

# The Metabolome Weakens RNA Helix Stability and Increases RNA Chemical Stability

Jacob P. Sieg<sup>1,2</sup>, Lauren McKinley<sup>1,2</sup>, Melanie Huot<sup>3,4</sup>, Neela H. Yennawar<sup>5</sup>, and Philip C. Bevilacqua<sup>\*1,2,3</sup>

<sup>1</sup>Department of Chemistry, Pennsylvania State University, University Park, PA 16802.

<sup>2</sup>Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802.

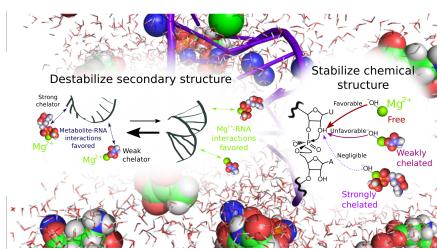
<sup>3</sup>Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802.

<sup>4</sup>Department of Biology, Pennsylvania State University, University Park, PA 16802.

<sup>5</sup>The Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA 16802.

**Keywords** Magnesium ion, Metabolites, Chelated magnesium, RNA folding, RNA function, Near-cellular condition

Herein, we examine the complex network of interactions among RNA, the metabolome, and divalent Mg<sup>2+</sup> in conditions that mimic the *E. coli* cytoplasm. We determined Mg<sup>2+</sup> binding constants for the top 15 *E. coli* metabolites, comprising 80% of the total metabolome, at physiological pH and monovalent ion concentrations. These data were used to inform the development of an artificial cytoplasm that mimics *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 Mg<sup>2+</sup> in the biologically relevant free Mg<sup>2+</sup> range of ~0.5 to 3 mM Mg<sup>2+</sup>, using a Mg<sup>2+</sup>-sensitive fluorescent dye. Effects of Eco80 conditions on the thermodynamic stability, chemical stability, structure, and catalysis of RNA were examined. We found that Eco80 conditions lead to opposing effects on the thermodynamics and chemical stability of RNA. In particular, the thermodynamic stability of RNA helices was weakened while the chemical stability and catalysis of RNA were enhanced, which can be understood using the speciation of Mg<sup>2+</sup> between weak and strong Mg<sup>2+</sup>-metabolite complexes in Eco80. Overall, the effects of Eco80 reflects RNA function *in vivo* and enhances the biological relevance of mechanistic studies of RNA *in-vitro*.



## Introduction

RNA serves as the conduit of genetic information in the Central Dogma of Molecular Biology and performs numerous functions in biology owing to the capacity of RNA to form complicated, diverse, and functional structures.<sup>1</sup> The development of genome wide structure-probing techniques *in vivo* has provided insight into RNA structure and function in cells.<sup>2-4</sup> However, most experimental techniques that provide insight into the mechanism and function of RNA cannot be readily performed in a cell, and are typically limited to simple conditions, usually 100 to 500 mM monovalent metal ions and 0.5 to 50 mM free divalent magnesium ions (Mg<sup>2+</sup>) with a dilute buffer.<sup>5</sup> *In vitro* studies of RNA in conditions that mimic a cell, so called *in vivo*-like conditions, provide a link between experiments that probe RNA structure *in vivo* and experiments that provide biologically-relevant insight *in vitro*.<sup>5</sup>

Many studies have investigated the effects of individual components of the cellular environment on nucleic acid structure, including small molecules and abiological crowders. Studies that use small molecules that are similar to metabolites indicate that these species interact strongly with the unfolded state of nucleic acids and destabilize secondary structure.<sup>1-5</sup> Studies that simulate the cellular macromolecules reveal stabilized RNA tertiary structures, increased folding cooperativity, and improved RNA function in crowded environments.<sup>6-10</sup> Thermodynamic

characterization of RNA helix formation in crowding conditions indicates that crowders destabilize helices.<sup>11-13</sup> In summary, using simple models to simulate the cellular environment has provided valuable insight into how the cell affects nucleic acids and motivated investigation of more complex and realistic artificial cytoplasms.<sup>14</sup>

A number of studies have provided mechanistic insight into proteins in complex environments,<sup>15</sup> ranging from cell lysates to live cells.<sup>16-19</sup> However, researchers sacrifice control over the environment that is provided by a simple system. Mechanistic studies of RNA in the cells or lysates have two additional problems. The first is the propensity of cells to degrade foreign RNA.<sup>19,20</sup> The second is the lack of control over Mg<sup>2+</sup> speciation between free and chelated Mg<sup>2+</sup>.

Control over Mg<sup>2+</sup> speciation is crucial for mechanistic studies of RNA because of the sensitivity of RNA folding to the concentration of Mg<sup>2+</sup> in the solution, as demonstrated by thousands of studies.<sup>21</sup> Furthermore, recent studies have demonstrated the importance of weak metabolite-Mg<sup>2+</sup> complexes to RNA function.<sup>22-24</sup> These studies consider effects of mixtures of one to three metabolites, which is a step forward, but still far from the true complexity of the cellular environment. In addition, Mg<sup>2+</sup> speciation was approximated in these studies, as Mg<sup>2+</sup> speciation was estimated assuming single-site-binding, meaning that one metabolite interacts with one Mg<sup>2+</sup> ion, and binding constants were

extrapolated from published sources, reported at disparate ionic compositions and pHs.<sup>21</sup>

Herein, 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 Mg<sup>2+</sup> ions. We start by compiling metabolite concentrations in *E. coli*, simplify to the 15 most abundant metabolites, determine metabolite-Mg<sup>2+</sup> binding constants at the relevant pH and ionic strength, and lastly determine the total Mg<sup>2+</sup> concentration in the final mixture of metabolites. This *aufbau* approach allows us to study the effects of the metabolite and metal ion species that comprise a major part of the interactions that RNA experiences in *E. coli* cells.

**Table 1. Eco80: The 15 most abundant metabolites, which comprise 80% of the *E. coli* metabolome.**

| Metabolite                 | Conc. (mM) <sup>a</sup> | K <sub>D</sub> (mM) <sup>b</sup> | Chelation strength <sup>c</sup> |
|----------------------------|-------------------------|----------------------------------|---------------------------------|
| ATP                        | 9.63<br>(0.963)         | 0.28<br>(0.01) <sup>b</sup>      | NTPCM<br>(Strong)               |
| UTP                        | 8.29<br>(0.829)         | 0.248<br>(0.004) <sup>b</sup>    | NTPCM<br>(Strong)               |
| GTP                        | 4.87<br>(0.487)         | 0.201<br>(0.007) <sup>b</sup>    | NTPCM<br>(Strong)               |
| dTTP                       | 4.62<br>(0.462)         | 0.160<br>(0.003) <sup>b</sup>    | NTPCM<br>(Strong)               |
| L-Glutamic acid            | 96<br>(9.6)             | 520<br>(50) <sup>c</sup>         | WMCM<br>(Weak)                  |
| Glutathione                | 16.6<br>(1.66)          | NA <sup>d</sup>                  | WMCM<br>(Weak)                  |
| Fructose 1,6-bisphosphate  | 15.2<br>(1.52)          | 5.9<br>(0.1) <sup>b</sup>        | WMCM<br>(Weak)                  |
| UDP-N-acetylglucosamine    | 9.24<br>(0.924)         | 29<br>(2) <sup>b</sup>           | WMCM<br>(Weak)                  |
| Glucose 6-phosphate        | 7.88<br>(0.788)         | 17.3<br>(0.2) <sup>b</sup>       | WMCM<br>(Weak)                  |
| L-Aspartic acid            | 4.23<br>(0.423)         | 465<br>(12) <sup>c</sup>         | WMCM<br>(Weak)                  |
| L-Valine                   | 4.02<br>(0.402)         | NA <sup>d</sup>                  | WMCM<br>(Weak)                  |
| L-Glutamine                | 3.81<br>(0.381)         | NA <sup>d</sup>                  | WMCM<br>(Weak)                  |
| 6-Phosphogluconic acid     | 3.77<br>(0.377)         | 14.4<br>(0.2) <sup>b</sup>       | WMCM<br>(Weak)                  |
| Pyruvic acid               | 3.66<br>(0.366)         | 15.8<br>(0.9) <sup>c</sup>       | WMCM<br>(Weak)                  |
| Dihydroxyacetone phosphate | 3.06<br>(0.306)         | 20<br>(1) <sup>b</sup>           | WMCM<br>(Weak)                  |

<sup>a</sup>Uncertainty propagated from uncertainties in reagent masses and volumes used during sample preparation. Extra significant digits included to avoid systematic rounding errors in the statistical model. <sup>b</sup>Determined at 37 °C with ITC as measured in SI figure 1 and SI table 2. Error is the propagated standard error in the fit parameters. <sup>c</sup>Determined at 37 °C with HQS emission as measured in the SI figure 2 and SI table 3. Error is the propagated standard error in the fit parameters. <sup>d</sup>No binding observed as per SI figure 2. <sup>e</sup>Metabolites with K<sub>D</sub>s for Mg<sup>2+</sup> less than 2 mM are considered strong Mg<sup>2+</sup> chelators and K<sub>D</sub>s greater than 2 mM are considered weak Mg<sup>2+</sup> chelators. NTPCM and WMCM are sub-artificial cytoplasmas composing Eco80,

nucleotide triphosphate-chelated Mg<sup>2+</sup> and weak metabolite-chelated Mg<sup>2+</sup>, respectively.

