JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2019JG005435

Key Points:
- Mass balances were used to assess if microbial transformation explains incomplete recovery of resazurin-resorufin in field studies
- Experiments conducted with different microorganisms indicate complete mass recovery during all conversion stages
- Complete mass recovery at the microbial scale suggests sorption is responsible for incomplete recovery in mesocosm and field experiments

Supporting Information:
- Supporting Information S1

Correspondence to:
R. González-Pinzón,
gonzaric@unm.edu

Citation:
Dallan, E., Regier, P., Marion, A., & González-Pinzón, R. (2020). Does the mass balance of the reactive tracers resazurin and resorufin close at the microbial scale?. *Journal of Geophysical Research: Biogeosciences*, *125*, e2019JG005435. https://doi.org/10.1029/2019JG005435

Received 22 AUG 2019
Accepted 8 FEB 2020
Accepted article online 11 FEB 2020

Does the Mass Balance of the Reactive Tracers Resazurin and Resorufin Close at the Microbial Scale?

Eleonora Dallan[1,2], Peter Regier[2], Andrea Marion[1], and Ricardo González-Pinzón[2]

[1]Department of Industrial Engineering, University of Padua, Padua, Italy, [2]Department of Civil, Construction and Environmental Engineering, University of New Mexico, Albuquerque, NM, USA

Abstract Resazurin (Raz) is a phenoxazine dye that can be reduced irreversibly to the daughter compound resorufin (Rru) by aerobic respiration. Previous hydrologic studies using the Raz-Rru reactive tracer system to quantify water-sediment interactions and metabolic activity have reported that dilution-corrected masses of Raz and Rru recovered are smaller than the mass of Raz injected. This lack of mass balance closure has been reported as a nonideality of this tracer system and, to date, it is still unclear what drives incomplete recovery. We used controlled laboratory experiments varying the initial concentrations of Raz, the duration of the experiments, and the type of microbial communities present to quantify mass balances of Raz and Rru under conditions that removed other suspected causes of incomplete recovery in field experiments, i.e., sorption to sediments and photodecay. We used the summation of Raz and Rru concentrations over time to assess mass recovery and variability and found mass recoveries in the range of 85.6–110.4%, with a maximum standard deviation of 7.5%. In three of the four experiments, no strong temporal trend in mass recovery is present. In an experiment with *Bacillus subtilis* bacteria, lower recovery and evidence of a temporal trend in recovery only occurred after 13 hr past the complete transformation of Raz (i.e., beyond the duration of most field experiments). These results suggest that the lack of mass recovery in field studies is likely associated with physical or chemical mechanisms rather than biological interactions with the Raz-Rru tracer system.

Plain Language Summary Resazurin (Raz) is a fluorescent dye that can be transformed into resorufin (Rru) by microorganisms. It has been used for a decade to study how water interacts with sediments and microbial communities. However, most of these studies recover less Raz and Rru than the Raz that was added, prompting the question: Does the mass balance of Raz and Rru close at the microbial scale? We conducted multiple experiments with different microbes to answer this question. Our results generally show full mass recovery, meaning that the incomplete recovery reported in field studies is governed by physical or chemical factors.

1. Introduction

Resazurin (Raz) is a redox-sensitive phenoxazine dye frequently used to estimate biological activity. In appropriate reducing conditions, Raz (blue) irreversibly loses an oxygen ion to become resorufin (Rru). Rru (pink) also can undergo a further reduction to colorless dihydroresorufin, but this reaction is reversible by sample exposure to oxygen. Importantly for research in aquatic ecosystems, the Raz-Rru tracer system mimics the behavior of a binary nano-switch capable of indicating water interactions with metabolic hot-spots. Raz, the tracer injected, begins in state 0 when it is added and remains in that state unless it enters a reducing environment where it is irreversibly transformed to Rru and, after that, registers state 1. States 1 and 0 can be identified at low concentrations (currently parts per billion) using fluorescent signatures. Field and laboratory studies have shown that the transformation of Raz in filtered water is negligible compared to that in sediments (Haggerty et al., 2008, 2009). Also, studies have shown that Raz can be reduced to Rru by strict aerobes, facultative anaerobes, aerotolerant, and microaerophile organisms but not by strict anaerobes and thus preferentially indicates aerobic metabolism (Guerin et al., 2001; González-Pinzon et al., 2012, 2014; Karakashev et al., 2003; Knapp et al., 2018; Mariscal et al., 2009; McNicholl et al., 2007; Min & Kang, 2011; O'Brien et al., 2000; Strotmann et al., 1993; Ziegler et al., 2011). Combined, these ideal properties of the Raz-Rru system have helped hydrologists and ecologists quantify parameters and fluxes related to exchange with, and storage within, metabolically active transient storage zones in the near

subsurface (e.g., benthic, hyporheic, and riparian zones) (González-Pinzón et al., 2014, 2015, 2016; Haggerty et al., 2009; Knapp et al., 2017, 2018).

Like most hydrologic tracers, Raz and Rru have nonidealities that affect their mass balance. First, it is unclear if the irreversible transformation of Raz uniquely yields Rru or also other unidentified by-products that are not fluorescent (O'Brien et al., 2000) and thus cannot be quantified through fluorescence spectroscopy, which is currently the most accurate method available to quantify Raz and Rru. Second, both Raz and Rru undergo sorption, particularly at lower pH values (Lemke et al., 2014), and this may cause transient mass retention operating at timescales longer than the duration of field studies, effectively causing incomplete recovery of the tracers. Haggerty et al. (2009) and Lemke et al. (2014) conducted nonequilibrium (kinetic) and equilibrium sorption analyses and found that linear sorption models are adequate but emphasized differences between the sorption distribution coefficients of Raz and Rru. Third, being fluorescent tracers, Raz and Rru may undergo photodecay. However, studies have shown that the timescales of photodecay are several tens of hours for Rru and hundreds of hours for Raz (Haggerty et al., 2008, 2009), suggesting that it is usually negligible for the duration of most field studies.

