Does It Pay Off to Explicitly Link Functional Gene Expression to Denitrification Rates in Reaction Models?

Anna Störiko[1]*, Holger Pagel[2], Adrian Mellage[1] and Olaf A. Cirpka[1]*

[1] *Center for Applied Geoscience, University of Tübingen, Tübingen, Germany,* [2] *Biogeophysics, Institute of Soil Science and Land Evaluation, University of Hohenheim, Stuttgart, Germany*

Environmental omics and molecular-biological data have been proposed to yield improved quantitative predictions of biogeochemical processes. The abundances of functional genes and transcripts relate to the number of cells and activity of microorganisms. However, whether molecular-biological data can be quantitatively linked to reaction rates remains an open question. We present an enzyme-based denitrification model that simulates concentrations of transcription factors, functional-gene transcripts, enzymes, and solutes. We calibrated the model using experimental data from a well-controlled batch experiment with the denitrifier *Paracoccous denitrificans*. The model accurately predicts denitrification rates and measured transcript dynamics. The relationship between simulated transcript concentrations and reaction rates exhibits strong non-linearity and hysteresis related to the faster dynamics of gene transcription and substrate consumption, relative to enzyme production and decay. Hence, assuming a unique relationship between transcript-to-gene ratios and reaction rates, as frequently suggested, may be an erroneous simplification. Comparing model results of our enzyme-based model to those of a classical Monod-type model reveals that both formulations perform equally well with respect to nitrogen species, indicating only a low benefit of integrating molecular-biological data for estimating denitrification rates. Nonetheless, the enzyme-based model is a valuable tool to improve our mechanistic understanding of the relationship between biomolecular quantities and reaction rates. Furthermore, our results highlight that both enzyme kinetics (i.e., substrate limitation and inhibition) and gene expression or enzyme dynamics are important controls on denitrification rates.

Keywords: denitrification, biogeochemical modeling, functional genes, transcripts, enzyme-based modeling, bayesian inference

1. INTRODUCTION

Modern high-throughput molecular-biological techniques and omics methods provide insights into the abundance, activity, and metabolic function of microorganisms in environmental systems (Bouchez et al., 2016; Starke et al., 2019). Because microbial transformations play a key role in biogeochemical cycling (Kuypers et al., 2018) and contaminant degradation (Dong et al., 2019),

reactive-transport models require an adequate quantification of microbial dynamics. Numerous laboratory and field studies (Bælum et al., 2008; Bowen et al., 2014; Anantharaman et al., 2016; Wegner et al., 2019) indicate that quantifying the abundance of genes and gene transcripts has the potential to improve the description of microbially mediated reactions implemented in reactive-transport models.

Transcript-to-gene ratios and relative gene expression levels have been suggested as a direct measure of microbial activity (Nicolaisen et al., 2008; Freitag and Prosser, 2009; Brow et al., 2013; Monard et al., 2013; Rohe et al., 2020). However, assuming a unique (e.g., linear) relationship neglects important factors that modulate reaction rates such as post-transcriptional and post-translational regulation, different time-scales of transcript and enzyme dynamics, and substrate limitation (Moran et al., 2013). Therefore, integrating a mechanistic description of the aforementioned processes into models could help to better establish a mechanistic link between transcript concentrations and reaction rates. Most reactive-transport models do not explicitly account for genes or transcription and thus cannot integrate molecular-biological datasets into quantitative validation frameworks. Addressing this challenge and integrating omics data into reactive-transport modeling has been proposed as a way to improve our understanding of the dynamic behavior of biogeochemical systems (Li et al., 2017a).

In a first approach to integrate genetic information, genome-scale metabolic models have been coupled to reactive transport, successfully predicting reaction rates under variable environmental conditions without the need of calibration (Scheibe et al., 2009). Though computationally intensive, genome-scale metabolic networks provide a powerful approach for cases in which reactions are dominated by a single, well-known organism or a small group of organisms, as in the case of uranium reduction. Other studies have focused on a single metabolic pathway of an organism, simulating the regulatory chain, including transcription factors, mRNA, enzymes, substrate consumption, and growth kinetics (Koutinas et al., 2011; Bælum et al., 2013).

Natural biogeochemical cycles, such as the nitrogen cycle, are mediated by a diverse group of organisms, of which only a small percentage can be cultivated. Under these circumstances, genome-scale models and models focusing on the functionality of a single microbial strain become impractical. Instead, approaches that target metabolic pathways rather than processes at the scale of a single cell may capture the explicit contribution of the non-cultivatable majority within a natural microbial community. Thus, the alternative *gene-centric* approach focuses on the functionality represented by certain genes instead of specific organisms. By using functional-gene abundances as a proxy for biomass with a certain function, gene-centric approaches can easily be integrated into existing biogeochemical and reactive-transport models (Reed et al., 2014, 2015). Genomic data provide information on metabolic potential, but not on actual microbial activity. To account for activity, the gene-centric approach has been further developed to predict concentrations of functional transcripts and enzymes (Louca et al., 2016; Li et al., 2017c;

Song et al., 2017). In one of the studies, the model was validated using metatranscriptomic and metaproteomic data (Louca et al., 2016). However, because the approach of these authors computes transcript and enzyme concentrations by postprocessing of reactions rates, it only provides a partially mechanistic link between rates, transcription, and enzyme production. Hence, this approach lacks the explicit integration of transcript and enzyme regulatory feedbacks into the biogeochemical model.

The functional-enzyme-based approach, by contrast, simulates concentrations of specific enzymes based on energetic (Li et al., 2017c) or cybernetic (Song and Liu, 2015) considerations, and enzyme concentrations directly regulate reaction rates. None of these modeling approaches represent the actual mechanisms of transcriptional regulation with transcription factors, having the advantage that they do not require special knowledge about the regulatory system. However, the validation of the enzyme-based modeling approach relies on quantitative enzyme data that remain challenging to obtain, particularly for environmental samples. Conversely, the quantification of mRNA via reverse transcription quantitative polymerase chain reaction (RT-qPCR) is well established for both laboratory and field setups.

While reactive-transport models that integrate molecular-biological data are more mechanistic, they unavoidably are more complex than traditional Monod-type formulations. Whether the added complexity actually improves model predictability remains understudied.