## Results

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

*E. coli* cells contain hundreds of metabolites (~240 mM total),<sup>25</sup> which is too many to test systematically. However, the 15 most abundant metabolites in *E. coli*, an experimentally manageable number, comprise 80% (195 mM) of the total metabolites (Figure 1A). We thus sought to prepare Eco80, an artificial cytoplasm containing biological concentrations of the 15 most abundant metabolites in *E. coli* (Table 1).

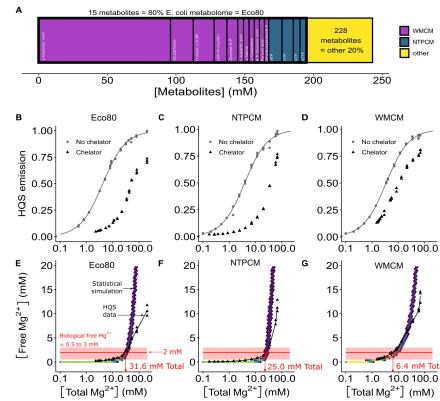
Eco80 was prepared at a 2x concentration so that it could be diluted into other reagents and contain physiological concentrations of monovalent metal ions at pH 7.0 (see supplementary information (SI) table 1 for details). Briefly, all metabolites in Eco80 are zwitterions or negatively charged near physiological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

Next, we considered how metabolites affect the speciation of free and chelated Mg<sup>2+</sup>. All 15 Eco80 metabolites have functional groups, carboxylates and phosphates, that drive chelating interactions with Mg<sup>2+</sup> ions (Table 1).<sup>21</sup> We sought to precisely characterize Mg<sup>2+</sup> chelation by the metabolites in Eco80, at the physiological background, since Mg<sup>2+</sup> binding affinity is dependent on environmental factors such as pH, composition of background ions, and temperature.<sup>26-32</sup>

We determined apparent dissociation constants (K<sub>D</sub>) for Eco80 metabolites in a background of 240 mM NaCl, 140 mM KCl, pH 7.0 buffer at 37 °C (Table 1). Isothermal titration calorimetry (ITC) was used to measure the K<sub>D</sub>s for phosphorylated metabolites (SI figure 1, SI table 2). A fluorescence assay, which measures the free Mg<sup>2+</sup> concentration in a sample using the metal ion-binding dye 8-Hydroxy-5-quinolinesulfonic (HQS) acid,<sup>33</sup> was used to estimate the K<sub>D</sub>s for metabolites that did not produce enough heat on binding Mg<sup>2+</sup> to measure with ITC (SI figure 2, SI table 3). For this assay, Mg<sup>2+</sup> is titrated into HQS solutions in the absence and presence of Mg<sup>2+</sup> chelators. Emission of HQS as a function of the total Mg<sup>2+</sup> in the absence of chelators is then fit to a binding model for the binding of Mg<sup>2+</sup> to HQS (SI figure 2A, top black data and fit). The free Mg<sup>2+</sup> concentration, which is equal to the total Mg<sup>2+</sup> concentration in the absence of chelator, is then associated with the fluorescence emission for each data point using the binding model. This process is repeated in the presence of a chelator, using the no-chelator data to obtain the free Mg<sup>2+</sup> concentration at any total concentration of Mg<sup>2+</sup> (SI figure 2A, bottom). Note that free and total Mg<sup>2+</sup> concentrations are the same, y=x, in the absence of chelators, and that the data are right-shifted in the presence of chelators. The affinity of Mg<sup>2+</sup> binding by

metabolites is thus obtained by fitting the free  $Mg^{2+}$  concentration as a function of the total  $Mg^{2+}$  concentration.

The binding affinity for Eco80 metabolites and  $Mg^{2+}$  ranged from strong to negligible. The four nucleotide triphosphates, ATP, UTP, GTP, and dTTP, were classified as strong  $Mg^{2+}$  binders, with  $K_D$  values ranging from 0.160 to 0.28 mM, less than the approximate free  $Mg^{2+}$  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-glutamic acid pyruvic acid, and dihydroxyacetone phosphate--were classified as weak  $Mg^{2+}$  binders with  $K_D$  values greater than 2 mM (Table 1). Three metabolites--glutathione, L-valine, and L-glutamine--had negligible  $Mg^{2+}$  binding properties, as measured with HQS (SI figure 2). In an effort to understand the effects of Eco80 on RNA mechanistically, we created two sub-artificial cytoplasms: NTP-chelated  $Mg^{2+}$  (NTPCM) and weak metabolite-chelated  $Mg^{2+}$  (WMCM), comprised of the strong  $Mg^{2+}$  chelators (NTPs) and weak  $Mg^{2+}$  chelators, respectively (Table 1).



**Figure 1** Analysis of  $Mg^{2+}$  speciation in *E. coli* metabolite mixtures. **(A)** *E. coli* metabolome molar composition. Eco80 contains the 15 most abundant metabolites that comprise 80% of the *E. coli* metabolome. NTPCM contains four strong  $Mg^{2+}$  chelating NTPs, and WMCM contains 11 other weak  $Mg^{2+}$  binding metabolites. **(B-D)** Effect of  $Mg^{2+}$  on HQS emission with and without mixtures of metabolites that chelate  $Mg^{2+}$ . Grey lines represent fits to determine the binding constant for  $Mg^{2+}$  and HQS. **(E-G)** Relationship between the total  $Mg^{2+}$  concentration and the free  $Mg^{2+}$  concentration with mixtures of metabolites that chelate  $Mg^{2+}$ . Hex bins represent a simulated artificial cytoplasm assuming single-site binding. Triangle data points are free  $Mg^{2+}$  concentrations calculated using HQS emission. Black lines were generated using polynomial regression. The red shaded region is the biological free  $Mg^{2+}$  range of 0.5 to 3 mM. The red line is the approximate free  $Mg^{2+}$  concentration in *E. coli* of 2 mM. Red arrows represent the total  $Mg^{2+}$  required to maintain 2 mM free  $Mg^{2+}$ .

We used two methods to estimate how Eco80 metabolites affect the speciation of free and chelated  $Mg^{2+}$ . The first method was the HQS assay that we used to estimate binding constants for metabolites, based on calculating the free  $Mg^{2+}$

concentration in the presence of metabolites using HQS fluorescence emission (Figure 1B-D, SI table 4). This method directly determined free  $Mg^{2+}$  but did not report on speciation of  $Mg^{2+}$  to different metabolites. The second method used a statistical model that accounts for experimental uncertainty in metabolite concentrations and uncertainty in  $K_D$  determination, and estimated  $Mg^{2+}$  speciation assuming single-site binding (meaning that one metabolite associated one  $Mg^{2+}$  ion). The second method approximated  $Mg^{2+}$  speciation to different metabolites but did not directly determine free  $Mg^{2+}$  concentration. The statistical model is described in detail in the Supporting Methods. Briefly, concentration errors were propagated from uncertainties in reagent masses and volumes used during sample preparation, and  $K_D$  uncertainties were obtained from the fits (Table 1). Both uncertainties were then randomly seeded into Equation 1 1000 times to create a series of virtual artificial cytoplasm with different errors, where  $[Mg]_T$  is the total  $Mg^{2+}$  concentration,  $[Mg]$  is the free  $Mg^{2+}$  concentration, “*i*” is an integer representing each metabolite in a mixture,  $N$  is the total number of metabolites in a mixture,  $[L_i]_T$  is the concentration of the “*i*th” metabolite in a mixture, and  $K_D$  is the dissociation constant.

$$[Mg]_T = [Mg] + \sum_{i=1}^N \frac{[L_i]_T [Mg]}{K_D + [Mg]} \quad (1)$$

Then, equation 1 was solved numerically to determine the free  $Mg^{2+}$  concentration produced at a given total  $Mg^{2+}$  concentration in a virtual artificial cytoplasm.

On the basis of agreement between the HQS data and the statistical simulation, methods 1 and 2, respectively, the two methods support a model in which  $Mg^{2+}$  speciated in artificial cytoplasms largely according to single-site binding within or below the biological free  $Mg^{2+}$  concentration range of 0.5 to 3 mM  $Mg^{2+}$ . However,  $Mg^{2+}$  did not speciate according to a single-site model at higher free  $Mg^{2+}$  concentrations (Figure 1E-G).