As a result of nonidealities, most hydrologic studies where the Raz-Rru system has been used have consistently found that the total mass recovered (Raz and Rru) is smaller than the mass of Raz injected after accounting for dilution. Also, this difference has been typically larger than the uncertainty in the quantification of the tracer concentrations. For example, Haggerty et al. (2009) reported 15% loss of mass, Argerich et al. (2011) found mass loss of 62.8%, Stanaway et al. (2012) obtained a mass balance range of 49–77% in column experiments, and Yakirevich et al. (2017) reported mass recovery of 13.7% (3.9% Raz + 9.8% Rru). To date, it is unclear if the unclosed Raz mass balance can be explained by the existence of additional reaction products that are produced during the transformation of Raz to Rru, the long-term sorption of the tracers, some degree of photodecay, all combined, or some other mechanism of uptake at the microbial scale.

This study seeks to answer the question: Does the mass balance of Raz and Rru close at the microbial scale? In our study, we minimized the effects that sorption and photodecay may have on the mass recovery of the tracers and used different microbial communities to investigate recovery patterns independent of specific microbial species. Our findings indicate a near complete mass recovery at the microbial scale and suggest that incomplete recovery is likely related to sorption and transient storage processes retaining the tracers over timescales longer than the field experiments.

2. Materials and Methods

2.1. Microbial-scale Experiments

In our experiments, we used (1) yeast cells (*Saccharomyces cerevisiae*) from commercial dry yeast as generic facultative anaerobe cells, (2) naturally present microorganisms from the Rio Grande (surface water sampled at locations near Albuquerque, New Mexico, USA), and (3) *Bacillus subtilis* as facultative anaerobes, soil- and water-dwelling bacteria with demonstrated success in previous Raz-Rru studies (González-Pinzon et al., 2012). Yeast cells and *B. subtilis* were used to determine responses in specific cell experiments, and the microbial communities from the Rio Grande were used to capture the responses of more complex communities.

All the experiments had the same experimental setup, except for the preparation of the initial solutions used with the different types of microorganisms. Yeast solution was prepared from 0.8 g of dry yeast and 8.0 g of sugar in 4 L of ultrapure water (18 MΩ) at room temperature (~21°C). This solution was stirred for 15 min to dissolve the sugar and activate the yeast. For the experiment with unfiltered river water, ~10 L of surface water was collected from the Rio Grande and stored overnight to allow suspended solids to settle out. The day of the experiment, a volume of about 4 L was collected from the superficial layer of the settled water. For the experiment using *B. subtilis*, cells were purchased from ATCC (ATCC® 23857) as freeze-dried stock, revived following manufacturer recommendations (i.e., hydrated in 10 ml of ATCC® medium 415), used to inoculate autoclaved vials containing medium 415, and incubated at 26°C for 24 hr. The final solution for the experiment was prepared using a total of 200 ml of incubated cultures, in a solution of 500 ml medium and 3.3 L of ultrapure water at 26°C to encourage the highest possible levels of bacterial activity. Although the other experiments were all started at room temperature, the *B. subtilis* experiment was started at 26°C to avoid shocking the incubated cultures and was allowed to return to room temperature (21°C) through the experiment.

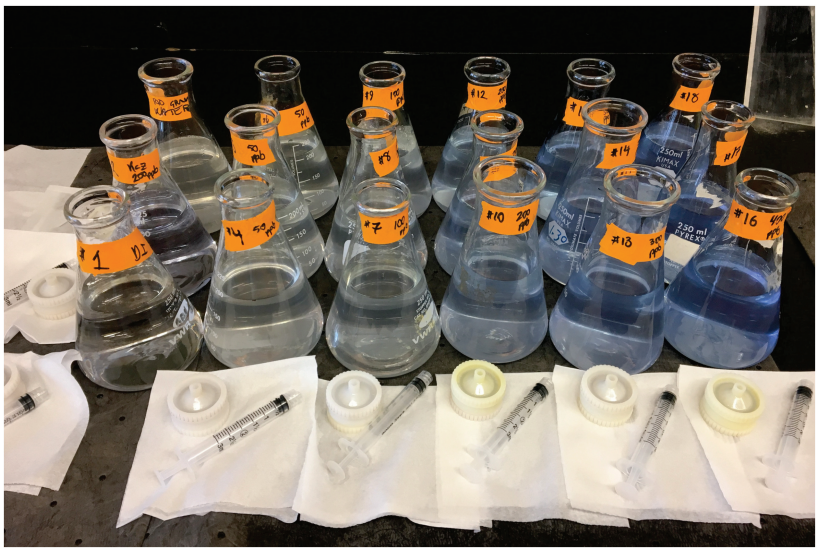

Figure 1. Experimental setup after Raz additions: 18 flasks include three controls (flasks # 1-2-3) and 5 Raz concentrations series, each with three replicates (flasks # 4–18).

The setup for each experiment was composed of 18 flasks, each with a starting volume of 205 ml, following the design presented in Figure 1 and Table 1. The 4 L of culture solution was split into 200 ml of aliquots and placed into acid-washed 250 ml Erlenmeyer flasks. Three flasks were used as controls in each experiment: one with medium only, one with medium and 200 ppb of Raz but no culture, and one with medium and culture but no Raz. The remaining 15 flasks were dosed as triplicates of 5 Raz concentrations (50, 100, 200, 300, and 400 ppb). Flasks were wrapped in foil to eliminate potential photodecay, placed on a rotary shaker table, and agitated at 110 rpm for the duration of the experiment to minimize the potential for anaerobic conditions. Dosing flasks was conducted as quickly as possible (within 15 min) to minimize variability between flasks and facilitate sampling as soon as possible after dosing to establish baseline readings. Four different experiments were carried out in this study (Table 2). The first two experiments used the yeast solution, where the first (named *Yeast-5h*) was run for about 5 hr and the second (named *Yeast-71h*) for about 3 days. The third experiment, using settled water from the Rio Grande (named *River*), was carried out for about 4 days, and the fourth experiment, conducted with *B. subtilis* (named *B. Sub*.), lasted about 1.5 days. The *Yeast-5h* experiment was a proof of concept, which allowed us to validate our experimental design and sampling protocols. Because full Raz conversion was not reached during this experiment, we conducted the *Yeast-71h* experiment. In the third and fourth experiments, we sampled until complete Raz conversion (*B. Sub*.) or until logistical constraints required ending the experiment (*River*).