In this study, we integrate transcript data into a denitrification reaction model by linking the expression of functional genes to process rates. We explicitly account for transcriptional regulation of denitrification via transcription factors, translating the current conceptual understanding of the regulatory system in *Paracoccus denitrificans* into a quantitative model. The main pathways of nitrogen transformations—denitrification, nitrification, N-fixation, annamox, and dissimilatory nitrate reduction to ammonium (DNRA)—have been extensively studied due to the relevance of reactive nitrogen compounds for ecosystem functioning (Steffen et al., 2015), groundwater contamination (Gutiérrez et al., 2018), eutrophication (Howarth, 2008), and greenhouse gas emissions (Liu and Greaver, 2009). The nitrogen cycle is thus an ideal test case for developing and testing new, enzyme-based models informed by measurements of functional genes and transcripts. Previous gene-centric and enzyme-based modeling studies focused on systems with slow dynamics (kinetics) (Li et al., 2017c; Song et al., 2017; Chavez Rodriguez et al., 2020) or at steady state (Reed et al., 2014; Louca et al., 2016). Here, we apply our model to a dynamic reactive system, that is, one with shifts in the predominant electron accepting species, and inform it with a highly temporally resolved dataset of transcript abundances. Previous studies have highlighted a potentially hysteretic relationship between transcript concentrations and reaction rates (Bælum et al., 2008; Chavez Rodriguez et al., 2020). Our model allows us to further explore the relationship between transcripts, enzymes, and reaction rates and develop mechanistic interpretations of the observations. Via a comparison with a classical Monod-type model formulation that does not integrate

FIGURE 1 | Schematic representation of processes considered in the models. **(A)** Model system and processes implemented common to both the Monod-type and enzyme-based models. **(B)** Simplified representation of gene expression and its link to reaction rates in the enzyme-based model. Narrow arrows represent inducing effects, whereas lines ending in a square denote inhibition. In the model, NO_2^- instead of NO is used as activating compound for NNR since we do not simulate NO concentrations.

transcriptional regulation, we shed light on the potential added benefits of integrating transcript data into (denitrification) reaction models.

2. METHODS

2.1. Conceptual Model Description

We set up a model to simulate the experiments of Qu et al. (2015) performed in well-mixed batch reactors. Qu et al. (2015) monitored aerobic respiration and denitrification coupled to succinate oxidation by the denitrifying organism *P. denitrificans*. Briefly, a series of batch reactors, inoculated with *P. denitrificans*, prepared in an aerobic medium, were amended with 5 mM succinate and 2 mM of NO_3^-. Automated headspace gas measurements (nitrogen gases, O_2) as well as aqueous-phase measurements of nitrite and optical density (OD) facilitated monitoring time series of concentrations driven by reaction and growth kinetics. In addition, and partly driven by headspace gas concentrations, cells were periodically harvested to measure the concentrations of the functional genes *narG*, *nirS*, *norB*, and *nosZ*. Following mRNA extraction, transcript numbers were determined via RT-qPCR using the standard curve method. For further details regarding the experimental procedures, we refer to the original publication (Qu et al., 2015).

In the experiment, the contribution of the intermediates NO and N_2O to the mass balance was always less than 1‰. From this experimental observation, we conclude that nitrite reduction was the rate-limiting step. Therefore, we set up a simplified model of denitrification simulating the specific experiments by assuming a two-step process (**Figure 1**). Therein, the reduction of nitrite via NO and N_2O to N_2 was treated as a single reaction step. Denitrification was coupled to the oxidation of succinate as the sole carbon source and electron donor:

$$7NO_3^- + C_4H_6O_4 \xrightarrow{narG} 7NO_2^- + 4CO_2 + 3H_2O \,, \tag{1}$$

$$14NO_2^- + 14H^+ + 3C_4H_6O_4 \xrightarrow{nirS} 7N_2 + 12CO_2 + 16H_2O \,. \tag{2}$$
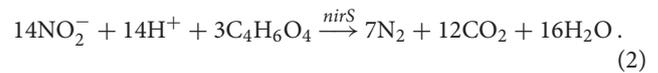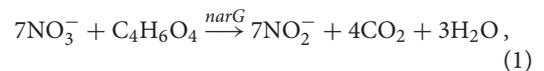
In the reaction equations above, the names of the functional genes linked to the reduction of the nitrogen compounds by *P. denitrificans* are given above the arrows. In the presence of oxygen, aerobic respiration is energetically favored over denitrification:

$$2C_4H_6O4 + 7O_2 \longrightarrow 8CO_2 + 6H_2O \,. \tag{3}$$

As shown in **Figure 1**, the presence of oxygen is assumed to inhibit denitrification. The electron donor, succinate, was assumed to be primarily assimilated for energy conservation by *P. denitrificans*, and thus we do not consider the incorporation of carbon into biomass during growth. Furthermore, the model considers mass transfer between the liquid and the gas phases via a linear driving force approximation, and the dilution of the gas-phase concentrations by sampling (see section 5 in the **Supplementary Material**).

We set up two models for comparison: (1) an enzyme-based and (2) a standard Monod-type model. The enzyme-based model considers transcripts and enzymes involved in denitrification reactions as state variables. Reaction rates are directly proportional to enzyme concentrations. The Monod-type model is much simpler and describes denitrification rates using the Monod equation. Therein, the regulation mechanisms and dynamics of catalyzing enzymes are not explicitly reflected, instead the reaction rates are assumed to be directly proportional to the biomass of the denitrifying bacteria.

2.2. Governing Equations

Our model represents the transcriptional regulation of denitrification genes by simulating transcription factor concentrations in response to oxygen and nitrogen substrates and their effect on transcript concentrations. For comparison, we also set up a simplified model that omits the explicit representation of transcription factors so that transcription directly depends on the concentrations of signaling molecules (see section 6 in the **Supplementary Material**). This simplified model can also be interpreted mechanistically (as discussed in section 6.3 of the **Supplementary Material**), but the assumptions regarding transcriptional regulation differ slightly from the ones presented in the following section.

Several transcription factors regulate the transcription of denitrification genes in *P. denitrificans*, sensing oxygen and nitrogen oxides. The *narG* gene is activated by FnrP and NarR, whereas transcription of *nirS* requires NNR (Gaimster et al., 2018). The FnrP protein directly reacts with oxygen that leads to its inactivation (Crack et al., 2016). NarR responds to nitrate and nitrite, although the underlying mechanism is currently unknown and the sensing might be indirect (Wood et al., 2001; Spiro, 2017). NNR is activated by NO and deactivated by O_2 (Lee et al., 2006; Gaimster et al., 2018).

We assume that the sum of active and inactive transcription factors remains constant throughout the experiment and simulate all active transcription factors relative to the total concentration. Genes that encode the transcriptional regulators (*fnrP*, *narR*, and *nnrR*) have been shown to be expressed at similar levels under aerobic and anaerobic conditions in *P. denitrificans* (Giannopoulos et al., 2017), supporting our assumption. We model the fraction of active NarR, X_{NarR} [−], as

$$\frac{dX_{NarR}}{dt} = \left(a_{NarR}^{NO_3^-} C_{NO_3^-} + a_{NarR}^{NO_2^-} C_{NO_2^-} \right) (1 - X_{NarR}) - k_{dec}^{NarR} X_{NarR}, \tag{4}$$

where $a_{NarR}^{NO_3^-}$ and $a_{NarR}^{NO_2^-}$ [M^{-1} s^{-1}] are the rate coefficients for activation of NarR by nitrate and nitrite, respectively, and k_{dec}^{NarR} [s^{-1}] is the dissociation constant of active NarR. The term $(1 - X_{NarR})$ represents the inactive fraction of the transcription factor. Binding of oxygen to the transcription factors is described at equilibrium using a Hill function (Gesztelyi et al., 2012). The active fraction of FnrP is