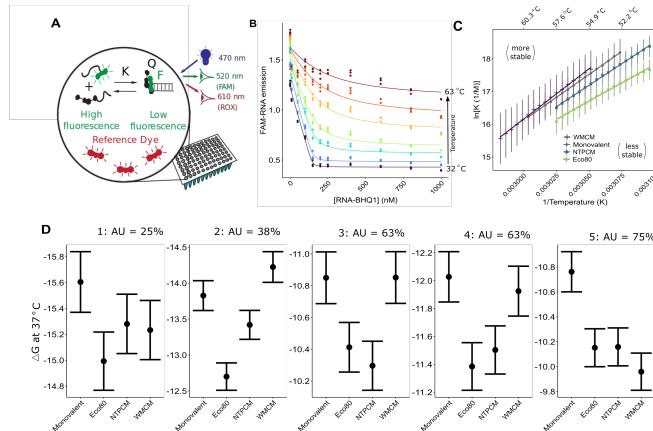
In Eco80, the statistical model suggests that the metabolites should buffer the free  $Mg^{2+}$  concentration in the biological  $Mg^{2+}$  range, where a 20 mM increase in the total  $Mg^{2+}$  from 20 to 40 mM leads to only a 2.5 mM increase in free  $Mg^{2+}$  from 0.5 to 3 mM (Figure 1E, hex bins). Free  $Mg^{2+}$  concentrations measured in this range with HQS emission are consistent with this single-site behavior (Figure 1E, black data points). At higher free  $Mg^{2+}$  concentrations, Eco80 should lose its free  $Mg^{2+}$  buffering capacity as chelators become saturated, and the free  $Mg^{2+}$  should increase with the total  $Mg^{2+}$  (model in Figure 1E, hex bins). The free  $Mg^{2+}$  concentration measured with HQS does not increase as fast as the statistical model predicts above 3 mM free  $Mg^{2+}$  (Figure 1E, compare black data points and hex bins). Free  $Mg^{2+}$  in Eco80 is expected to increase from 3 mM to ~100 mM as the total  $Mg^{2+}$  concentration is increased from 40 mM to 200 mM (Figure 1E, hex bins). However, the free  $Mg^{2+}$  concentration measured with HQS only increases from 3 mM to ~10 mM (Figure 1E, data points). One possibility is that multivalent interactions, where several  $Mg^{2+}$  saturated metabolites interact with additional  $Mg^{2+}$  molecules, dominate the equilibrium. Such non-single-site behavior above 3 mM free  $Mg^{2+}$  is also observed in the NTPCM and WMCM artificial cytoplasms (Figure 1F & G), and was observed previously.<sup>21</sup>

**Commented [BPC1]:** for me: could you just have fit the data in the upper row to a quadratic version of eq 1? Of course you need your technique for complex mixtures but did you really need it here? Did you ever try my alternative method to see if it agrees with yours?

**Commented [SJP2R1]:** Yes

**Table 2. Mg<sup>2+</sup> concentrations used to obtain 2 mM free Mg<sup>2+</sup> in artificial cytoplasm.**

| Condition | [Total Mg <sup>2+</sup> ] (mM) | [Chelated Mg <sup>2+</sup> ] (mM) | [Free Mg <sup>2+</sup> ] (mM) |
|-----------|--------------------------------|-----------------------------------|-------------------------------|
| Eco80     | 31.6                           | 29.6                              | 2.0                           |
| NTPCM     | 25.0                           | 23.0                              | 2.0                           |
| WMCM      | 6.4                            | 4.4                               | 2.0                           |



**Figure 2** *E. coli* metabolite Mg<sup>2+</sup> mixtures destabilize RNA helices. **(A)** Layout of a fluorescence binding isotherm assay in a Real-Time PCR machine. **(B)** Raw fluorescence binding isotherms fit to determine equilibrium constants with MeltR. Data points represent raw data. Curves represent curve fits. Colors represent different temperatures (purple: 32.3, blue: 41.8, teal: 51.3, green: 54.6, yellow: 58.4, orange: 60.7, red: 63.1 °C). **(C)** Van't Hoff relationship between the helix association equilibrium constant and temperature for helix 2:GGCAUCU/AGGAUGCG folding in background monovalent metal ions, Eco80, NTPCM, and WMCM. All conditions contain 2 mM free Mg<sup>2+</sup>, 240 mM Na<sup>+</sup> 140 mM K<sup>+</sup>. Points and error bars represent association constants and standard errors propagated from the fit (using MeltR). Lines represent the fits to the Van't Hoff equation that MeltR uses to calculate folding energies. A shift in the Van't Hoff relationship, down and to the right of the plot area, indicates that Eco80 destabilizes the helix. **(D)** The Gibbs free energy at 37 °C ( $\Delta G^{37\text{°C}}$ ) in Eco80, NTPCM, and WMCM compared to the  $\Delta G^{37\text{°C}}$  in background monovalent metal ions for five RNA helices. All conditions contain 2 mM free Mg<sup>2+</sup> and errors were propagated assuming 1.5% uncertainty in the Gibb's free energy at 37 °C (see methods for error analysis).

#### Thermodynamic analysis of RNA helices in Eco80 by fluorescence binding isotherms

We sought to understand how Eco80 affects the thermodynamic stability of RNA. Stability of RNA helices is traditionally measured with UV-absorbance-detected melting curves, typically monitored at 260 or 280 nm.<sup>34,35</sup> However, such absorbance melting curves cannot measure helix stability in Eco80 because of the high absorptivity of the nucleotide metabolites. Thus, we pursued a fluorescence-detected binding isotherm assay.

Helix stability was monitored 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 it is bound in duplex with RNA-BHQ1. We used a binding isotherm method, where RNA-BHQ1 is titrated into a constant concentration of FAM-RNA (SI figure 3), resulting

Lastly, we sought to empirically determine how much total Mg<sup>2+</sup> is required to attain a free Mg<sup>2+</sup> concentration of 2 mM in Eco80, NTPCM, and WMCM. The relationship between the free Mg<sup>2+</sup> calculated from HQS emission and the total Mg<sup>2+</sup> concentration in each artificial cytoplasms was fit to a polynomial to empirically approximate the data (Figure 1 E-G, blank lines), and the total Mg<sup>2+</sup> concentration required to produce 2 mM Free Mg<sup>2+</sup> was calculated from the polynomial fit (see methods for details). This resulted in predicted 31.6, 25.0, and 6.4 mM total Mg<sup>2+</sup> to produce 2 mM free Mg<sup>2+</sup> in Eco80, NTPCM, and WMCM, respectively (Table 2).

in a binding isotherm (Figure 2B). We favored binding isotherms over fluorescence-detected melts because of the dependence of FAM emission on temperature.<sup>36-38</sup> Emission of FAM was monitored at different temperatures, resulting in an isotherm every 0.5 °C from 20 to 80 °C (Figure 2B).

Raw fluorescence was fit with MeltR, a program created by the authors, to determine folding energies. MeltR is a package of functions in the R programming language that allows facile conversion of raw data to folding energies (see Methods for details). MeltR calculates folding energies using two Van't Hoff methods: (1) directly fitting a Van't Hoff plot as a function of temperature (Figure 3C) and (2) globally fitting raw fluorescence emission.

#### Eco80 thermodynamically destabilizes RNA helices

We used fluorescence binding isotherms to determine helix folding energies in a background monovalent metal ion control (240 mM NaCl 140 mM KCl), Eco80, NTPCM, and WMCM for a set of five representative eight base-pair RNA

**Commented [BPC3]:** To allow better comparison across the oligo sets, try plotting the data as follows:

- 1.) See Monovalent at 0.0
- 2.) Plot data (adjust y-max and y-min) such that less stable oligos are higher on the plot.
- 3.) The error bars confuse me here. Are they coming from experiment or just a plot of a fixed percent, which you say is 1% but looks more like 10% to me. In any case, it should come from something else?

**Commented [SJP4R3]:** I'm confused we should talk in person

- 1.) You want me to plot ddG?
- 2.) This isn't possible with ddG?
- 2.) 10% of 15.6 kcal/mol is 1.56 kcal/mol. 1% is 0.156 kcal/mol.

helices; all solutions contain 2 mM free Mg<sup>2+</sup>, as per Table 2. This helix set was designed to contain representatives of all 10 Watson-Crick nearest neighbor parameters and vary in AU content from 25% to 75%. Both of the aforementioned methods to determine folding energies in MelT<sup>R</sup> agreed (SI table 5) so the results from the Van't Hoff plot were reported. Results are summarized in Table 3, which is ranked according to the AU content of the duplex. Errors in the main text are reported 1.5% in terms of the ΔG<sup>37°C</sup> and a detailed error analysis is in the supplemental methods.

All five representative helices were significantly destabilized in Eco80 relative to the background monovalent condition, meaning the ΔΔG<sup>37°C</sup> between the background monovalent condition and Eco80 was larger than its propagated uncertainty (Table 3, Figure 2D). We did not observe a clear relationship between AU content and destabilization (SI

figure 4). Thus, Eco80 destabilizes RNA helices but the underlying sequence dependence is not apparent.

To better understand how the various components of Eco80 contribute to destabilizing RNA helices, we analyzed the effects of the strong and weak Mg<sup>2+</sup>-chelating metabolites separately. NTPCM, which is comprised of strong Mg<sup>2+</sup>-chelating metabolites, consistently destabilized RNA helices (Figure 2D), by 0.32 to 0.60 kcal/mol (Table 3). The destabilizing effect of Eco80 appears to be related to the AU content of the helix with destabilization increasing linearly from +0.32 kcal/mol at 25% AU content to 0.60 kcal/mol at 75% AU content ( $R^2 = 0.99$ , SI figure 4).

In contrast, WMCM, which is composed of weak Mg<sup>2+</sup>-chelating metabolites, destabilized, had no effect, or stabilized RNA helices in a fashion that did not depend on AU content (Figure 2D, Table 3). Similar to Eco80, the sequence dependence of stabilization or destabilization is not clear.