Table 1
Experimental Setup: Individual Experiment Arrangement

Flask number	Main solution	Raz initial concentration (ppb)
1	Medium[a]	0
2	Medium + Raz	200
3	Initial solution[b]	0
4-5-6	Initial solution + Raz	50
7-8-9	Initial solution + Raz	100
10-11-12	Initial solution + Raz	200
13-14-15	Initial solution + Raz	300
16-17-18	Initial solution + Raz	400

[a]Medium is the nonreactive solution used as control, made with ultrapure water for *Yeast-5h*, *Yeast-71h*, and *River* experiments and broth (ATCC® medium 415) for the *B. Sub*. experiment (see Table 2). [b]Initial solution is the combination of the respective medium and cells.

2.2. Experimental Sampling, Storage, and Readings

The experiments were conducted at room temperature (21°C). Samples were taken after the addition of Raz at increasing time intervals, with higher frequency during the first 5–6 hr (subhourly) and with decreasing frequency after that. For *Yeast-71h*, sampling started after 5 hr in an attempt to evaluate recovery behavior beyond the timescales measured in *Yeast-5h*. For each sample, a small aliquot (~3 ml) was removed from each flask, with 2 ml filtered through 0.2 μm polycarbonate filters (Whatman) into cuvettes and buffered with 1 M sodium phosphate to reach pH of 8.5 (González-Pinzón et al., 2012; Haggerty et al., 2008). All samples were refrigerated immediately after collection until analysis. Samples were analyzed using a Varian Cary Eclipse fluorescence spectrophotometer (Agilent), and all samples for a given experiment were analyzed on the same day to minimize day-to-day variation in instrument performance.

Table 2
Experimental Setup: Type of Cells, Concentrations, and Experiment Durations across Experiments

Experiment	Initial solution	Raz initial concentrations (ppb)	Experiment duration
Yeast-5h	Yeast	0, 50, 100, 200, 300, 400	5.3 hr
Yeast-71h	Yeast	0, 50, 100, 200, 300, 400	71 hr
River	Rio Grande water	0, 50, 100, 200, 300, 400	91 hr
B. Sub.	*B. subtilis*	0, 50, 100, 200, 300, 400	31 hr

2.3. Data Analysis and Mass Recovery

The concentrations of Raz and Rru were calculated from the fluorescence readings relative to standard curves using available MATLAB codes (Knapp et al., 2018) to obtain Raz and Rru concentration values, $C^{\mathrm{Raz}}_{i,j,k}$ and $C^{\mathrm{Rru}}_{i,j,k}$ expressed in (μmol/L), where index i = 1–5 indicates concentration series (50, 100, 200, 300, and 400 ppb, respectively), index j = 1–3 represents the relevant flask number for each triplicate (see Table 1), and index k refers to the sample number, corresponding to the time the sample was taken (see tables S2–S5 in "Recovery_data.xlsx" file in Dallan et al., 2020).

The total concentration of a flask at a given sampling time was calculated as

$$C^{\mathrm{Tot}}_{i,j,k} = C^{\mathrm{Raz}}_{i,j,k} + C^{\mathrm{Rru}}_{i,j,k}. \tag{1}$$

2.4. Statistics

Mean values of Raz, Rru, and combined (Raz + Rru) concentrations for sampling time k were calculated using the three replicates (j = 1,2,3) of the same initial concentration series i:

$$\widetilde{C}^{\mathrm{tracer}}_{i,k} = \frac{\sum_{j=1}^{3} C^{\mathrm{tracer}}_{i,j,k}}{3}, \tag{2}$$

where tracer represents either Raz, Rru or the summation of both (i.e., Raz + Rru, referred to as Tot).

The mean total concentration for each initial concentration series i, for all sampling times k, was

$$\widetilde{C}^{\mathrm{Tot}}_{i} = \frac{\sum_{k=1}^{n} \widetilde{C}^{\mathrm{Tot}}_{i,k}}{n}, \tag{3}$$

where n is the number of samples for each experiment.

For each initial concentration series i, we quantified the sample variability within comparable flasks (i.e., for j = 1,2,3) for each sampling event k, i.e., $v_{i,k}$, and across sampling events (k = 1,2,...,n), i.e., v_i, as

$$v_{i,k} = \frac{std\left(C^{\mathrm{Tot}}_{i,j,k}\right)}{\widetilde{C}^{\mathrm{Tot}}_{i,k}}, \tag{4}$$

$$v_{i} = \frac{std\left(\widetilde{C}^{\mathrm{Tot}}_{i,k}\right)}{\widetilde{C}^{\mathrm{Tot}}_{i}}. \tag{5}$$

Recovery percentages were estimated for each initial concentration series i, at sampling event k as

$$R^{\mathrm{tracer}}_{i,k} = 100 \cdot \frac{\widetilde{C}^{\mathrm{tracer}}_{i,k}}{\widetilde{C}^{\mathrm{Tot}}_{i}}. \tag{6}$$

The variability associated with $R^{\mathrm{tot}}_{i,k}$ was calculated for each series i, i.e., V_i, and for each complete experiment (i.e., V for all five series i), as

$$V_i = \mathrm{std}\left(R^{\mathrm{Tot}}_{i,k}\right), \tag{7}$$

$$V = \mathrm{mean}(V_i). \tag{8}$$

We examined several metrics to statistically quantify the mass recovery of Raz injected compared to the Raz + Rru retrieved. First, we conducted experiments to evaluate errors associated with sample preparation and spectrofluorometer readings (see Supporting Information S1, section "Error analysis"). The standard deviation (σ) associated with these errors was 5.1%, and this value was used to establish a statistical