$$X_{FnrP} = \frac{I_{FnrP}^p}{I_{FnrP}^p + C_{O_2}^p}, \tag{5}$$

with inhibition constant I_{FnrP} [M] and Hill coefficient p [−]. Experiments have shown that NNR quickly reacts with oxygen, but that its reaction to NO is slower (Lee et al., 2006). We therefore model the inactivation process at equilibrium while accounting for temporal evolution in the activation process. Although NNR actually senses NO, we simulate its activation by nitrite as we do not explicitly model NO concentrations in our two-step representation of denitrification (see conceptual model description above and **Figure 1**). The activation of NNR is described as follows:

$$\frac{d\hat{X}_{NNR}}{dt} = a_{NNR} C_{NO_2^-} \left(1 - \hat{X}_{NNR} \right) - k_{dec}^{NNR} \hat{X}_{NNR}, \tag{6}$$

where $\hat{X}_{NNR}$ [−] is the fraction of NNR activated by nitrite (without accounting for inactivation by oxygen), a_{NNR} [M^{-1} s^{-1}] is the activation rate constant and k_{dec}^{NNR} [s^{-1}] is the dissociation constant of activated NNR. Taking oxygen inhibition into account, the active fraction of NNR X_{NNR} [−] is given by

$$X_{NNR} = \frac{I_{NNR}^q}{I_{NNR}^q + C_{O_2}^q} \hat{X}_{NNR}, \tag{7}$$

with oxygen inhibition constant I_{NNR} [M] and Hill coefficient q.

We assume that the transcription rate for gene i scales with the fraction of operator sites where all activating transcription factors are bound f_{act}^i (Ingalls, 2013):

$$r_{transcription}^i = \alpha_i f_{act}^i B, \tag{8}$$

where α_i [transcripts cell^{-1} s^{-1}] is the maximum transcription rate for gene i and B is the cell density [cells L^{-1}]. For *narG* and *nirS*, the fractions of active operator sites are given by equations 9 and 10, respectively:

$$f_{act}^{nar} = \frac{\frac{X_{FnrP} X_{NarR}}{K_{FnrP} K_{NarR}}}{1 + \frac{X_{FnrP}}{K_{FnrP}} + \frac{X_{NarR}}{K_{NarR}} + \frac{X_{FnrP} X_{NarR}}{K_{FnrP} K_{NarR}}}, \tag{9}$$

$$f_{act}^{nir} = \frac{X_{NNR}}{X_{NNR} + K_{NNR}}, \tag{10}$$

where K_{FnrP}, K_{NarR} and K_{NNR} (all dimensionless) are the half-saturation constants for transcription factor binding to the operator site relative to the total transcription factor concentration.

Translation of mRNA into enzymes is described by first-order kinetics:

$$r_{translation}^i = k_{translation}^i T_i, \tag{11}$$

in which $k_{translation}$ [enzymes transcript^{-1} s^{-1}] is the first-order translation coefficient and T_i [transcripts L^{-1}] is the transcript concentration of gene i. Enzymes (E) and transcripts (T) undergo first-order decay with decay coefficients k_{dec}^T [s^{-1}] and k_{dec}^E [s^{-1}], respectively:

$$r_{decay,T}^i = k_{dec}^T T_i, \tag{12}$$

$$r_{decay,E}^i = k_{dec}^E E_i. \tag{13}$$

Dynamics of transcripts are usually fast, with transcript half-lives of only a few minutes (Härtig and Zumft, 1999; Bernstein et al., 2002), compared to enzyme half-lives and cell doubling times of several hours (Blaszczyk, 1993; Maier et al., 2011). Therefore, for transcript concentrations, we assumed a quasi-steady state (qss) (Ingalls, 2013):

$$\frac{dT_i}{dt} = r_{transcription}^i - r_{decay,T}^i = 0, \tag{14}$$

which yields the quasi-steady state transcript concentrations T_i^{qss} [transcripts L^{-1}]:

$$T_i^{\text{qss}} = \beta_T^i f_{\text{act}}^i B, \tag{15}$$

where $\beta_T^i = \alpha_i / k_{\text{dec}}^T$ is the number of transcripts per cell at maximum transcription, that is when f_{act}^i equals 1. Enzyme concentrations are assumed to be governed by mRNA-translation as well as first-order decay:

$$\frac{dE_i}{dt} = r_{\text{translation}}^i - r_{\text{decay,E}}^i. \tag{16}$$

The translation rate constant is difficult to measure and the prior range of this parameter is therefore essentially unconstrained. However, we can express it in terms of parameters that are easier to estimate by computing the quasi-steady state (qss) of enzyme concentrations, E_i^{qss} [enzymes L^{-1}], given by

$$E_i^{\text{qss}} = \frac{k_{\text{translation}}^i \beta_T^i}{k_{\text{dec}}^E} f_{\text{act}}^i B. \tag{17}$$

We defined $\beta_E = \frac{k_{\text{translation}}^i \beta_T^i}{k_{\text{dec}}^E}$ as the maximum quasi-steady state enzyme concentration per cell, a parameter that is easier to constrain based on literature data. We can then express the translation rate constant as

$$k_{\text{translation}}^i = k_{\text{dec}}^E \frac{\beta_E}{\beta_T^i}. \tag{18}$$

The transformation rate of nitrogen substrate j (NO_3^- or NO_2^-) with corresponding enzyme i (NAR or NIR) is described by Michaelis–Menten kinetics considering kinetic inhibition of the reaction by O_2:

$$r_j = r_{\text{max}}^i \frac{C_j}{C_j + K_j} \frac{I_{\text{reac}}^i}{I_{\text{reac}}^i + C_{O_2}}, \tag{19}$$

in which the maximum rate constant $r_{\text{max}}^i \left[\text{mol L}^{-1} \text{s}^{-1}\right]$ is the theoretical rate at full substrate saturation without inhibition. The second and third term of equation 19 describe substrate limitation and oxygen inhibition of the enzyme, in which c_j [M] is the concentration of substrate j, K_j is the half-saturation constant [M] of substrate j and I_{reac}^i [M] is the oxygen-inhibition constant for the corresponding enzyme (i).

In a Monod-type model, the enzyme concentration inside a cell is assumed to be constant and r_{max} is proportional to the cell density:

$$r_{\text{max}}^i = v_{\text{max}}^i B, \tag{20}$$

in which v_{max}^i [mol $\text{cell}^{-1} \text{s}^{-1}$] represents a constant, cell-specific maximum substrate-consumption rate. In the enzyme-based model, the transformation rate of substrate i is given by

$$r_{\text{max}}^i = k_{\text{max}}^i E_i, \tag{21}$$

in which the turnover number k_{max}^i [s^{-1}] is the maximum amount of substrate that a single enzyme molecule can transform per unit time.