**Table 3. Stability of RNA helices in *E. coli* metabolite mixtures.**

| Sequence <sup>a</sup>                   | AU content (%) | Condition <sup>b</sup> | ΔG (kcal/mol) <sup>c</sup> | ΔΔG (kcal/mol) <sup>c</sup> |
|---|----------------|------------------------|----------------------------|-----------------------------|
| 1:<br>5'-CGGAUGGC-3'<br>3'-GCCUACCG-5'  | 25%            | 2 mM free              | -15.60 (0.23)              | --                          |
|   |                | Eco80                  | -15.00 (0.23)              | +0.60 (0.32)                |
|   |                | NTPCM                  | -15.28 (0.23)              | +0.32 (0.33)                |
|   |                | WMCM                   | -15.20 (0.23)              | +0.40 (0.33)                |
| 2:<br>5'-CGCAUCCU-3'<br>3'-GCGUAGGA-5'  | 38%            | 2 mM free              | -13.82 (0.21)              | --                          |
|   |                | Eco80                  | -12.70 (0.19)              | +1.12 (0.28)                |
|   |                | NTPCM                  | -13.41 (0.20)              | +0.41 (0.29)                |
|   |                | WMCM                   | -14.22 (0.21)              | -0.40 (0.30)                |
| 3:<br>5'-CGUAUGUA-3'<br>3'-GCAUACAU-5'  | 63%            | 2 mM free              | -10.85 (0.16)              | --                          |
|   |                | Eco80                  | -10.41 (0.16)              | +0.44 (0.23)                |
|   |                | NTPCM                  | -10.30 (0.15)              | +0.55 (0.22)                |
|   |                | WMCM                   | -10.85 (0.16)              | 0.00 (0.23)                 |
| 4:<br>5'-CCAUAUCA-3'<br>3'-GGUUAUAGU-5' | 63%            | 2 mM free              | -12.02 (0.18)              | --                          |
|   |                | Eco80                  | -11.38 (0.17)              | +0.64 (0.25)                |
|   |                | NTPCM                  | -11.50 (0.17)              | +0.52 (0.25)                |
|   |                | WMCM                   | -11.90 (0.18)              | +0.12 (0.25)                |
| 5:<br>5'-CCAUUUUA-3'<br>3'-GGUUAUAAU-5' | 75%            | 2 mM free              | -10.76 (0.16)              | --                          |
|   |                | Eco80                  | -10.15 (0.15)              | +0.61 (0.22)                |
|   |                | NTPCM                  | -10.16 (0.15)              | +0.60 (0.22)                |
|   |                | WMCM                   | -9.94 (0.15)               | +0.82 (0.22)                |

<sup>a</sup>The first sequence was 5'-FAM labeled and the second sequence was 3'-BHQ1 labeled. <sup>b</sup>All solutions contain 2 mM Free Mg<sup>2+</sup> 240 mM Na<sup>+</sup> 140 mM K<sup>+</sup>. <sup>c</sup>Extra significant digits were included to avoid propagating rounding errors.

**Commented [BPC5]:** Add 1 row before this row and enter 1 M NaCl numbers. Repeat for Sequences 2-5.

By so-doing we can show that Eco80 is even more destabilized relative to traditional parameters in 1 M NaCl.

I think you are measured all these numbers. Moreover, it would be good to have a Supporting Table that has 1 M NaCl numbers with columns of: Predicted, UV Melt Measured, FDBI Measured.

**Commented [SJP6R5]:** I have 1 M NaCl numbers with FDBI and predicted. The melts would be very expensive because they use about 50x as much material.

I don't want to go after 1 M NaCl numbers in this paper because I think it can only serve to antagonize 1 M NaCl melters and a broader audience is not interested. I believe a the appropriate comparison for a broader audience is low monovalent and 2 mM free Mg<sup>2+</sup> to low monovalent and 2 mM free Mg<sup>2+</sup> and metabolites.

I want to go after the 1 M NaCl numbers in the second paper, where we will have more data, folding predictions, and a comparison to *in vivo* reactivity data.

Overall, the net effect of Eco80 on RNA helices is destabilization, with AU-content-dependent destabilizing interactions dominating for strong  $Mg^{2+}$ -chelating metabolites, and a mixture of stabilizing and destabilizing interactions for weak  $Mg^{2+}$ -chelating metabolites. Apparently weak  $Mg^{2+}$  chelation gives rise to a hidden sequence dependence that carries over to Eco80 (see Discussion).

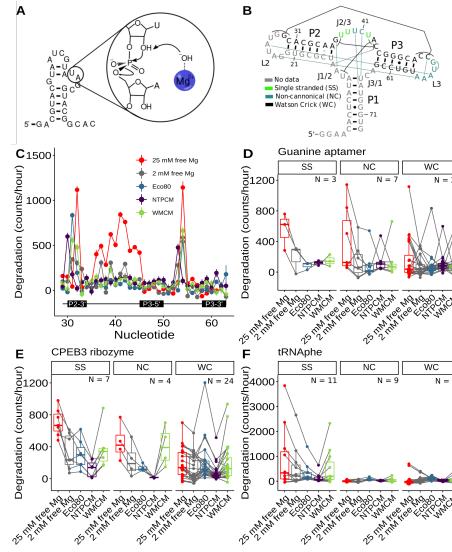
#### Eco80 protects RNA from chemical degradation

Several studies indicate that weak and strong  $Mg^{2+}$ -chelating metabolites reduce  $Mg^{2+}$ -mediated RNA degradation.<sup>22,39</sup> To assess whether Eco80 stabilizes the chemical structure of RNA, we used an in-line probing (ILP) assay, which takes advantage of the natural susceptibility of the RNA phosphodiester backbone to cleavage.<sup>40</sup> For ILP, the 2'-hydroxyl is deprotonated by a  $Mg^{2+}$ -hydroxide ( $Mg^{2+}\text{-OH}$ ), and serves as a nucleophile to attack the adjacent phosphate 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.<sup>40</sup> For this assay, 5'-<sup>32</sup>P RNAs are incubated at 37 °C for about 90 h to facilitate in-line cleavage, with time points taken regularly. RNA fragments are then fractionated on a denaturing PAGE gel (SI figure 5), providing single nucleotide resolution of degradation rates measured by the increase in counts for a given band. In-line degradation rates for biological RNA in Eco80, NTPCM, and WMCM with enough total  $Mg^{2+}$  to maintain 2 mM free  $Mg^{2+}$  were compared to degradation rates in a 2 mM free  $Mg^{2+}$  and 25 mM free  $Mg^{2+}$  condition. All conditions contained 240 mM K<sup>+</sup> and 140 mM Na<sup>+</sup>. The 25 mM free  $Mg^{2+}$  condition was chosen because it is a common free  $Mg^{2+}$  condition *in-vitro* and is similar to the 25 and 31.6 total  $Mg^{2+}$  condition used for NTPCM and Eco80, respectively (Table 2).

We first determined degradation rates for the guanine riboswitch aptamer in different artificial cytoplasmas (Figure 3B). The guanine riboswitch aptamer has been studied extensively, providing structural and mechanistic information.<sup>41–43</sup> We chose to study the guanine riboswitch in its guanine-ligand-unbound, apo, state for experimental simplicity. The expression platform was trimmed to prevent structural switching and the guanine ligand was not added to the solution to favor the apo state. Moreover, guanine binding to the aptamer induces structural changes only at nucleotides directly mediating the guanine binding site,<sup>41</sup> indicating that information provided by X-ray crystal structures of the ligand bound aptamer is relevant for a structural analysis of degradation rates.

Great care was taken in our analysis to confirm that the guanine aptamer adopted a similar structure between conditions. 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 (L3), indicating that the apo guanine riboswitch aptamer adopts a similar structure in these conditions (Figure 3C, SI figure 6). The 25 mM free  $Mg^{2+}$  condition exhibited higher degradation rates than the other conditions in the J2/3 junction (Figure 3C). This pattern is similar to ILP data published for another guanine riboswitch at a higher pH and  $Mg^{2+}$  concentration,<sup>44</sup> indicating that the increase in degradation rates in the 25 mM free  $Mg^{2+}$

condition is dependent on the presence of  $Mg^{2+}\text{-OH}$  complexes (SI figure 6). To further confirm that the guanine aptamer adopts similar structures in all conditions, we collected small angle X-ray scattering (SAXS) data on the apo form of the aptamer. Bell shaped Kratky plots overlay one another, indicating that the structure of the guanine aptamer is folded and similar between conditions (SI figure 7A). Guinier analysis and p(r) analysis, where the maximum is the radius of gyration (SI figure 7B), and porod analysis, reveal similar radius of gyration,  $D_{\max}$ , and porod volume between solution conditions, with slight compaction in Eco80 (SI table 6). This compaction is similar to the previously reported increased compaction of tertiary structure in the presence of  $Mg^{2+}$  and crowders.<sup>10</sup> Lastly, electron density reconstructions and bead model reconstructions are consistent with the crystal structure of the guanine riboswitch aptamer in every condition (SI figure 7C-G). 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 activity of  $Mg^{2+}\text{-OH}$  complexes rather than changes in structure.



**Figure 3** *E. coli* metabolite and  $Mg^{2+}$  mixtures stabilize the chemical structure of RNA. (A) ILP degradation mechanism facilitated by  $Mg^{2+}$ . (B) Secondary structure of the guanine riboswitch aptamer with tertiary contacts. (C) Degradation rate, at each residue in different solution conditions. Groupings are based on analysis of crystal structures. SS: Single stranded, the base was not participating in hydrogen bonding interactions with other residues. NC: non-canonical, the base was forming non-canonical hydrogen bonding interactions in the tertiary structure. WC: Watson-Crick, the base was in a helix composed mostly of Watson-Crick base pairs.

We sought to better characterize the structural dependence of RNA degradation in different conditions. We therefore extended the study to the CPEB3 ribozyme and yeast

**Commented [BPC9]:** Should all three panels range from 0-1200? This would require you to use a break for the tRNA for a few points.

**Commented [SJP10R9]:** I don't think they should have the same range because they are P32-labeled samples ran on different gels. If the RNA was hotter or the gel was exposed longer, it will have a steeper slope I think.