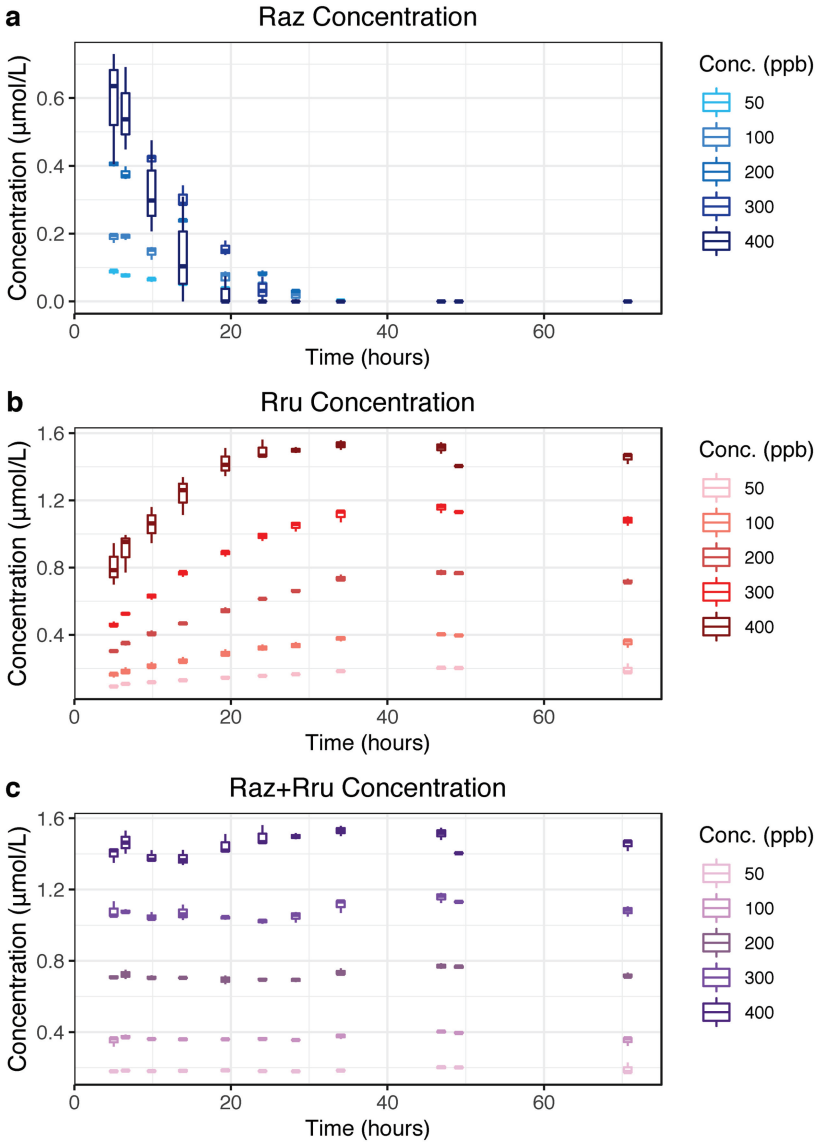

Figure 2. Example of the evolutions of (a) Raz, (b) Rru,and (c) Raz + Rru molar concentrations $C^{\mathrm{Raz}}_{i,j,k}$, $C^{\mathrm{Rru}}_{i,j,k}$, $C^{\mathrm{Tot}}_{i,j,k}$ during the *Yeast-71h* experiment (Table 3). Flasks with different starting concentrations of Raz are shown in different colors. Boxplots represent the variability among triplicate flasks for each Raz concentration, with whiskers representing experiment-specific standard deviations.

complete recovery range of 100% ± 5.1%. Therefore, values greater than 100% ± 10.2% (2σ) are interpreted as incomplete recovery or invalid. We also calculated standard deviations for the results of each experiment, following Equations (7) and (8), to compare the relative variation within a given experiment to other experiments. Finally, we assessed trends by fitting regression lines and report regression statistics in the results and discussion. Outliers were identified by total concentration variability $v_{i,k}$ larger than 100 ± 15.3% (3σ), with a total of three values in *Yeast-71h* removed, all associated with improper sample preparation or handling (# 16-17-18 at t = 49.1 hr).

3. Results and Discussion

3.1. Mass Recovery

Figure 2 shows an example of the evolution of the concentrations of Raz, Rru, and total molar concentrations (Raz + Rru) obtained for the long experiment with yeast (*Yeast-71h*). Equivalent plots for *Yeast-5h*, *River*,

Table 3
Summary of Recovery Results for Each Experiment

Yeast-5h	**Raz 50**	**Raz 100**	**Raz 200**	**Raz 300**	**Raz 400**
$\widetilde{C}_i^{\mathrm{Tot}}$ (μmol/L)	0.204	0.409	0.802	1.188	1.572
$min\ R_{i,k}^{\mathrm{tot}}$ (%)	96.5%	96.4%	98.3%	98.0%	96.9%
$max\ R_{i,k}^{\mathrm{tot}}$ (%)	101.9%	102.9%	102.1%	103.0%	101.9%
V_i (%)	1.8%	2.2%	1.2%	1.8%	1.7%
V (%)			1.7%		
Yeast-71h	**Raz 50**	**Raz 100**	**Raz 200**	**Raz 300**	**Raz 400**
$\widetilde{C}_i^{\mathrm{Tot}}$ (μmol/L)	0.995	1.018	1.005	0.999	1.006
$min\ R_{i,k}^{\mathrm{tot}}$ (%)	96.3%	95.1%	96.2%	94.7%	94.4%
$max\ R_{i,k}^{\mathrm{tot}}$ (%)	107.7%	109.0%	107.0%	107.3%	105.1%
V_i (%)	4.3%	4.7%	3.8%	3.8%	3.8%
V (%)			4.0%		
River	**Raz 50**	**Raz 100**	**Raz 200**	**Raz 300**	**Raz 400**
$\widetilde{C}_i^{\mathrm{Tot}}$ (μmol/L)	0.203	0.398	0.784	1.160	1.533
$min\ R_{i,k}^{\mathrm{tot}}$ (%)	93.2%	95.7%	97.6%	97.6%	96.3%
$max\ R_{i,k}^{\mathrm{tot}}$ (%)	103.3%	106.4%	106.8%	103.2%	104.3%
V_i (%)	2.7%	2.9%	2.3%	1.7%	2.3%
V (%)			2.3%		
B. Sub.	**Raz 50**	**Raz 100**	**Raz 200**	**Raz 300**	**Raz 400**
$\widetilde{C}_i^{\mathrm{Tot}}$ (μmol/L)	0.949	1.013	1.011	1.009	1.044
$min\ R_{i,k}^{\mathrm{tot}}$ (%)	88.5%	87.8%	87.7%	85.6%	85.7%
$max\ R_{i,k}^{\mathrm{tot}}$ (%)	109.5%	109.8%	110.2%	110.4%	110.1%
V_i (%)	7.4%	7.5%	7.5%	8.5%	7.8%
V (%)			7.5%		

Note. Mean Raz + Rru concentration $\widetilde{C}_i^{\mathrm{Tot}}$ for each series i, minimum and maximum recovery percentages $R_{i,k}^{\mathrm{tot}}$ for each series i, recoveries variability V_i for each series i, and recovery variability V for each set of experiments with a given microbial group.

and *B. Sub*. are presented in Supporting Information S1 (Figures S1–S3). Consistent patterns of decreasing Raz concentration were generally observed through time, with steeper slopes for higher concentrations. Similarly, Rru concentrations increased, while Raz + Rru concentrations were steady (Figure 2).