In both models (enzyme-based and Monod-type), aerobic respiration is described by a Monod rate law:

$$r_{O_2} = v_{\text{max}}^{O_2} B \frac{C_{O_2}}{C_{O_2} + K_{O_2}}, \tag{22}$$

with the maximum cell-specific rate $v_{\text{max}}^{O_2}$ [mol $\text{cell}^{-1} \text{s}^{-1}$] and half-saturation constant K_{O_2} [M]. *P. denitrificans* can utilize both aerobic respiration and denitrification for growth. However, typically reported growth yields for aerobic respiration are higher than those reported for denitrification (Boogerd et al., 1984) and, under the reported experimental conditions (Qu et al., 2015), growth using oxygen as the electron acceptor was the dominant process. Therefore, we assume that *P. denitrificans* only grows on aerobic respiration and that denitrification steps do not result in growth:

$$\frac{dB}{dt} = Y_{O_2} \frac{n_{C_4H_6O_4}}{n_{O_2}} r_{O_2}, \tag{23}$$

in which Y_{O_2} [cells mol^{-1} of succinate] is the growth yield and $n_{C_4H_6O_4} = 2$ and $n_{O_2} = 7$ are the stoichiometric coefficients of succinate and oxygen in the energy reaction, respectively.

The flasks used in the experiment had a headspace. We accounted for the mass transfer rates r_{tr}^i of gaseous compounds (N_2 and oxygen) between the water and the gas phase. Furthermore, we also considered dilution of the gas phase by sampling. Details can be found in the **Supplementary Material**. The dynamics of solute concentrations of oxygen and nitrogen species are given by.

$$\frac{dC_{O_2}}{dt} = -r_{O_2} + r_{\text{tr}}^{O_2}, \tag{24}$$

$$\frac{dC_{NO_3^-}}{dt} = -r_{NO_3^-}, \tag{25}$$

$$\frac{dC_{NO_2^-}}{dt} = r_{NO_3^-} - r_{NO_2^-}, \tag{26}$$

$$\frac{dC_{N_2}}{dt} = \frac{1}{2} r_{NO_2^-} + r_{\text{tr}}^{N_2}. \tag{27}$$

2.3. Parameter Identification

A subset of the model parameters was fixed, either because these parameters were known from the experimental setup, or they could be well constrained by literature values. In total, we estimated 14 parameters of the Monod-type model and 32 parameters of the enzyme-based model with a Bayesian approach. Given the evidence of observed data and prior knowledge, we obtain the *posterior* probability distribution of parameter values. Following Bayes' law, the conditional probability density of parameters $\boldsymbol{\theta}$ given the data $\boldsymbol{y}$, $p(\boldsymbol{\theta} \mid \boldsymbol{y})$, is proportional to the prior probability density $p(\boldsymbol{\theta})$ of the parameters and the likelihood of the measured data $p(\boldsymbol{y} \mid \boldsymbol{\theta})$:

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto p(\boldsymbol{y} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}). \tag{28}$$

The prior distribution constrains the parameters on knowledge uninformed by the data, such as literature values or physical constraints. The likelihood describes the probability density of the measured data if the model parameters were correct. For a perfect model, the likelihood only describes measurement errors, but in most applications it also include the effects of conceptual errors on meeting the data. We used the likelihood function

$$p(\mathbf{y} \mid \boldsymbol{\theta}) = \prod_{i,j} T(y_{ij} \mid \mu_{ij}, \sigma_{ij}, \nu), \qquad (29)$$

in which T is the Student's t-distribution with scale σ_{ij}, degrees of freedom ν, and location

$$\mu_{ij} = \text{boxcox}(c_{ij} + b_i, \lambda_i). \qquad (30)$$

The index i refers to the measured variable, j indicates the measurement number, and boxcox is the Box–Cox transform (Box and Cox, 1964). All measured data from the experiment were Box–Cox transformed to ensure homoscedasticity of the residuals. We added a constant background value b_i to the simulated concentration c_{ij} of oxygen and nitrite because measured concentrations never reached zero in the experiment, but stayed at a very low background value ($<10^{-7}$ M for NO_2^- and $<2 \times 10^{-7}$ M for O_2). The errors σ_{ij} are the sum of data errors computed from the standard deviations of triplicate measurements and a constant error that accounts for model structural errors, which we chose to be an estimation parameter. We chose a Student's t-distribution with $\nu = 10$ instead of a normal distribution so that outliers do not get too much weight.

The model was implemented in the programming language Python and the code is freely available (Störiko et al., 2021). We solved the system of ordinary differential equations (ODEs) using the Python package Sunode (Seyboldt, 2020), a Python wrapper to the CVODES library (Hindmarsh et al., 2005). Obtaining numerically stable results over a broad range of parameter values was difficult. Therefore, the ODE was log-transformed and solved using a backward differentiation formula of variable order (Byrne and Hindmarsh, 1975; Jackson and Sacks-Davis, 1980) with small tolerances (10^{-12} for the absolute and relative tolerance in the forward problem, 10^{-8} in the adjoint problem). Samples from the posterior distribution were drawn using the Python package PyMC3 (Salvatier et al., 2016) and analyzed with ArviZ (Kumar et al., 2019).

We sampled all models with the Hamiltonian MCMC algorithm No-U-Turn Sampler (NUTS; Hoffman and Gelman, 2014) as implemented in PyMC3, with adjusted settings for the tuning phase (Foreman-Mackey, 2020). It uses gradient information that we obtained by solving the adjoint sensitivity equations. We used the $\widehat{R}$-criterion (Vehtari et al., 2020) that compares the variance between different chains to the in-chain-variance to assess convergence. The largest $\widehat{R}$ is 1.005 for the Monod-type model and 1.021 for the enzyme-based model—the values close to one indicate convergence.

Supplementary Table 2 lists fixed parameter values, prior distributions, and statistics about the posterior distributions. **Supplementary Table 1** provides initial values of the simulations.

Enzyme concentrations were not measured in the experiment and were only indirectly constrained through reaction rates. As all rate laws contain the product of enzyme concentrations and the maximum enzyme-specific turnover rate $k_{\max}$, it is impossible to estimate both $k_{\max}$ and β_E without direct measurements of enzyme concentrations. Therefore, we fixed β_E to an arbitrary value of 1125 enzymes cell^{-1} based on measured intracellular enzyme concentrations in the literature (Maier et al., 2011). This implies that the fitted values of $k_{\max}$ and the simulated enzyme concentrations within bacterial cells have to be interpreted with care, as they are conditioned on the arbitrary choice of β_E.

3. RESULTS AND DISCUSSION

3.1. Modeled Concentration Time Series

Figure 2 shows a comparison of the simulated concentrations of the enzyme-based and the Monod-type models with the experimental data (Qu et al., 2015). **Figure 2A** shows the concentrations of nitrogen species (nitrate, nitrite, N$_2$) of the models and the experiments, **Figure 2B** shows the oxygen concentrations and cell densities, **Figure 2C** shows the transcripts of *narG* and *nirS*, respectively, whereas **Figure 2D** contains simulated enzyme concentrations for which no measurements were available. Fitted parameter values and their uncertainties are outlined in **Supplementary Table 2**, and simulated transcription factor concentrations are shown in **Supplementary Figure 4**.