**Commented [BPC7]:** Not just this region. L3? (also label it L5?)

Are you sure that J2/3 and L3/L5 kiss forms in the apo? Maybe fleetingly? We need to understand this to write this convincingly and if so, then we can use this to justify choosing the apo over the holo.

**Commented [SJP8R7]:** Yes I am sure. No ILP changes occur in the L2 and L3 stem loop ranges on guanine binding indicating that the kissing interaction is robust in the apo state.

tRNA<sup>Phe</sup> (SI figure 8, 9, & 10). We inspected the crystal structure of these two RNAs, plus the original guanine aptamer, and manually classified each residue as single stranded (SS), meaning that the base was not participating in hydrogen bonding interactions with other residues, non-canonical (NC), meaning that the base was forming non-canonical hydrogen bonding interactions with other residues in the tertiary structure, and Watson-Crick (WC), meaning that the base was in a helix composed mostly of Watson-Crick base pairs (SI table 7). Rates of ILP were then analyzed in box plots (Figure 3D-F).

We begin box plot analysis with the guanine aptamer (Figure 3D). We had data for 3 single stranded nucleotides without accompanying non-canonical hydrogen bonding interactions. We observed decreased degradation rates at the single stranded (SS) nucleotides in 2 mM free Mg<sup>2+</sup>, Eco80, NTPCM, and WMCM, in comparison to the 25 mM free Mg<sup>2+</sup> condition. Likewise, we observed an overall decrease in reactivity in 2 mM free Mg<sup>2+</sup>, Eco80, NTPCM, and WMCM in comparison to the 25 mM free Mg<sup>2+</sup> condition for nucleotides involved in non-canonical tertiary interactions (NC). In contrast, degradation rates for nucleotides participating in Watson-Crick base pairing interactions were independent of solution conditions. Thus, we observed a trend of protection from degradation in artificial cytoplasm, even with similar amounts of total Mg<sup>2+</sup> in solution as the 25 mM free Mg<sup>2+</sup> condition.

We repeated our degradation assay with the cleaved-CPEB3 ribozyme and yeast tRNA<sup>Phe</sup>, to test if the protection from degradation in artificial cytoplasm was broadly applicable (SI figure 8, 9, & 10). For the cleaved-CPEB3 ribozyme, degradation rates at single-stranded residues were reduced in 2 mM free Mg<sup>2+</sup>, Eco80, and NTPCM conditions in comparison to the 25 mM free Mg<sup>2+</sup> condition (Figure 3E). Interestingly, the degradation rates of single-stranded residues recovered in WMCM, indicating that degradation rates are partially dependent on the strength of Mg<sup>2+</sup> chelation. Likewise, the degradation rates for residues that were predicted to participate in non-canonical tertiary contacts were reduced in 2 mM free Mg<sup>2+</sup>, Eco80, and NTPCM, but not WMCM, in comparison to the 25 mM free Mg<sup>2+</sup> condition, further indicating that degradation rates are dependent on the strength of Mg<sup>2+</sup> chelation. Degradation rates were similar for nucleotides participating in Watson-Crick base-pairs between all conditions.

Yeast tRNA<sup>Phe</sup> exhibits almost no in-line degradation except for the single stranded nucleotides in the P3 stem loop, termed the anticodon loop (SI figure 9D). Single stranded nucleotides showed reduced degradation rates in the 2 mM free Mg<sup>2+</sup>, Eco80, and NTPCM conditions in comparison to the 25 mM free condition, and degradation rates recovered in WMCM. Degradation rates were constant across conditions for nucleotides that form Watson-Crick base pairs and for nucleotides that form non-canonical contacts, which is different than the increased degradation observed for nucleotides that form non-canonical contacts in the guanine riboswitch aptamer and the CPEB3 ribozyme. One possible explanation is that the tertiary structure of tRNA<sup>Phe</sup> is less dynamic than the tertiary structure of the guanine riboswitch aptamer and the CPEB3 ribozyme, thus reducing the degradation rates in regions that participate in non-

canonical tertiary interactions to the baseline levels observed for nucleotides participating in Watson-Crick base pairs.

Overall, the in-line degradation assay indicated that Eco80 and NTPCM protect RNA from Mg<sup>2+</sup>-OH-mediated degradation in structural regions that are susceptible to in-line cleavage, even though both artificial cytoplasmas have relatively high concentrations of total Mg<sup>2+</sup>. WMCM shows an intermediate effect between the high degradation rates in the 25 mM free Mg<sup>2+</sup> condition and the low degradation rates in 2 mM free Mg<sup>2+</sup>, Eco80, and NTPCM, indicating that degradation rates are dependent on Mg<sup>2+</sup> chelation strength (see Discussion).

#### Eco80 supports RNA catalysis

Weak metabolite-chelated Mg<sup>2+</sup> is known to promote catalysis by ribozymes. For example, CPEB3 ribozyme catalysis is enhanced by about 1.6-fold by 2 mM free Mg<sup>2+</sup> in solution with an estimated 11.3 mM glutamate-chelated Mg<sup>2+</sup>, in comparison to catalysis in 2 mM free Mg<sup>2+</sup> alone.<sup>22</sup> Thus, we also hypothesized that Eco80 metabolites would also promote CPEB3 catalysis.

We compared CPEB3 ribozyme cleavage rates in 2 mM free Mg<sup>2+</sup> and 25 mM free Mg<sup>2+</sup> to Eco80, NTPCM, and WMCM containing enough total Mg<sup>2+</sup> to produce 2 mM free Mg<sup>2+</sup> (Table 2). All conditions contain 240 mM Na<sup>+</sup> and 140 mM K<sup>+</sup>. Briefly, we purified full length CPEB3 ribozyme (Figure 4A), incubated CPEB3 in artificial cytoplasm, fractionated 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 11). The fraction cleaved as a function of time was fit to a single exponential equation to estimate the reaction rate constant (Figure 4B). CPEB3 ribozyme catalysis was reduced in all conditions in comparison to the 25 mM free Mg<sup>2+</sup> control (Figure 4C). Surprisingly, CPEB3 catalysis was reduced in Eco80 by about 1/2 in comparison to the 2 mM free Mg<sup>2+</sup> control, despite the 31.6 mM total Mg<sup>2+</sup> in Eco80. CPEB3 catalysis was reduced by about 1/3 in NTPCM in comparison to the 2 mM free Mg<sup>2+</sup> 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 Mg<sup>2+</sup>. In summary, Eco80 supports RNA catalysis but does not enhance catalysis in comparison to the 2 mM free Mg<sup>2+</sup> condition. CPEB3 reaction rates in Eco80 are between the rates in WMCM and NTPCM. WMCM likely has exposed Mg<sup>2+</sup> to help fold the RNA, while NTPCM does not, and moreover may denature the RNA as per Figure 2 (see Discussion).

WMCM may be more biologically relevant than Eco80 for studying CPEB3 ribozyme activity. We performed an analysis of absolute metabolite concentrations in yeast and mammalian imbk cells, which have a closer evolutionary relationship to human cells where CPEB3 exists (Figure 4D). Absolute metabolite concentrations were compiled from the literature and the 11 most abundant metabolites that compose 80% of the Yeast and mammalian metabolome were selected to compose hypothetical Yeast80 and Mammal80 artificial cytoplasmas.<sup>45</sup> Estimated metabolite/Mg<sup>2+</sup> binding constants<sup>21</sup> were used to classify each metabolite in Yeast80 and Mammal80 as strong (NTPCM) or weak Mg<sup>2+</sup>

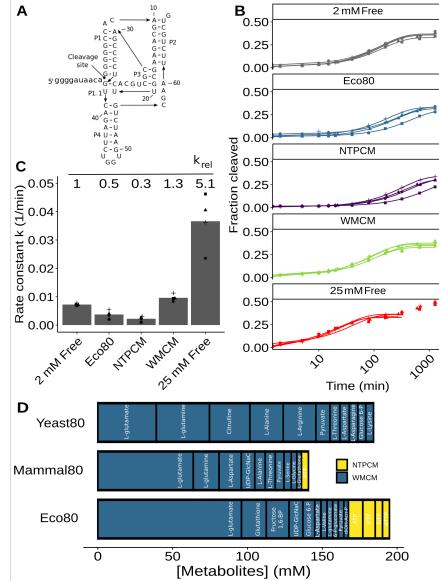
**Commented [BPC11]:** how? Inspection? Some software, and if so written by you or someone else?

These classes should be provided in the SI as a table. Provide nt 1, 2, 3, ... and in the next column list as SS, NC, or WC. Provide the pdb(s) used for this. In Figs E-G, provide the number of such bases in parentheses next to "SS", "NC", and "WC" at the top of the chart.

**Commented [BPC12]:** Choose the hammerhead result instead? Is 2 mM the only Mg<sup>2+</sup> data you trust?

**Commented [SJP13R12]:** I'm choosing Ryota's CPEB3 ribozyme because it is the same RNA and because it has a similar effect.

chelators. We found that Yeast80 and Mammal80 would be depleted in strongly chelated-Mg<sup>2+</sup>, with Yeast80 having no strong Mg<sup>2+</sup> chelators and Mammal80 having ~4 mM strong Mg<sup>2+</sup> chelators. (Figure 4D). Thus, the 1.3-fold rate enhancement in is more relevant to CPEB3 function in human cells than the rate decrease in Eco80.