Recovery percentage statistics are summarized in Table 3 (see Tables S2–S5 in the "Recovery_data.xlsx" file in Dallan et al., 2020 for full data sets and statistics), and recovery time series are plotted in Figure 3, along with complete recovery ranges, experiment-specific standard deviations, and linear regression trendlines. For the *Yeast-5h* experiment, the mass recovery $R_{i,k}^{\mathrm{tot}}$ ranged from 96.4% to 103.0%, with a mean standard deviation V of 1.7% (Table 3). All values were within the complete recovery range ($\pm\sigma$), and no significant trend was observed ($R^2 = 0.00$, $p = 0.7831$), indicating complete recovery throughout the experiment.

The *Yeast-71h* experiment expanded the timescale of *Yeast-5h* and showed mass recoveries $R_{i,k}^{\mathrm{tot}}$ in the range 94.4–109.0% and a standard deviation V of 3.9%. The standard deviation is less than the standard deviation for experimental errors (5.1%), and most individual points fell within the range of $\pm\sigma$, with no values greater than $\pm 2\sigma$. Statistically, there is a weak trend ($R^2 = 0.16$, $p = 0.0026$) driven by abnormally increasing values between hours 30 and 50, but this behavior disappeared before the end of the experiment. As such, *Yeast-71h* generally suggests mass closure in timescales approaching the longest published field experiments conducted with Raz (e.g., Argerich et al., 2011; González-Pinzón et al., 2014). In the two experiments with yeast cells, their durations had no effect on the recovery results; i.e., for both of them, the standard deviation was lower than the defined standard deviation of errors, and the recovery percentages did not show trends. The weak trend in the longest experiment, as considered above, is likely driven by two higher values around 50 hr.

For the *River* experiment, we obtained recoveries $R_{i,k}^{\mathrm{tot}}$ in the range of 93.2–106.8%, with a mean standard deviation V of 2.3%. As for experiments *Yeast-5h* and *Yeast-71h*, this value is within the complete recovery range. Moreover, only 3 of 75 measurements are greater than $\pm\sigma$, with no values greater than $\pm 2\sigma$. A low R^2 and relatively high p value indicate that no significant trend is present over time ($R^2 = 0.04$, $p = 0.0797$). As such, the *River* experiment consistently exhibits complete recovery of Raz using natural microbial communities, even at multiday timescales.

Finally, mass recoveries $R_{i,k}^{\mathrm{tot}}$ in experiment *B. Sub*. were in the range 85.6–110.4%, with a standard deviation V of 7.5%. This experiment was characterized by higher variability than the previous three, and we observed incomplete recovery toward the end of the experiment (after 25 hr). Unlike the previous three experiments, the standard deviation for the experiment is higher than σ, suggesting high variability during the experiment and several points fell outside the $\pm 2\sigma$ range (although no points fell outside the $\pm 3\sigma$ range). As shown in Figure 3 and quantified by linear regression of mass recovery over time ($R^2 = 0.53$, $p < 0.0001$), the higher standard deviation is driven by incomplete recovery toward the end of the experiment. During the first ~20 hr of Exp. 4, a similar pattern to experiment *Yeast-71h* is observed, where values increase and then return to baseline, yielding no linear trend ($R^2 = 0.00$, $p = 0.7722$, $n = 45$) and a standard deviation (4.5%) less than the standard deviation of the experimental error.

From the four experiments conducted, mass recoveries $R_{i,k}^{\mathrm{tot}}$ ranged from 85.6% to 110.4%, with a maximum standard deviation for a single time series of 8.5% (Table 3). Across the four experiments, the average standard deviation was 3.9%, which is lower than the experimentally determined variability within the experimental error of 5.1%. As such, findings from this study generally suggest that the mass balance of