Concentration times series results in **Figures 2A,B** are present as several draws of the posterior distribution for the enzyme based model and the median output for the Monod-type model. Results for the Monod-type model exhibited a similarly narrow spread of simulated concentrations, and thus multiple draws were omitted. Both model formulations fit the concentration data equally well. The excellent fit supports the validity of our conceptual model and the underlying assumptions outlined above. Oxygen is consumed before nitrate and nitrite are reduced to molecular nitrogen, which confirms the inhibition of denitrification by oxygen. The simulations predict that nitrate is fully converted to nitrite before the onset of the second denitrification step. Both models capture the sharp increase in nitrite concentrations, peaking after 25 h at 2 mM, followed by a gradual decrease over 40 h, and the subsequent production of N$_2$ gas, which is the final product of denitrification. The posterior distributions of solute and gas concentrations in both models were narrow, reflecting the small measurement variance. Even though nitrate concentrations were not measured by Qu et al. (2015), the nitrate simulation results are indirectly well constrained by nitrite data and the mass conservation assumption encoded in the model equations.

At the time of highest nitrite concentrations, the measured cell densities reached peak values and gradually decreased (**Figure 2B**). Neither of the models was able to capture the slightly non-monotonic behavior in cell densities with a peak at 25 h and a subsequent modest decline. The slight discrepancy between simulated and measured cell densities could be due to several assumptions of our model. We assumed that *P. denitrificans* only grows on oxygen, even though the organism is known to

FIGURE 2 | Simulated dynamics, 100 draws from the posterior of the enzyme-based model (solid lines) and posterior median of the Monod-type model (dashed lines), and measurements from three experimental replicates (open symbols). **(A)** Nitrogen compounds, **(B)** oxygen and cell densities, **(C)** transcripts, and **(D)** enzymes and maximum rates.
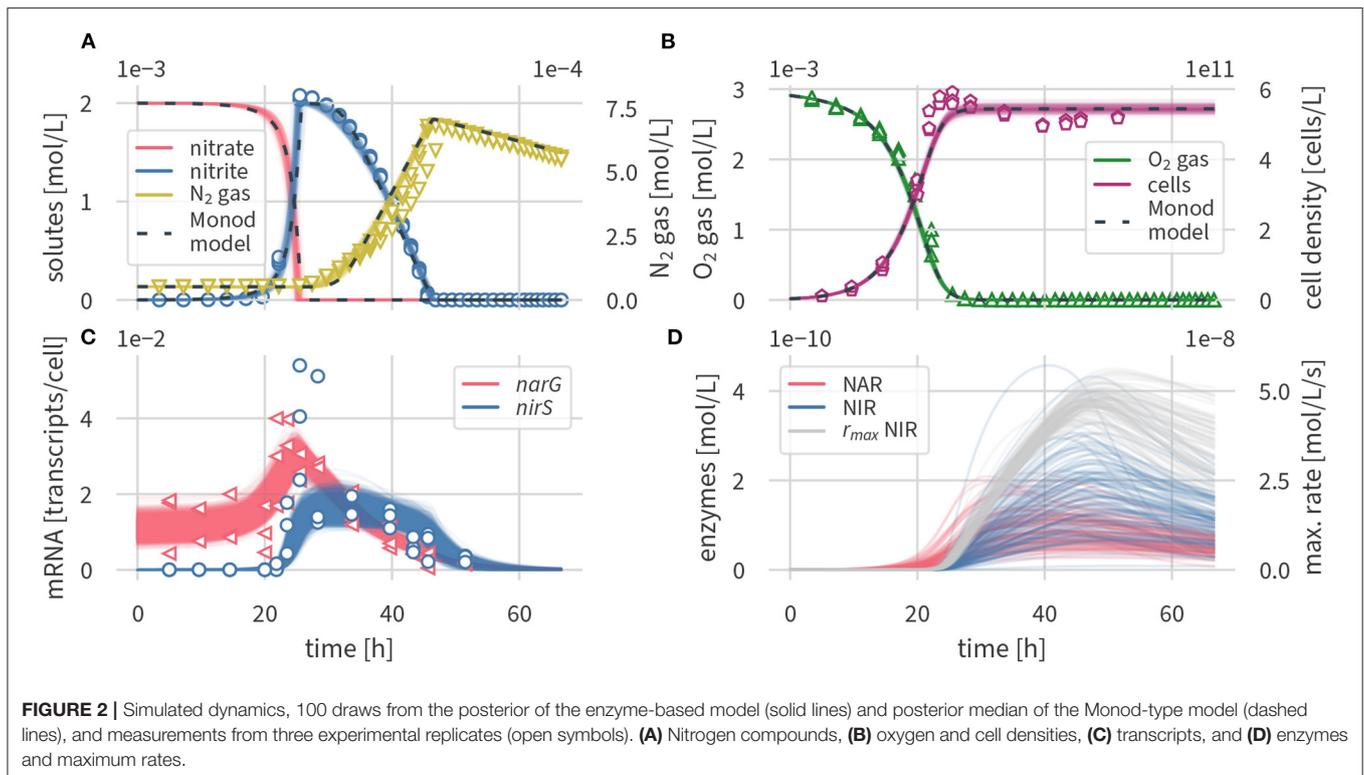
grow on both oxygen and nitrate (Boogerd et al., 1984), which may explain growth beyond the time of oxygen availability. The subsequent slight decrease in cell densities, if not a measurement artifact, may be a toxicity effect of nitrite or one of the non-modeled intermediates. Nitrite is known to be toxic to *P. denitrificans* and other microorganisms at millimolar concentrations (van Verseveld et al., 1977; Stouthamer, 1980; Zhang et al., 2013). Accounting for nitrite toxicity has been postulated as an explanation for delayed growth and increased cell lysis of *Shewanella oneidensis* while reducing nitrite to ammonium (Mellage et al., 2019). Altogether, we deemed the misfit in cell densities not significant enough to justify making the model more complex by introducing additional processes with difficult-to-constrain parameters.

3.2. Transcript and Enzyme Dynamics

Overall, the enzyme-based model formulation captured the dynamics of measured *narG* and *nirS* concentrations (**Figure 2C**). Both simulated and measured *narG* transcript concentrations fluctuated around a baseline value of 10^{-2} transcripts cell^{-1}, and *nirS* levels were very low during the oxic phase. Following a drop in oxygen levels, simulated *nar* transcripts rose, followed by those of *nir*. The model satisfactorily captured the peak of *narG* transcripts at 23 h and the following gradual decrease to low, but detectable transcript levels at the last measurement time point, at about 50 h. The model also predicted the very abrupt increase of *nir* transcripts at 24 h. While peak levels of measured *nir* transcripts at 25 h could not be matched, the measurement uncertainty of the transcript data at that time point was very large (2.4×10^{-2} transcripts cell^{-1} to

5.4×10^{-2} transcripts cell^{-1}), and thus we considered these peak values as outliers. Our assumption is supported by additional data of a parallel experiment with the same setup, but using butyrate instead of succinate as the carbon source, where a peak in *nir* transcripts was not detected (Qu et al., 2015). Analogous to measured data, both of the simulated transcripts remained at concentrations above the limit of detection for a few hours after complete denitrification.

As seen in **Figure 2D**, enzyme concentrations were subject to considerable uncertainty—an expected result, as there are no data to constrain them. However, the maximum cell-specific turnover rate of NIR $r_{\max}^{\mathrm{NIR}}$, given by the product of enzyme concentrations and the maximum enzyme specific turnover rate, was well constrained, provided that substrate was being consumed. The solute data that we used for conditioning could apparently provide sufficient information about reaction rates to estimate the maximum rate.

