ABSTRACT: Herein, we examine the complicated network of interactions among RNA, the metabolome, and the metalome in conditions that mimic the *E. coli* cytoplasm. First, 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. Then, we used this information to inform creation 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 sit binding behavior towards Mg2+ in the biologically relevant free Mg2+ range of ~0.5 to 3 mM Mg2+, using a Mg2+ binding fluorescent dye (8-Hydroxy-5-quinolinesulfonic acid). Furthermore, we examined the effects of Eco80 conditions on the thermodynamic stability, chemical stability, catalysis, and compactness of RNA. We find that these Eco80conditions lead to opposing effects, wherein thermodynamic stability of RNA helices were weakened but chemical stability, compactness, and catalysis were enhanced. We propose a mechanism where increased RNA compactness and catalysis is facilitated in Eco80.

**Introduction**

Summary of progress on in vivo-like conditions.

Studies that consider cellular components one at a time.

Studies that consider cellular components together in artificial cytoplasm.

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 relavent 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 the network of interactions that RNA experiences in *E. coli* cells.

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

*E. coli* cells contain hundreds of metabolites (about 243 mM total), which is too many metabolites to test systematically. However, 15 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).

All of the metabolites in Eco80 are zwitter ions or negatively charged near physiological pH (~7) and require electrostatic neutralization with metal ions. Thus, we prepared Eco80 so that the final monovalent ion concentration was the physiological value of 240 mM Na+ and 140 mM K+ (Supplementary information (SI) Table 1). Metabolite salts and free acids were prepared to a final 2xconcentration, and the amount of Na+ and K+ added with each metabolite was recorded. 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 was added to a final 240 mM Na+ and 140 mM K+. We thus created Eco80, at a 2x final concentration so that it could be diluted into other reagents for experiments.

Next, we considered how metabolites effect the speciation of free and chelated Mg2+. All 15 Eco80 metabolites have functional groups, carboxylates and phosphates, that drive chelating interactions with divalent 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 putative as Mg2+ chelation strength is dependent on environmental factors such as pH, ionic strength, the composition of background ions, and temperature. Thus, we sought to better characterize Mg2+ chelation by the metabolites inEco80,at the physiological background of 240 mM Na+, 140 mM K+, pH 7.0, and 37 °C.

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 determine 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+ metabolites that did not produce enough heat on binding 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 2B, bottom). Mg2+ binding by metabolites is thus observed 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 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,* 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). We thus broke Eco80 down into two other artificial cytoplasms, NTP chelated Mg2+ (NTPCM), and weak metabolite chelated Mg2+ (WMCM), composed of the strong Mg2+ chelators (nucleotide triphosphates), and weak Mg2+ chelators, respectively (Table 1).

We next used two methods to estimate how the metabolites effect the speciation of free and chelated Mg2+ as a mixture. The first method was experimental, using HQS emission to estimate the free Mg2+ concentration in the presence of metabolites (Figure 1B-C, SI Table 3). The second method was a statistical model that accounts for experimental uncertainties in metabolite concentrations and uncertainty in KD determination, based on single-site binding (meaning that one metabolite associates one Mg2+). The statistical 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 a 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+, but not at higher free Mg2+ concentrations (Figure 1 E-F). For example, in Eco80, the statistical model suggests that the free Mg2+ should increase slowly as the total Mg2+ concentration is increased, until the strong Mg2+ chelators (NTPs) become saturated at about 27 mM total Mg2+ (Figure 1E, hex bins). At total Mg2+ concentrations higher than 27 mM, the free Mg2+ should increase faster because the NTPs are saturated by Mg2+ and the weak chelators sequester less Mg2+. Free Mg2+, calculated using HQS emission shows a similar trend to the statistical model below 3 mM Mg2+ free (Figure 1E, data points). However , the free Mg2+ concentration calculated from HQS emission does not increase with the total Mg2+ as fast as the single-site model would predict above 3 mM free Mg2+, indicating that multivalent interactions, where one metabolite interacts with several Mg2+ molecules, dominate the equilibrium. Non-single-site behavior above 3 mM free Mg2+ is also observed in the NTPCM and WMCM artificial cytoplasms (Figure 1 F & G).

Lastly, we sought to empirically determine how much total Mg2+ is required to maintain 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 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. This resulted in a 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).

**Eco80 destabilizes RNA helices**

*Paragraph 1: Transition*

We sought to understand how Eco80 effects the stability of RNA helices composed of Watson-Crick base pairs.

Could not use the traditional method for thermodynamics, absorbance melting curves because of the high absorbtivity of ATP, UTP, GTP, dTTP, and UDP-N-acetylglucosamine.

Used fluorescence binding isotherms because the method is orthogonal to the optical properties of metabolites.

*Paragraph 2: Fluorescence isotherms-intro*

Described layout of assay, reagent conc, temperature control, expected results (Figure 2A)

Describe dependence of temperature and fit (Figure 2B)

Describe how the Van’t hoff plot can be used to extract thermodynamic parameters

*Paragraph 3: MeltR*

Raw fluorecence was fit with a new program called MeltR to determine thermodynamic parameters

MeltR handles two sources of experimental error that can effect results: (1) uncertainties in RNA concentration determination and (2) inacurate KDs collected at low and high temperatures.

*Paragraph 4: MeltR concentration optimization algorithm*

Fit quality is highly dependent in the determination of RNA concentration, which is uncertain because of diextinction coefficient uncertainty, but errors are propagated systematically (SI Figure 3)

For example consider modeled data, assuming perfect data (SI Figure 4 A, left panel)

Now consider modeled data with +20% seeded error (SI Figure 4 A, middle panel)

Use low temperature isotherm to determine the correct R.

Will find the accurate dG even with 50% error in concentration determination.

*Paragraph 5: MeltR only fits data from the most accurate isotherms*

sffkpksdfpksd[pfk’[sdlk

**Eco80 protects RNA from degradation**