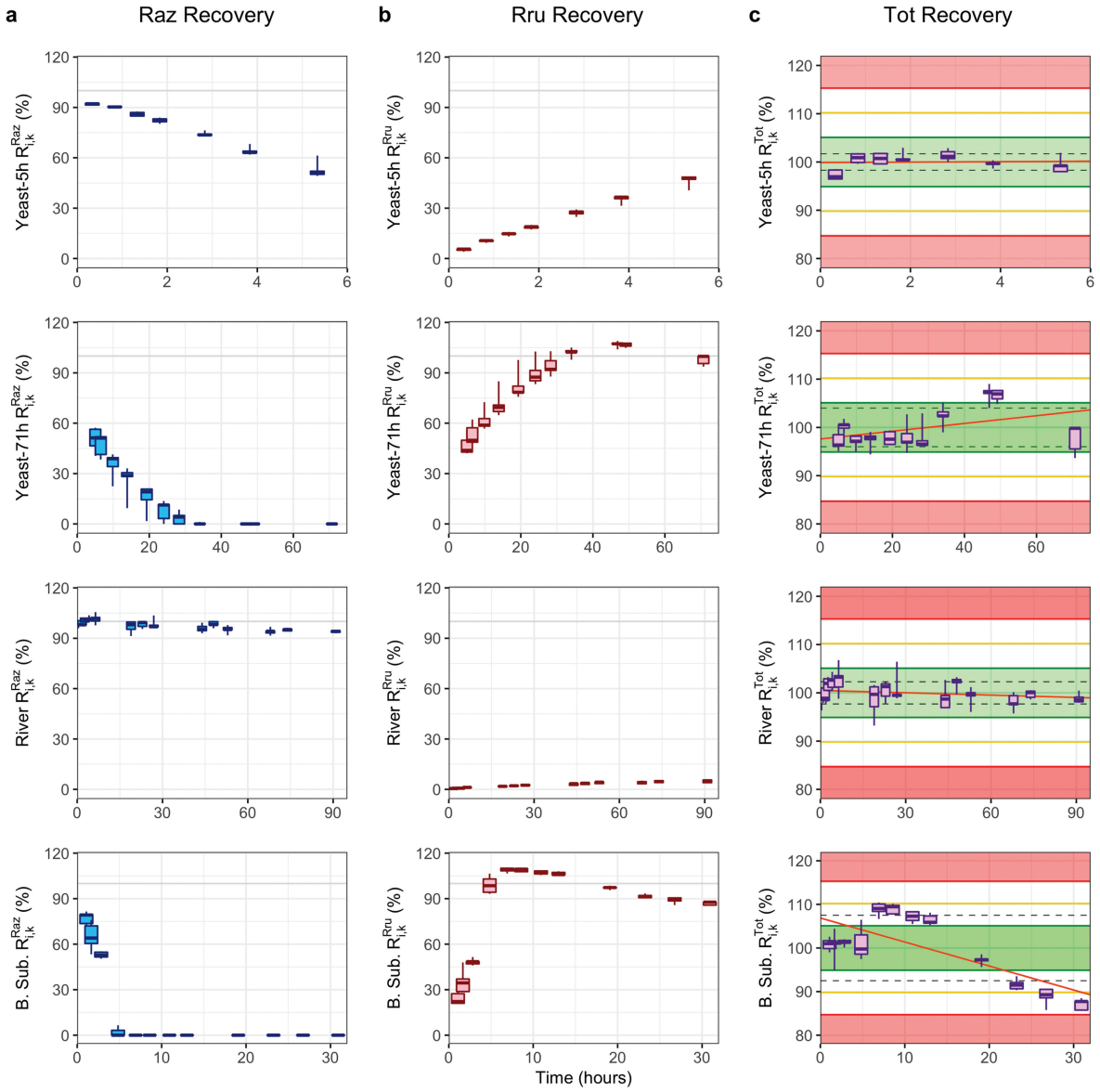

Figure 3. Percent recoveries $R^{tracer}_{i,k}$ for the four experiments for (a) Raz, (b) Rru, and (c) Raz + Rru (Tot). Each boxplot includes all Raz concentrations for each sampling point. Panel c is color coded to recovery, where the green-shaded region represents the range of complete recovery ($\pm\sigma$), while yellow lines show $\pm2\sigma$ (suggesting incomplete recovery). Red-shaded regions represent $\pm3\sigma$. Dashed black lines represent experiment-specific standard deviations. Solid red lines show best-fit linear regression lines.

the Raz-Rru system closes at the microbial scale. The *B. Sub*. experiment indicated lower recoveries (85.6–95.6%, table S5 in "Recovery_data.xlsx" file in Dallan et al., 2020) starting after ~13 hr of total Raz consumption. This timescale is arguably much longer than most field experiments (typically <8 hr), particularly considering that microbial-scale experiments, unlike field experiments, feature near-perfect mixing conditions and contact times between solutes and bacteria that approach the duration of the experiments. Furthermore, *B. Sub*. was conducted under conditions optimized for bacterial activity (e.g., removal of nutrient and temperature limitations), which are not common in natural systems. In contrast to *B. Sub*., the experiment conducted with natural microbial communities (*River*) showed very low Raz transformation and Rru production, even after 90 hr. The low values for the conversion of Raz to Rru obtained in our *River* experiment may be partially explained by lower microbial activity during the winter,

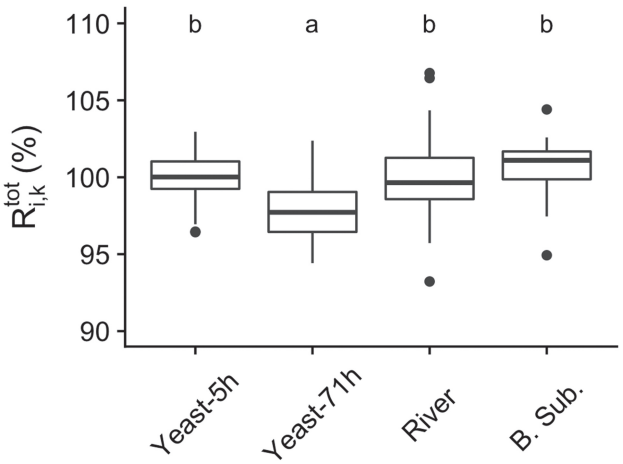

Figure 4. Percent recoveries $R^{tot}_{i,k}$ for the four experiments. Each boxplot includes all Raz + Rru recovery percentages for all sampling points, while Raz concentration was greater than 0 for a given experiment. Letters indicate results of Turkey-Kramer tests comparing the means for each pair of experiments, where different letters indicate significantly ($p < 0.05$) different means.

when the samples were collected, and strongly suggest that the conditions required to replicate the incomplete recovery observed in the *B. Sub.* experiment do not represent the activity of natural microbial communities observed during the *River* experiment.

To test for differences between cell types, we applied the Tukey-Kramer statistical test to subsets of each experiment when Raz concentrations were above 0. This allowed us to compare recovery during the period of active Raz transformation across each experiment. We found that *Yeast-5h*, *River*, and *B. Sub.* experiments were not significantly different, showing that our results were independent of cell type (Figure 4). *Yeast-71h* was significantly lower than the other three experiments for this subset of data ($p < 0.05$), due to slightly lower recovery (average: 97.8%) driven by two abnormally higher values around 50 hr. However, based on the experimental error present, we consider this difference negligible for all practical purposes.