We also set up a model version with a more simplified description of transcriptional regulation, for comparison. The model, described in the **Supplementary Material**, yielded similar results (**Supplementary Figure 6**) to those outlined in **Figure 2**. However, the simplified model formulation was unable to capture the sudden increase in *nirS* transcripts (at $t = 24$ h). Therefore, we opted for the more complex transcription factor based formulation.

3.3. Posterior Parameter Uncertainty

Figure 3 shows prior and posterior distributions of selected parameters. In our application, most posterior parameter distributions were much narrower than the prior ranges

FIGURE 3 | Selected parameter distributions. **(A)** Kernel density estimates of the marginal distributions of selected parameters (marginal distributions of all parameters can be found in the **Supplementary Material**). Densities are cut off at the 94 % highest density intervals, circular markers indicate the mean. Parameters $k_{max}^{NO_2^-}$, $t_{1/2}^E$, I_{NNR} and $a_{NarR}^{NO_3^-}$ are not present in the Monod-type model such that the posterior is only shown for the enzyme-based formulation. Posterior distributions of Y_{O_2} are similar in both model formulations and cannot be distinguished in the plot. **(B)** Pairwise joint posterior distributions of parameters with strong correlation.

(**Figure 3A**), indicating that they were strongly constrained by the available data (e.g., Y_{O_2}). Where applicable, the Monod-type and enzyme-based models mostly resulted in similar parameter distributions. However, the estimated enzyme inhibition constants I_{reac}^{NAR} and I_{reac}^{NIR} were smaller in the Monod-type model compared to the enzyme-based model because the Monod-type model compensates the lacking regulation of enzyme levels with a stronger inhibition of enzymes by oxygen. Some parameters exhibit considerable spread even after conditioning to the data, partly ranging over several orders of magnitude. This holds mainly for parameters that are related to transcription factor, transcript and enzyme concentrations such as transcription factor activation rate constants (e.g., $a_{NarR}^{NO_3^-}$) or the enzyme half-life $t_{1/2}^E$.

In some cases, part of the spread can be attributed to the correlation among the conditional parameters, as shown in **Figure 3B** for selected examples. **Supplementary Figures 4, 5** of the **Supplementary Material** contain all correlation coefficients of the posterior distributions. **Figure 3B** illustrates that the

enzyme half-life $t_{1/2}^E$ strongly correlates with the "efficiency" of NIR enzymes expressed by $k_{max}^{NO_2^-}$. The enzyme half-life influences the concentration of the enzymes. Because the total reaction rate depends on the product of the enzyme concentration with the efficiency of a single enzyme, changing the enzyme half-life can be compensated by also changing the enzyme efficiency. Thus, if we had enzyme data, we could significantly decrease the uncertainty in enzyme half-lives and the enzyme-specific maximum turnover rate $k_{max}^{NO_2^-}$ of nitrite.

A strong correlation can also be observed between the Monod parameters related to aerobic respiration (**Figure 3B**), which has been previously reported (Liu and Zachara, 2001). Reparameterization alleviated the correlation between $v_{max}^{O_2}$ and K_{O_2}. The parameter $v_{max}^{O_2}$ is the cell-specific respiration rate at the limit of the substrate concentration approaching infinity. Instead, we used the cell-specific respiration rate at a fixed concentration $v_{fix}^{O_2}$ as a parameter (see Section 4 in the **Supplementary Material**). $v_{fix}^{O_2}$ and K_{O_2} were less correlated than before the reparametrization, but the correlation of $v_{fix}^{O_2}$ with the growth yield was larger.

3.4. Relationship Between Reaction Rates and Transcript Concentrations

It has been previously suggested to use transcript concentrations as a proxy for reaction rates (Nicolaisen et al., 2008; Freitag and Prosser, 2009; Brow et al., 2013; Achermann et al., 2020; Rohe et al., 2020). To test if such a relationship would be a valid assumption, we plot the simulated transcript concentrations and corresponding reaction rates against each other in **Figures 4A,D**. Cell-specific rates were calculated by applying the rate laws to the simulated concentrations and normalizing the result by the cell densities. Measured transcript concentrations plotted against the reaction rates based on the Monod-type model are shown for comparison.

Transcript concentrations of *nar* and *nir* are upregulated before rates increase (**Figures 4A,D**). Between 10 h to 25 h, *nar* transcripts show a positive, but non-linear correlation with reaction rates, before the rates drop to zero once nitrate is depleted. Nitrite reduction rates, in contrast, only start increasing when *nir* transcript levels have nearly reached their maximum (**Figure 4D**). Rates continue to increase when *nir* transcript levels start to drop such that rates and transcript levels are anti-correlated around 40 h, just before reaction rates drop abruptly upon substrate depletion. After the rapid drop in reaction rates, both transcripts remain present for several hours. We also looked at the relationship of enzyme concentrations with denitrification rates (**Figures 4B,E**). Enzyme concentrations are not well constrained by the data and their posterior uncertainty propagates to the relationship with reaction rates. We therefore plotted the product of enzyme concentrations and their "efficiency" $k_{max}^{NO_2^-}$, corresponding to the maximum potential reaction rate, without accounting for enzyme inhibition or substrate limitation (**Figures 4C,F**). The maximum rate of NIR is well constrained and shows an almost linear relationship with
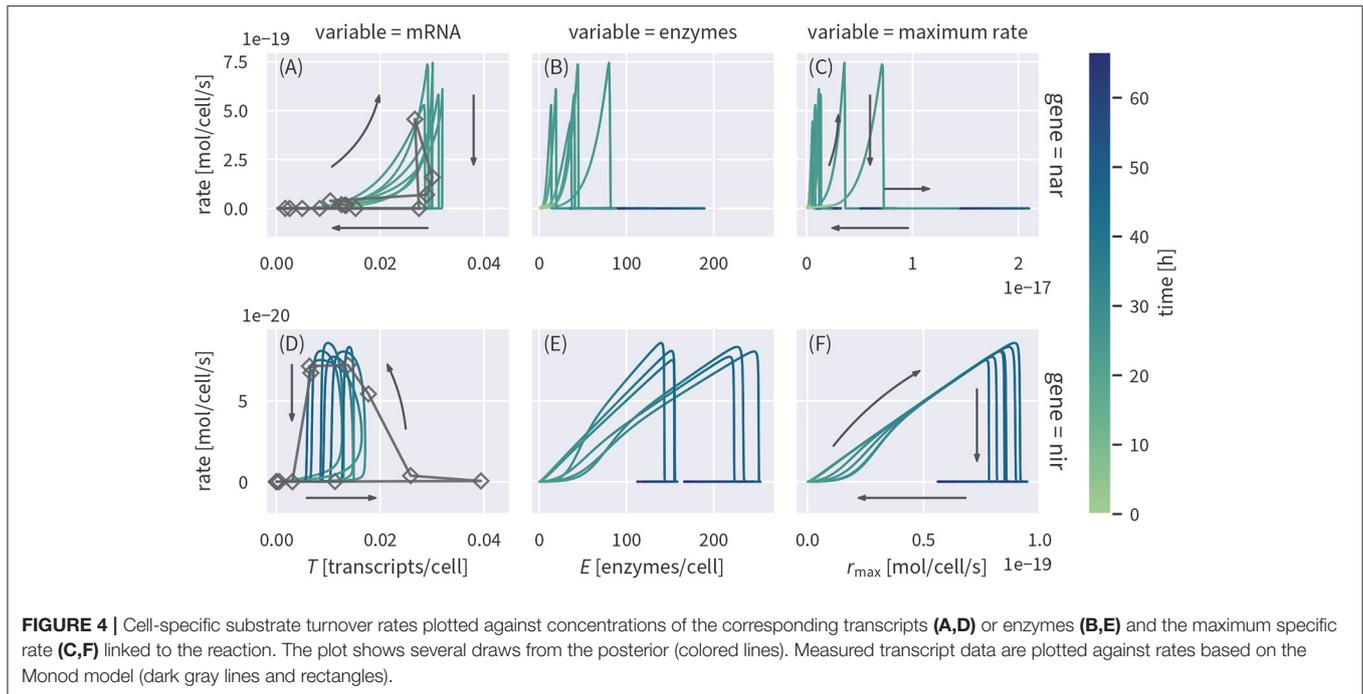
FIGURE 4 | Cell-specific substrate turnover rates plotted against concentrations of the corresponding transcripts **(A,D)** or enzymes **(B,E)** and the maximum specific rate **(C,F)** linked to the reaction. The plot shows several draws from the posterior (colored lines). Measured transcript data are plotted against rates based on the Monod model (dark gray lines and rectangles).