**Figure 4** Eco80 supports CPEB3 ribozyme catalysis. **(A)** Secondary structure of the un-cleaved CPEB3 ribozyme. **(B)** Fraction of cleaved CPEB3 as a function of time fit to a single exponential. Four technical replicates are displayed. 2 mM and 25 mM free refers to the Mg<sup>2+</sup> concentration. All conditions contain a background of 240 mM Na<sup>+</sup> and 140 mM K<sup>+</sup>. Enough total Mg<sup>2+</sup> was added to Eco80, NTPCM, and WMCM to maintain a 2 mM free Mg<sup>2+</sup> concentration. **(C)** Rate constant ( $k$ ) for the CPEB3 ribozyme in different conditions.  $k_{rel}$  is the relative rate constant in comparison to the 2 mM free Mg<sup>2+</sup> condition. **(D)** Composition of artificial cytoplasmas composed of 80% of yeast and mammalian iMBK metabolites, termed Yeast80 and Mammal80, respectively, compared to the composition of Eco80. Each box represents one abundant metabolite. NTPCM represents nucleotide metabolites and WMCM represents metabolites that are expected to weakly chelate Mg<sup>2+</sup> with K<sub>d</sub>s greater than 2 mM.

## Discussion

In summary, we have used a bottom-up, *aufbau*, approach to create a complex but manageable artificial cytoplasm, termed Eco80, which encapsulates 80% of the *E. coli* metabolome (Figure 5A). We also broke down Eco80 into sub-artificial cytoplasmas, which contain either metabolites that strongly chelate Mg<sup>2+</sup> (NTPs), or metabolites that weakly chelate Mg<sup>2+</sup>, providing mechanistic insight into the effects of Mg<sup>2+</sup> speciation on RNA in cells.

Importantly, we have adopted the Mg<sup>2+</sup> sensitive dye, HQS,<sup>33</sup> to measure Mg<sup>2+</sup> speciation in artificial cytoplasmas. A key challenge to studying RNA in *in vivo*-like conditions is knowing how components will affect the speciation of Mg<sup>2+</sup> between free and chelated Mg<sup>2+</sup>. Published binding constants for cellular components are unreliable, as they apply to solutions with specific ionic character.<sup>20</sup> More often, binding constants are not known.<sup>21</sup> Lastly, predicting Mg<sup>2+</sup> speciation using binding constants requires making assumptions about the stoichiometry of Mg<sup>2+</sup>-component complexes, which may or may not be valid. For example, in this work, the free Mg<sup>2+</sup> concentration in Eco80, NTPCM, and WMCM approximated the free Mg<sup>2+</sup> concentration that was predicted using binding constants in the biological range of 0.5 to 3 mM free Mg<sup>2+</sup>, but the prediction was not accurate at higher free Mg<sup>2+</sup> concentrations (Figure 1 E-G). Thus, the HQS assay provides information on Mg<sup>2+</sup> speciation in biologically relevant solutions, without requiring assumptions or accurate Mg<sup>2+</sup> binding constants. Although we used this assay to directly measure Mg<sup>2+</sup> speciation in mixtures of metabolites, this assay can be applied to other biological molecules.

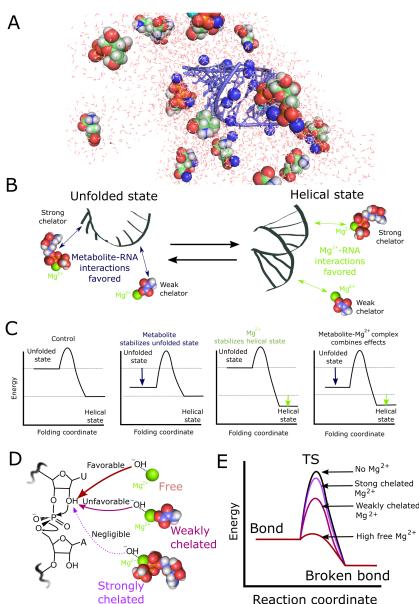
Our Mg<sup>2+</sup> speciation calculations and HQS experiments indicate that metabolites play an important role in buffering the free Mg<sup>2+</sup> concentration in cells. Recent theoretical and experimental work has demonstrated that the cellular environment buffers the concentration of biological molecules, thus reducing concentration noise *in vivo*.<sup>47,48</sup> Single-site-Mg<sup>2+</sup> interactions in Eco80 buffers the free Mg<sup>2+</sup> concentration between 0.5 and 3 mM Mg<sup>2+</sup>, given a total Mg<sup>2+</sup> change of 20 mM. This buffering effect is exaggerated at high total Mg<sup>2+</sup> concentrations in Eco80, where an increase in the total Mg<sup>2+</sup> concentration to 200 mM only increases the free Mg<sup>2+</sup> concentration to ~10 mM.

Our thermodynamic analysis of RNA helices in Eco80 indicates that the *E. coli* metabolome has a net destabilizing effect on RNA helices, with destabilizing effects dominating for NTPCM and a mixture of destabilizing and stabilizing effects observed for WMCM (Figure 2D). The effects of our artificial cytoplasm on RNA helix stability can be understood using a model that combines the established effects of polar small molecules and Mg<sup>2+</sup> on RNA helix stability. Polar small molecules are known to interact favorably with the exposed bases in the unfolded state (Figure 5B).<sup>49-52</sup> Likewise, Mg<sup>2+</sup> is known to interact favorably with the high density of negative charge in helical RNA. Thus, metabolites destabilize helices by favoring the unfolded state and Mg<sup>2+</sup> stabilizes helices by favoring the helical state (Figure 5C). The changes in helix formation energy in Mg<sup>2+</sup>/metabolite mixtures demonstrate a balance between metabolites favoring the unfolded state and Mg<sup>2+</sup> favoring the helical state (Figure 5C). For example, NTPCM strongly chelates Mg<sup>2+</sup>, thus sequestering Mg<sup>2+</sup> from interacting with the folded state so that the destabilizing interactions between NTPs and RNA dominate, leading to a consistent destabilization of RNA helices (Figure 2D). In contrast, WMCM only weakly sequesters Mg<sup>2+</sup>, so that Mg<sup>2+</sup> is available for favorable interactions with helices. This leads to the inconsistent destabilization/stabilization of RNA helices observed in WMCM, dependent on the relative strength of stabilizing Mg<sup>2+</sup>-RNA

interactions and destabilizing metabolite-RNA interactions (Figure 5C).

NTPCM destabilizes AU rich helices more than GC rich helices (SI figure 4). A similar destabilizing effect on RNA G-quadruplex structures has been observed for cytidine nucleotides.<sup>53</sup> Interestingly, other nucleotides (A and C) had a smaller destabilizing effect, indicating that G-quadruplexes are destabilized by favorable base-pairing interactions between cytidine nucleotides in solution and Gs in the unfolded state of the RNA. NTPCM is mostly composed of ATP, UTP, and dTTP (22.5 mM total versus 4.9 mM GTP). ATP, UTP, and dTTP are expected to form stronger transient hydrogen bonds with As and Us in the unfolded state of RNA, explaining the AU dependence of helix destabilization by NTPCM.

Our analysis of RNA degradation in Eco80 indicates that metabolites protect susceptible regions of RNA from Mg<sup>2+</sup>-OH-mediated degradation (Figure 3). Eco80 and NTPCM had the strongest protective effects, while WMCM had an intermediate protective effect, indicating that protection from degradation is dependent on the strength of the chelating interaction between metabolites and Mg<sup>2+</sup>. In this model, in-line cleavage of the RNA backbone is limited by the formation of Mg<sup>2+</sup>-OH<sup>-</sup> species, which is favorable for free Mg<sup>2+</sup>, unfavorable for weakly-chelated Mg<sup>2+</sup>, and negligible for strong NTP chelated-Mg<sup>2+</sup> (Figure 5D). Thus, RNA degradation reaction rates are weakly reduced by depletion of active Mg<sup>2+</sup>-OH<sup>-</sup> species in the presence of weak Mg<sup>2+</sup> chelators and strongly reduced by depletion of active Mg<sup>2+</sup>-OH<sup>-</sup> species in the presence of strong Mg<sup>2+</sup> chelators (Figure 5E).



**Figure 5** Models describing the destabilization of RNA helices and stabilization of RNA chemical structure by Eco80. **(A)**

Semi-quantitative molecular representation of an RNA helix in Eco80. The average number of molecules (colored sphere models) in Eco80 that would occupy a sphere with a 50 Å radius was placed randomly around an 8N RNA helix using Pymol (blue cartoon, PDB 1SDR). Mg<sup>2+</sup> ions are represented with teal spheres. Solvent (red wires) and K<sup>+</sup> (blue spheres) were modeled using WAXSIS.<sup>46</sup> **(B-C)** Mechanism for destabilization of helices by metabolites and stabilization of helices by Mg<sup>2+</sup>. Net effect of the metabolite-chelated Mg<sup>2+</sup> combines metabolite interactions favoring the unfolded state and Mg<sup>2+</sup> interactions favoring the helical state. **(D-E)** In-line degradation of the RNA backbone mediated by Mg<sup>2+</sup> hydroxide species.