3.2. Anomalies in Recovery Time Series and Potential Drivers of Incomplete Raz-Rru Mass Recovery

In the experiments *Yeast-71h* and *B. Sub.*, we observed slightly higher recovery percentages $R^{tot}_{i,k}$ after complete Raz consumption (i.e., after 30 hr for *Yeast-71h* and 5 hr for *B. Sub.*, Figure 3). We estimated a recovery increase of about 7% from the difference between the mean recovery percentage before Raz total consumption (97.9% for *Yeast-71h* between 5.0 and 28.3 hr and 100.9% for *B. Sub.* between 1.1 and 4.8 hr) and the mean recovery percentage for the peak (105.2% for *Yeast-71h* between 34 and 49.1 hr and 107.9% for *B. Sub.* between 6.9 and 13.0 hr). These percentages greater than 100% in *Yeast-71h* and *B. Sub.* are likely an artifact of the calculation of $\widetilde{C}^{tot}_i$, considering that recoveries $R^{tot}_{i,k}$ are calculated with respect to the mean concentration $\widetilde{C}^{tot}_i$ (Equation (6)), and lower recoveries toward the end would lower $\widetilde{C}^{tot}_i$.

In addition, the apparent incomplete recovery at the end of *Yeast-71h* and *B. Sub.* matches findings by O'Brien et al. (2000) and Chen et al. (2018), who demonstrated that a secondary reaction from Rru to hydroresorufin is favored after all Raz has been reduced (reducing total detectable Raz + Rru) and that this reaction can happen both inside and outside cells. Another study using recellularized kidney scaffolds observed a fluorescence peak that gradually decreased throughout the assay, decaying to zero by 24 hr, demonstrating that cells were capable of secondary reduction of Rru to hydroresorufin (Uzarski et al., 2017). Moreover, Vehniäinen et al. (2012) demonstrated that in the liver S9 fraction of some fish species, Rru was enzymatically reduced to a less fluorescent form. Together, these findings suggest that the decreasing Rru concentrations observed after complete Raz consumption in *Yeast-71h* and *B. Sub.* are indicative of additional nonfluorescent product formation that may not be completely associated with hydroresorufin formation, since sample exposure to oxygen during preparation and analysis did not result in a fully reversible reaction to Rru, as Karakashev et al. (2003) reported would happen for hydroresorufin. To advance current understanding, we recommend further investigation of the reversible Rru-hydroresorufin reaction pathway to find an accurate method to quantify and correct for currently unmeasured by-product formation.

In light of the mass balance closure generally observed in the four experiments conducted in this study, we do not find compelling evidence suggesting that microbial activity governs the lack of mass balance closure commonly observed in field studies. We note that our experimental design did not explicitly explore the role of benthic or hyporheic biofilm dynamics, where sorption and rate-limited mass transfer may occur. However, the fact that we observed complete mass recoveries suggests that biofilm formation played a minimum role in the *Yeast-5h*, *Yeast-71h*, and *River* experiments but may have played a role near the end of the *B. Sub.* experiments. Our results suggest that the incomplete Raz-Rru recovery observed in field and mesocosm experiments is largely associated with physical or chemical processes. Both Raz and Rru undergo photochemical decay, but Haggerty et al. (2008) demonstrated that the timeframe for photodecay is tens of hours for Raz and hundreds of hours for Rru. As such, minimizing photodecay during field experiments only requires storing samples in the dark, which is standard practice. Another potential explanation for

incomplete recovery that both laboratory and field studies most strongly support is sorption to particulates (Lemke et al., 2014; Knapp & Cirpka, 2017), particularly when the sorption timescales are much longer than the experimental timescales. Unfortunately, while an in situ or mesocosm experiment containing a sediment-water interface may more accurately represent the metabolic activity in stream compartments, such experiment would be affected by sorption onto sediments and would obscure microbial-scale responses.

4. Conclusions

We conducted laboratory experiments to investigate mass balance closure of the Raz-Rru system at the microbial scale. Four experiments were performed, with three different microbial groups (yeast cells, native microbial communities, and pure *B. subtilis* cells), different initial concentrations of Raz (0–400 ppb), and different duration of the experiments (5–91 hr). Our experiments generally indicated full mass recovery, even under optimal conditions for microbial activity. One experiment conducted with *B. subtilis* exhibited incomplete mass recovery (<15% loss) after 25 hr due to an unmeasurable transformation of Rru that occurred after all Raz was transformed; however, to date, a complete transformation of Raz to Rru has not been reported in any mesocosm or field studies. Natural microbial communities present in water collected from the Rio Grande converted less than 10% of Raz to Rru after 91 hr and exhibited complete mass recovery throughout. Interestingly, these results support previous studies showing that the conversion of Raz to Rru is negligible in the water column and that this fact can be used to investigate the extent of reactive transport in different compartments of the stream network (e.g., main channel vs. surface transient storage zones and benthic and hyporheic zones). Combining our results, which suggest complete mass recovery at the microbial scale, with those from other characterizations of nonidealities of the Raz-Rru tracer system (i.e., photodecay and sorption), we conclude that sorption is likely the main driver of incomplete mass recovery in field and mesocosm studies.

Acknowledgments
This material is based upon work supported by the National Science Foundation under Award Nos. EAR-1642399, HRD-1345169, and HRD-1720912. All data presented in the manuscript are freely available on CUAHSI-HIS repository (http://www.hydroshare.org/resource/291c24a5ce3f41579eb53b2950fb5a61). The authors thank Holly Wilkie, James Fluke, Justin Nichols, and Jancoba Dorley for assistance in conducting laboratory experiments.

References

Argerich, A., Haggerty, R., Martí, E., Sabater, F., & Zarnetske, J. (2011). Quantification of metabolically active transient storage (MATS) in two reaches with contrasting transient storage and ecosystem respiration. *Journal of Geophysical Research*, *116*, G03034. https://doi.org/10.1029/2010JG001379

Chen, J. L., Steele, T. W. J., & Stuckey, D. C. (2018). Metabolic reduction of resazurin; location within the cell for cytotoxicity assays. *Biotechnology and Bioengineering*, *115*, 351–358. https://doi.org/10.1002/bit.26475

Dallan, E., P. Regier, A. Marion, R. González-Pinzón (2020). Does the mass balance of the reactive tracers resazurin and resorufin close at the microbial scale?, HydroShare, http://www.hydroshare.org/resource/291c24a5ce3f41579eb53b2950fb5a61

González-Pinzón, R., Haggerty, R., & Argerich, A. (2014). Quantifying spatial differences in metabolism in headwater streams. *Freshwater Science*, *33*(3), 798–811. https://doi.org/10.1086/677555

González-Pinzón, R., Haggerty, R., & Myrold, D. D. (2012). Measuring aerobic respiration in stream ecosystems using the resazurin-resorufin system. *Journal of Geophysical Research*, *117*, G00N06. https://doi.org/10.1029/2012JG001965

González-Pinzón, R., Peipoch, M., Haggerty, R., Martí, E., & Fleckenstein, J. H. (2016). Nighttime and daytime respiration in a headwater stream. *Ecohydrology*, *9*(1), 93–100.