actual reaction rates up to the point where nitrite depletion cuts the rate (**Figure 4F**). The rate of nitrite reduction is essentially not limited by oxygen inhibition or substrate limitation. The maximum rate is thus the determining factor such that an increase in enzyme concentrations directly increases the rate. The relationship of NAR maximum rates and actual nitrate reduction rates (**Figure 4C**) is more complicated. The initial increase is strongly non-linear and the remaining posterior spread of the maximum rate leads to uncertainty in the scaling factor. Both findings can be explained by the initial inhibition of NAR enzymes by oxygen: NAR enzymes are produced while oxygen is still present so that r_{max} increases, but the actual rate does not, driven by the weaker inhibition of *enzyme production* relative to that of the *enzymatic catalysis*. The posterior uncertainty of the relationship is due to correlation between the oxygen inhibition parameter and the maximum rate. Different combinations of the parameters can yield the same reaction rate and an unambiguous assignment is not possible for a case when the data only constrain the effective rate.

Our results show that under the dynamic conditions of the experiment, transcript concentrations are not a good predictor of reaction rates. The complicated relationships are mainly caused by the asynchronism and different time scales of substrate dynamics, transcript and enzyme production, and decay. Under environmental conditions, changes in concentrations of redox-active species may be much slower than in the laboratory experiment simulated in this study. In groundwater, for example, the typical time scales are on the order of several months (Arora et al., 2013, 2016). If changes in redox-active species are slower than the enzyme dynamics (acting on the order of days), we would expect a better correlation between transcript and enzyme concentrations and reaction rates, potentially

allowing for further simplifications of model formulations. However, many environments in which nitrogen cycling plays a crucial role are very dynamic. For example, rapid water-content driven changes in the redox state of soils (Pronk et al., 2020) or diurnal fluctuations in river biogeochemistry (Kunz et al., 2017) can control nitrogen turnover. The lack of obvious correlations between transcripts of functional genes and reaction rates observed in our modeling results are likely representative for such dynamic environments. A non-linear, hysteretic relationship between transcript concentrations and reaction rates has also been found in a gene-centric model of pesticide degradation (Chavez Rodriguez et al., 2020). Interestingly, this is the case although pesticide degradation acts on much longer time scales than denitrification in this study and the observed relationship between rates and transcripts shows different patterns than the ones presented in **Figure 4** (e.g., direction of the hysteresis). This highlights the need to assess the relationship between gene expression and rates for each reaction system individually. Nevertheless, models that explicitly simulate transcript concentrations, as presented in this study, can help to understand the relationship between gene expression and rates and extrapolate it.

If enzyme concentrations were at quasi-steady state, the enzyme-based model could be simplified, reducing the number of state variables and parameters. To test if this assumption is valid, we ran the fully transient model version as shown in **Figure 2**) and computed the quasi-steady state concentrations based on the fully transient results. If the quasi-steady state assumption were valid, the two concentrations should cluster along the 1:1-line. As evident from **Figure 5**, this is not the case—enzyme concentrations exhibit a strongly hysteretic relationship and remain much lower than the corresponding quasi-steady state

FIGURE 5 | Quasi-steady enzyme state concentrations [enzymes L^{-1}] plotted against the concentrations simulated by the fully transient model. Although the exact relationship varies between different draws from the posterior (indicated by several lines), all of them are characterized by hysteresis.
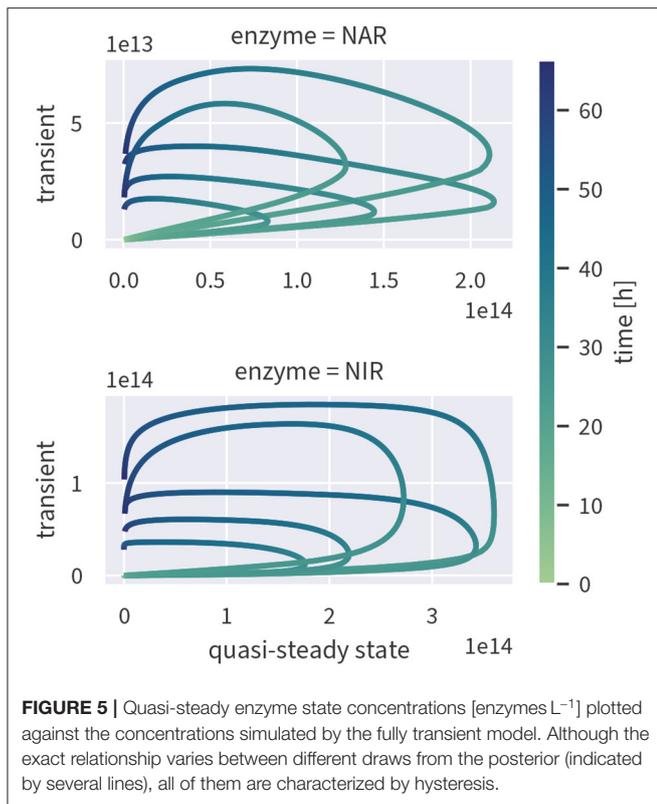
concentrations. Enzyme dynamics, as measured by estimated enzyme half-lives, are considerably slower than substrate dynamics such that the substrate stimulus is too short for enzyme concentrations to reach their quasi-steady state value.