Our analysis of CPEB3 catalysis in Eco80 indicates that metabolite/Mg<sup>2+</sup> mixtures support RNA catalysis. A previous study of hammerhead ribozyme catalysis in the presence of nucleotides found that reaction rates were enhanced by NDP chelated-Mg<sup>2+</sup>, a weakly chelated-Mg<sup>2+</sup> species, and that NTP chelated-Mg<sup>2+</sup> has no effect on reaction rates.<sup>23</sup> Similarly, our results in metabolite mixtures found that WMCM enhanced CPEB3 ribozyme catalysis while NTPCM inhibited CPEB3 ribozyme catalysis. A previous study of CPEB3 ribozyme in the presence of weak amino acid-chelated Mg<sup>2+</sup> indicated that rate enhancement was not driven by direct interactions between amino acid-chelated Mg<sup>2+</sup> and the catalytic site, but by stabilization of catalytically relevant CPEB3 ribozyme structure.<sup>22</sup> Likewise, the thermodynamic destabilization of helices observed in Eco80 and NTPCM indicate that reduction CPEB3 catalysis is caused by destabilization of the catalytically relevant structure. Thus, ribozyme rate enhancement *in vivo* is likely dependent on the presence of weak metabolite chelators that stabilize the catalytically relevant structure and depletion of strong chelators that destabilize the catalytically relevant structure.

Eco80 has opposing effects on RNA which reflects the complexity of the cellular environment. The thermodynamic stability of RNA helices was weakened by Eco80, the chemical stability of RNA was enhanced by Eco80, and the catalysis of RNA was supported by Eco80. These seemingly contradictory effects can be understood using the speciation of Mg<sup>2+</sup> between weak and strong Mg<sup>2+</sup>-metabolite complexes in the biologically relevant metabolite mixture. The effects of Eco80 reflect RNA function *in vivo* and enhance the biological relevance of mechanistic studies of RNA.

#### ASSOCIATED CONTENT

**Supplemental Information.** Supplemental methods, Supplemental figures, Supplemental tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### AUTHOR INFORMATION

##### Corresponding Author

\*Philip C. Bevilacqua. Email: pcb5@psu.edu

##### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

##### Funding Sources

This work was supported by National Institutes of Health Grant R35-GM127064 to PCB. SAXS data collection was supported by

SIG S10 of the National Institutes of Health under award number S10-OD028589 for the small angle X-ray scattering and S10 OD030490 for the Wyatt SEC-MALS-DLS system to NYH.

## ACKNOWLEDGMENT

The authors thank Allison Williams and Dr. Elizabeth Jolley for useful discussions and suggestions. We also wish to thank the assistance of Julia Fecko at the X-ray Crystallography core in collecting the SAXS data.

## ABBREVIATIONS

RNA, ribonucleic acid; DNA, deoxynucleic acid; Mg<sup>2+</sup>, divalent magnesium ion; HQS, 8-Hydroxy-5-quinolinesulfonic acid; Eco80, 80% of *E. coli* metabolites; NTPCM, nucleotide triphosphate chelated Mg<sup>2+</sup>; WMCM, weak metabolite chelated Mg<sup>2+</sup>; ITC, isothermal titration calorimetry; SAXS, small angle X-ray scattering

## REFERENCES

- (1) Cech, T. R.; Steitz, J. A. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell* **2014**, *157* (1), 77–94. <https://doi.org/10.1016/j.cell.2014.03.008>.
- (2) Ritchey, L. E.; Su, Z.; Tang, Y.; Tack, D. C.; Assmann, S. M.; Bevilacqua, P. C. Structure-Seq2: Sensitive and Accurate Genome-Wide Profiling of RNA Structure *In Vivo*. *Nucleic Acids Res.* **2017**, *45* (14), e135–e135. <https://doi.org/10.1093/nar/gkx533>.
- (3) Zubradt, M.; Gupta, P.; Persad, S.; Lambowitz, A. M.; Weissman, J. S.; Rouskin, S. DMS-MaPseq for Genome-Wide or Targeted RNA Structure Probing *In Vivo*. *Nat. Methods* **2017**, *14* (1), 75–82. <https://doi.org/10.1038/nmeth.4057>.
- (4) Lan, T. C. T.; Allan, M. F.; Malsick, L. E.; Woo, J. Z.; Zhu, C.; Zhang, F.; Khandwala, S.; Nyeo, S. S. Y.; Sun, Y.; Guo, J. U.; Bathe, M.; Näär, A.; Griffiths, A.; Rouskin, S. Secondary Structural Ensembles of the SARS-CoV-2 RNA Genome in Infected Cells. *Nat. Commun.* **2022**, *13* (1), 1128. <https://doi.org/10.1038/s41467-022-28603-2>.
- (5) Leamy, K. A.; Assmann, S. M.; Mathews, D. H.; Bevilacqua, P. C. Bridging the Gap between *In Vitro* and *In Vivo* RNA Folding. *Q. Rev. Biophys.* **2016**, *49*. <https://doi.org/10.1017/S003358351600007X>.
- (6) Nakano, S.; Karimata, H. T.; Kitagawa, Y.; Sugimoto, N. Facilitation of RNA Enzyme Activity in the Molecular Crowding Media of Cosolutes. *J. Am. Chem. Soc.* **2009**, *131* (46), 16881–16888. <https://doi.org/10.1021/ja9066628>.
- (7) Kilburn, D.; Roh, J. H.; Guo, L.; Briber, R. M.; Woodson, S. A. Molecular Crowding Stabilizes Folded RNA Structure by the Excluded Volume Effect. *J. Am. Chem. Soc.* **2010**, *132* (25), 8690–8696. <https://doi.org/10.1021/ja101500g>.
- (8) Kilburn, D.; Roh, J. H.; Behrouzi, R.; Briber, R. M.; Woodson, S. A. Crowders Perturb the Entropy of RNA Energy Landscapes to Favor Folding. *J. Am. Chem. Soc.* **2013**, *135* (27), 10055–10063. <https://doi.org/10.1021/ja4030098>.
- (9) Lee, H.-T.; Kilburn, D.; Behrouzi, R.; Briber, R. M.; Woodson, S. A. Molecular Crowding Overcomes the Destabilizing Effects of Mutations in a Bacterial Ribozyme. *Nucleic Acids Res.* **2015**, *43* (2), 1170–1176. <https://doi.org/10.1093/nar/gku1335>.
- (10) Leamy, K. A.; Yennawar, N. H.; Bevilacqua, P. C. Cooperative RNA Folding under Cellular Conditions Arises from Both Tertiary Structure Stabilization and Secondary Structure Destabilization. *Biochemistry* **2017**, *56* (27), 3422–3433. <https://doi.org/10.1021/acs.biochem.7b00325>.
- (11) Nakano, S.; Karimata, H.; Ohmichi, T.; Kawakami, J.; Sugimoto, N. The Effect of Molecular Crowding with Nucleotide Length and Cosolute Structure on DNA Duplex Stability. *J. Am. Chem. Soc.* **2004**, *126* (44), 14330–14331. <https://doi.org/10.1021/ja0463029>.
- (12) Ghosh, S.; Takahashi, S.; Ohyama, T.; Endoh, T.; Tateishi-Karimata, H.; Sugimoto, N. Nearest-Neighbor Parameters for Predicting DNA Duplex Stability in Diverse Molecular Crowding Conditions. *Proc. Natl. Acad. Sci.* **2020**, *117* (25), 14194–14201. <https://doi.org/10.1073/pnas.1920886117>.
- (13) Adams, M. S.; Znosko, B. M. Thermodynamic Characterization and Nearest Neighbor Parameters for RNA Duplexes under Molecular Crowding Conditions. *Nucleic Acids Res.* **2019**, *47* (7), 3658–3666. <https://doi.org/10.1093/nar/gkz019>.
- (14) Tyrrell, J.; Weeks, K. M.; Pielak, G. J. Challenge of Mimicking the Influences of the Cellular Environment on RNA Structure by PEG-Induced Macromolecular Crowding. *Biochemistry* **2015**, *54* (42), 6447–6453. <https://doi.org/10.1021/acs.biochem.5b00767>.
- (15) Stadtmiller, S. S.; Pielak, G. J. Protein-Complex Stability in Cells and *In Vitro* under Crowded Conditions. *Curr. Opin. Struct. Biol.* **2021**, *66*, 183–192. <https://doi.org/10.1016/j.sbi.2020.10.024>.
- (16) You, X.; Nguyen, A. W.; Jabaiah, A.; Sheff, M. A.; Thorn, K. S.; Daugherty, P. S. Intracellular Protein Interaction Mapping with FRET Hybrids. *Proc. Natl. Acad. Sci.* **2006**, *103* (49), 18458–18463. <https://doi.org/10.1073/pnas.0605422103>.
- (17) Phillip, Y.; Kiss, V.; Schreiber, G. Protein-Binding Dynamics Imaged in a Living Cell. *Proc. Natl. Acad. Sci.* **2012**, *109* (5), 1461–1466. <https://doi.org/10.1073/pnas.1112171109>.
- (18) Sukenik, S.; Ren, P.; Gruebele, M. Weak Protein-Protein Interactions in Live Cells Are Quantified by Cell-Volume Modulation. *Proc. Natl. Acad. Sci.* **2017**, *114* (26), 6776–6781. <https://doi.org/10.1073/pnas.1700818114>.
- (19) Hull, C. M.; Bevilacqua, P. C. Discriminating Self and Non-Self by RNA: Roles for RNA Structure, Misfolding, and Modification in Regulating the Innate Immune Sensor PKR. *Acc. Chem. Res.* **2016**, *49* (6), 1242–1249. <https://doi.org/10.1021/acs.acscounts.6b00151>.
- (20) Uehata, T.; Takeuchi, O. RNA Recognition and Immunity—Innate Immune Sensing and Its Posttranscriptional Regulation Mechanisms. *Cells* **2020**, *9* (7), 1701. <https://doi.org/10.3390/cells9071701>.
- (21) Yamagami, R.; Sieg, J. P.; Bevilacqua, P. C. Functional Roles of Chelated Magnesium Ions in RNA Folding and Function. *Biochemistry* **2021**, *60* (31), 2374–2386. <https://doi.org/10.1021/acs.biochem.1c00012>.
- (22) Yamagami, R.; Bingaman, J. L.; Frankel, E. A.; Bevilacqua, P. C. Cellular Conditions of Weakly Chelated Magnesium Ions Strongly Promote RNA Stability and Catalysis. *Nat. Commun.* **2018**, *9* (1), 2149. <https://doi.org/10.1038/s41467-018-04415-1>.
- (23) Yamagami, R.; Huang, R.; Bevilacqua, P. C. Cellular Concentrations of Nucleotide Diphosphate-Chelated Magnesium Ions Accelerate Catalysis by RNA and DNA Enzymes. *Biochemistry* **2019**, *58* (38), 3971–3979. <https://doi.org/10.1021/acs.biochem.9b00578>.
- (24) Leamy, K. A.; Yamagami, R.; Yennawar, N. H.; Bevilacqua, P. C. Single-Nucleotide Control of tRNA Folding Cooperativity under near-Cellular Conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (46), 23075–23082. <https://doi.org/10.1073/pnas.1913418116>.
- (25) Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D. Absolute Metabolite Concentrations and Implied Enzyme Active Site Occupancy in *Escherichia Coli*. *Nat. Chem. Biol.* **2009**, *5* (8), 593–599. <https://doi.org/10.1038/nchembio.186>.
- (26) Martell, A. E.; Smith, R. M. *Critical Stability Constants*; New York; London: Plenum Press, 1974; Vol. V1.
- (27) Martell, A. E.; Smith, R. M. *Critical Stability Constants*; New York, Plenum Press, 1974; Vol. V2.
- (28) Martell, A. E.; Smith, R. M. *Critical Stability Constants*; New York, Plenum Press, 1974; Vol. V3.