González-Pinzón, R., Ward, A. S., Hatch, C. E., Wlostowski, A. N., Singha, K., Gooseff, M. N., et al. (2015). A field comparison of multiple techniques to quantify groundwater–surface-water interactions. *Freshwater Science*, *34*(1).

Guerin, T., Mondido, M., McClenn, B., & Peasley, B. (2001). Application of resazurin for estimating abundance of contaminant-degrading microorganisms. *Letters in Applied Microbiology*, *32*(5), 340–345. https://doi.org/10.1046/j.1472765X.2001.00916.x

Haggerty, R., Argerich, A., & Martí, E. (2008). Development of a "smart" tracer for the assessment of microbiological activity and sediment-water interaction in natural waters: The resazurin-resorufin system. *Water Resources Research*, *44*, W00D01. https://doi.org/10.1029/2007WR006670

Haggerty, R., Martí, E., Argerich, A., von Schiller, D., & Grimm, N. (2009). Resazurin as a "smart" tracer for quantifying metabolically active transient storage in stream ecosystems. *Journal of Geophysical Research*, *114*, G03014. https://doi.org/10.1029/2008JG000942

Karakashev, D., Galabova, D., & Simeonov, I. (2003). A simple and rapid test for differentiation of aerobic from anaerobic bacteria. *World Journal of Microbiology and Biotechnology*, *19*, 233–238. https://doi.org/10.1023/A:1023674315047

Knapp, J. L. A., & Cirpka, O. A. (2017). Determination of hyporheic travel time distributions and other parameters from concurrent conservative and reactive tracer tests by local-in-global optimization. *Water Resources Research*, *53*(6), 4984–5001. https://doi.org/10.1002/2017WR020734

Knapp, J. L. A., González-Pinzón, R., Drummond, J. D., Larsen, L. G., Cirpka, O. A., & Harvey, J. W. (2017). Tracer-based characterization of hyporheic exchange and benthic biolayers in streams. *Water Resources Research*, *53*(2), 1575–1594. https://doi.org/10.1002/2016WR019393

Knapp, J. L. A., González-Pinzón, R., & Haggerty, R. (2018). The resazurin-resorufin system: Insights from a decade of "smart" tracer development for hydrologic applications. *Water Resources Research*, *54*(9), 6877–6889. https://doi.org/10.1029/2018WR023103

Lemke, D., González-Pinzón, R., Liao, Z., Wohling, T., Osenbruck, K., Haggerty, R., & Cirpka, O. A. (2014). Sorption and transformation of the reactive tracers resazurin and resorufin in natural river sediments. *Hydrology and Earth System Sciences*, *18*(8), 3151–3163. https://doi.org/10.5194/hess-18-3151-2014

Mariscal, A., Lopez-Gigosos, R., Carnero-Varo, M., & FernandezCrehuet, J. (2009). Fluorescent assay based on resazurin for detection of activity of disinfectants against bacterial biofilms. *Applied Microbiology and Biotechnology*, *82*(4), 773–783. https://doi.org/10.1007/s00253-009-1879-x

McNicholl, B. P., McGrath, J. W., & Quinn, J. P. (2007). Development and application of a resazurin-based biomass activity test for activated sludge plant management. *Water Research*, *41*(1), 127–133. https://doi.org/10.1016/j.watres.2006.10.002

Min, T. G., & Kang, W. S. (2011). Simple, quick and nondestructive method for Brassicaceae seed viability measurement with single seed base using resazurin, Hortic. *Horticulture, Environment and Biotechnology*, *52*(3), 240–245. https://doi.org/10.1007/s13580-011-0182-9

O'Brien, J., Wilson, I., Orton, T., & Pognan, F. (2000). Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *European Journal of Biochemistry*, *267*(17), 5421–5426. https://doi.org/10.1046/j.1432-1327.2000.01606.x

Stanaway, D., Haggerty, R., Benner, S., Flores, A., & Feris, K. (2012). Persistent metal contamination limits lotic ecosystem heterotrophic metabolism after more than 100 years of exposure: A novel application of the resazurin-resorufin smart tracer. *Environmental Science & Technology*, *46*, 9862–9871. https://doi.org/10.1021/Es3015666

Strotmann, U. J., Butz, B., & Bias, W. R. (1993). A dehydrogenase assay with resazurin-practical performance as a monitoring-system and pH-dependent toxicity of phenolic-compounds. *Ecotoxicology and Environmental Safety*, *25*(1), 79–89. https://doi.org/10.1006/eesa.1993.1009

Uzarski, J. S., DiVito, M. D., Wertheim, J. A., & Miller, W. M. (2017). Essential design considerations for the resazurin reduction assay to noninvasively quantify cell expansion within perfused extracellular matrix scaffolds. *Biomaterials*, *129*, 163–175. https://doi.org/10.1016/j.biomaterials.2017.02.015

Vehniäinen, E., Schultz, E., Lehtivuori, H., Ihalainen, J., & Oikari, A. (2012). More accuracy to the EROD measurements—The resorufin fluorescence differs between species and individuals. *Aquatic Toxicology*, 116–117. doi.org/10.1016/j.aquatox.2012.03.007

Yakirevich, A., Shelton, D., Hill, R., Kiefer, L., Stocker, M., Blaustein, R., et al. (2017). Transport of conservative and "smart" tracers in a first-order creek: Role of transient storage type. *Water*, *9*, 485.

Ziegler, V. G., Knaup, J., Stahl, D., Krammer, B., & Plaetzer, K. (2011). Fluorescence detection and depletion of T47D breast cancer cells from human mononuclear cell-enriched blood preparations by photodynamic treatment: Basic in vitro experiments towards the removal of circulating tumor cells. *Lasers in Surgery and Medicine*, *42*, 548.