplasm

-Complex but manageable

Free Mg2+ buffering in the cell

-Has been observed by Email Phil for Papers

Effects on helices

-Mg stabilizes (from lit)

-Polor osmolytes destabilize (from lit)

-Allison’s paper

-Here we analyze their effects together and see opposing effects

Propose model that summaries effects