- (29) Martell, A. E.; Smith, R. M. *Critical Stability Constants*; New York; London : Plenum, 1982; Vol. V5.
- (30) Smith, R. M. (Robert M.; Martell, A. E. *Critical Stability Constants*; New York; London : Plenum, 1989; Vol. V6.
- (31) Berthon, G. Critical evaluation of the stability constants of metal complexes of amino acids with polar side chains (Technical Report). *Pure Appl. Chem.* **1995**, *67* (7), 1117–1240. <https://doi.org/10.1351/pac199567071117>.
- (32) Martell, A. E.; Smith, R. M. *NIST46 Critically Selected Stability Constants of Metal Complexes*. NIST. <https://www.nist.gov/srd/nist46> (accessed 2021-05-25).
- (33) Grilley, D.; Soto, A. M.; Draper, D. E. Chapter 3 Direct Quantitation of Mg<sup>2+</sup>-RNA Interactions by Use of a Fluorescent Dye. In *Methods in Enzymology: Biothermodynamics*, Part A; Academic Press, 2009; Vol. 455, pp 71–94. [https://doi.org/10.1016/S0076-6879\(08\)04203-1](https://doi.org/10.1016/S0076-6879(08)04203-1).
- (34) Puglisi, J. D.; Tinoco, I. [22] Absorbance Melting Curves of RNA. In *Methods in Enzymology: RNA Processing Part A: General Methods*; Academic Press, 1989; Vol. 180, pp 304–325. [https://doi.org/10.1016/0076-6879\(89\)80108-9](https://doi.org/10.1016/0076-6879(89)80108-9).
- (35) Xia, T.; SantaLucia, J.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson–Crick Base Pairs †. *Biochemistry* **1998**, *37* (42), 14719–14735. <https://doi.org/10.1021/bi9809425>.
- (36) Liu, B.; Shankar, N.; Turner, D. H. Fluorescence Competition Assay Measurements of Free Energy Changes for RNA Pseudoknots. *Biochemistry* **2010**, *49* (3), 623–634. <https://doi.org/10.1021/bi091541j>.
- (37) Liu, B.; Diamond, J. M.; Mathews, D. H.; Turner, D. H. Fluorescence Competition and Optical Melting Measurements of RNA Three-Way Multibranch Loops Provide a Revised Model for Thermodynamic Parameters. *Biochemistry* **2011**, *50* (5), 640–653. <https://doi.org/10.1021/bi101470n>.
- (38) You, Y.; Tataurov, A. V.; Owczarzy, R. Measuring Thermodynamic Details of DNA Hybridization Using Fluorescence. *Biopolymers* **2011**, *95* (7), 472–486. <https://doi.org/10.1002/bip.21615>.
- (39) Adamala, K.; Szostak, J. W. Non-Enzymatic Template-Directed RNA Synthesis inside Model Protocells. *Science* **2013**, *342* (6162), 1098–1100. <https://doi.org/10.1126/science.1241888>.
- (40) Soukup, G. A.; Breaker, R. R. Relationship between Inter-nucleotide Linkage Geometry and the Stability of RNA. *RNA* **1999**, *5* (10), 1308–1325.
- (41) Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R. Riboswitches Control Fundamental Biochemical Pathways in *Bacillus Subtilis* and Other Bacteria. *Cell* **2003**, *113* (5), 577–586. [https://doi.org/10.1016/S0092-8674\(03\)00391-X](https://doi.org/10.1016/S0092-8674(03)00391-X).
- (42) Batey, R. T.; Gilbert, S. D.; Montange, R. K. Structure of a Natural Guanine-Responsive Riboswitch Complexed with the Metabolite Hypoxanthine. *Nature* **2004**, *432* (7015), 411. <https://doi.org/10.1038/nature03037>.
- (43) Gilbert, S. D.; Love, C. E.; Edwards, A. L.; Batey, R. T. Mutational Analysis of the Purine Riboswitch Aptamer Domain †. *Biochemistry* **2007**, *46* (46), 13297–13309. <https://doi.org/10.1021/bi700410g>.
- (44) Poudyal, R. R.; Sieg, J. P.; Portz, B.; Keating, C. D.; Bevilacqua, P. C. RNA Sequence and Structure Control Assembly and Function of RNA Condensates. *RNA* **2021**, *27* (12), 1589–1601. <https://doi.org/10.1261/rna.078875.121>.
- (45) Park, J. O.; Rubin, S. A.; Xu, Y.-F.; Amador-Noguez, D.; Fan, J.; Shlomi, T.; Rabinowitz, J. D. Metabolite Concentrations, Fluxes and Free Energies Imply Efficient Enzyme Usage. *Nat. Chem. Biol.* **2016**, *12* (7), 482–489. <https://doi.org/10.1038/nchembio.2077>.
- (46) Knight, C. J.; Hub, J. S. WAXSiS: A Web Server for the Calculation of SAXS/WAXS Curves Based on Explicit-Solvent Molecular Dynamics. *Nucleic Acids Res.* **2015**, *43* (W1), W225–W230. <https://doi.org/10.1093/nar/gkv309>.
- (47) Klozin, A.; Oltsch, F.; Harmon, T.; Honigmann, A.; Jülicher, F.; Hyman, A. A.; Zechner, C. Phase Separation Provides a Mechanism to Reduce Noise in Cells. *Science* **2020**, *367* (6476), 464–468. <https://doi.org/10.1126/science.aav6691>.
- (48) Riback, J. A.; Zhu, L.; Ferrolino, M. C.; Tolbert, M.; Mitrea, D. M.; Sanders, D. W.; Wei, M.-T.; Kriwacki, R. W.; Brangwynne, C. P. Composition-Dependent Thermodynamics of Intracellular Phase Separation. *Nature* **2020**, *581* (7807), 209–214. <https://doi.org/10.1038/s41586-020-2256-2>.
- (49) Lambert, D.; Draper, D. E. Effects of Osmolytes on RNA Secondary and Tertiary Structure Stabilities and RNA-Mg<sup>2+</sup> Interactions. *J. Mol. Biol.* **2007**, *370* (5), 993–1005. <https://doi.org/10.1016/j.jmb.2007.03.080>.
- (50) Pogram, L. M.; Wendorff, T.; Erdmann, R.; Shkel, I.; Bellissimo, D.; Felitsky, D. J.; Record, M. T. Why Hofmeister Effects of Many Salts Favor Protein Folding but Not DNA Helix Formation. *Proc. Natl. Acad. Sci.* **2010**, *107* (17), 7716–7721. <https://doi.org/10.1073/pnas.0913376107>.
- (51) Lambert, D.; Draper, D. E. Denaturation of RNA Secondary and Tertiary Structure by Urea: Simple Unfolded State Models and Free Energy Parameters Account for Measured m-Values. *Biochemistry* **2012**, *51* (44), 9014–9026. <https://doi.org/10.1021/bi301103j>.
- (52) Cheng, X.; Shkel, I. A.; Molzahn, C.; Lambert, D.; Karim, R.; Record, M. T. Quantifying Interactions of Nucleobase Atoms with Model Compounds for the Peptide Backbone and Glutamine and Asparagine Side Chains in Water. *Biochemistry* **2018**, *57* (15), 2227–2237. <https://doi.org/10.1021/acs.biochem.8b00087>.
- (53) Williams, A. M.; Dickson, T.; Lagoa-Miguel, C.; Bevilacqua, P. C. Biological Solution Conditions and Flanking Sequence Modulate LLPS of RNA G-Quadruplex Structures. *RNA* **2022**, *rna.079196.122*. <https://doi.org/10.1261/rna.079196.122>.