3.5. Overall Model Performance

Our parsimonious enzyme-based model describes expressional patterns well and simulates denitrification rates very accurately. The model quantitatively links concentrations of functional transcripts to turnover rates. In reproducing the solute-concentration data of Qu et al. (2015), however, the Monod-type model is as good as the enzyme-based model. Small data errors lead to narrow posterior distributions of the solute concentrations in both models, whereas the much more uncertain transcript data do not help to further constrain them. While the enzyme-based model can be calibrated using more types of data than the Monod-type model, it also requires estimating more parameters, which appear to be weakly constrained. The strength of the enzyme-based model lies in improving our mechanistic understanding in a quantitative way. For example, it helps to decompose the contributions of transcriptional regulation and enzyme inhibition to the down-regulation of denitrification rates by oxygen. It confirms that the conceptual model of transcriptional regulation of *narG* and *nirS* through FnrP, NarR, and NNR can also explain expressional patterns quantitatively. Finally, the quantitative model helps to understand the non-linear relationship between rates and transcript or enzyme concentrations. It can be used to generate

hypotheses about the behavior in other systems and develop sampling strategies for these quantities.

Transcription-factor regulation and enzyme dynamics need to be explicitly described as these processes take place at similar or even larger time scales than the actual enzymatic reactions and the corresponding substrate dynamics. The relation between transcript concentrations and reaction rates exhibits a strong hysteresis. Enzyme concentrations show a better correlation with reaction rates at early times but are uncorrelated at later times due to substrate limitations. Thus, enzyme kinetics and the abundance of enzymes present likewise have an impact on reaction rates and need to be taken into account.

3.6. Transferability to Environmental Systems

While the experimental conditions during the experiment we simulated are not representative for many environments (high succinate concentrations, single organism), the controlled conditions allowed us to set up a detailed and mechanistic model of the denitrification reactions. The presented model equations hereafter can be integrated in a model that better represents natural environments, for example, by accounting for transport or carbon limitation. Representing all denitrifiers by a single organism in the model certainly does not account for all interactions that occur in natural microbial communities. Considering a single pool of denitrifiers in quantitative reaction models, however, is a common practice to keep the simulations computationally and conceptually tractable (Kinzelbach et al., 1991; Sanz-Prat et al., 2015; Yan et al., 2016). In addition, *P. denitrificans* is a denitrifying organism that is actually present in natural environments like soils (Nokhal and Schlegel, 1983). The exact mechanism of transcriptional regulation is not the same in all denitrifiers, but some basic principles are similar in various studied organisms (Gaimster et al., 2018). Further, the exact pattern of gene expressional dynamics does not matter for reaction rates because they are smoothed out by the much slower enzyme dynamics. A simplified representation of the transcriptional regulation as we present in the **Supplementary Material** might describe the behavior of a diverse microbial community sufficiently well, where the effective expressional patterns would be some smoothed average of the individual patterns produced by different regulation mechanisms.

An alternative approach to modeling transcriptional regulation that does not require exact knowledge of the underlying mechanism has been presented previously by Song et al. (2017). The latter authors applied a cybernetic approach (Song and Liu, 2015) that relies on the principle of return-on-investment: Microorganisms will transcribe those genes that maximize their metabolic "profits" (though it is often unclear how to express that in quantitative terms). Compared to the simulation results of the enzyme-based denitrification model presented earlier (Song et al., 2017), we can observe some key differences. While transcript concentrations respond relatively quickly to changes in substrate concentrations, Song et al. (2017) predict a long delay in the response of

transcripts. At the same time, enzyme concentrations respond to transcripts quickly in their model. While this is one possibility to explain the metabolic lag they observed, it is not necessary to ascribe the delay to slow *transcription* dynamics. Transcript dynamics in their model are not constrained by data so the actual lag could also be due to delayed *translation*. Differences between the model outcomes might be attributed to the different environmental systems. The laboratory system that we simulated exhibited faster kinetics and was more dynamic than the sediment that Song et al. (2017) analyzed (complete denitrification within 2 days with the pronounced accumulation of reactive intermediates, compared to a week). In contrast to the aforementioned study (Song et al., 2017), our simulations and transcript data support an immediate transcriptional response to denitrifying conditions, justifying the omission of a resource pool which first builds up before transcription starts in the model formulation.

The authors of the experimental data used here measured biomass in cell densities, which we also chose as state variable. In the natural environment, functional biomass is difficult to measure and a different approach is needed. A gene-centric approach that uses gene concentrations as proxies for functional biomass (Reed et al., 2014; Pagel et al., 2016) could be readily integrated into our model. Gene-centric approaches have been mostly applied to marine environments and their application to soils or groundwater might need extensions to account for relic DNA, which has been shown to represent as much as 40 % (Carini et al., 2016).

3.7. Implications for Biogeochemical Modeling

We informed our model with highly resolved time series of all solute concentrations in a controlled experiment. Under these conditions, the quantitative prediction of reaction rates was not improved by integrating transcript data. Since substrate and product concentrations are easier to obtain than transcript concentrations, quantitative rate predictions can be more effectively improved by additional chemical measurements than by integrating transcript concentrations. By contrast, transcript data can support the identification of key reactions, particularly if intermediate products have low concentrations and cannot be detected. In addition, transcript-based models can fill a key gap in our predictive capabilities within natural systems where the difficulty in acquiring cell density information impedes the validation of biomass-explicit models.

The advancement of enzyme-based models requires appropriate data for evaluating models and testing hypotheses. Datasets with quantitative information about functional genes, transcripts *and* enzymes could help to find accurate and parsimonious descriptions of transcript and enzyme dynamics. Measuring enzyme concentrations (e.g., by targeted quantification of functional enzymes, Li et al., 2017b) seems more promising than transcript concentrations as they more closely relate to actual turnover rates. However, measuring specific proteins in environmental samples may be restricted by major challenges, such as the efficient extraction of proteins from environmental samples, hindering accurate quantification and limiting the wider application of functional enzyme measurements (Starke et al., 2019).

We are convinced that further improvements of enzyme-based models in environmental systems can be achieved by integrating data from experiments carried out under different environmental conditions, expanding the sensitivity range of simulated processes and parameters. While batch experiments with high cell densities and nutrient concentrations usually exhibit fast dynamics there is a need to condition models to data from less dynamic systems such as chemostats and flow-through columns, which may more closely relate to natural environmental conditions.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: www.doi.org/10.5281/zenodo.4620167.

AUTHOR CONTRIBUTIONS

AS developed and implemented the models, and prepared the original draft, supervised by HP and OC. OC was responsible for funding acquisition. AS, HP, AM, and OC contributed to conceptualization, visualization, study design, writing, and revisions of the original draft. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.684146/full#supplementary-material

REFERENCES

Achermann, S., Mansfeldt, C. B., Müller, M., Johnson, D. R., and Fenner, K. (2020). Relating metatranscriptomic profiles to the micropollutant biotransformation potential of complex microbial communities. *Environ. Sci. Technol.* 54, 235–244. doi: 10.1021/acs.est.9b05421

Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* 7:13219. doi: 10.1038/ncomms13219

Arora, B., Dwivedi, D., Hubbard, S. S., Steefel, C. I., and Williams, K. H. (2016). Identifying geochemical hot moments and their controls on a contaminated river floodplain system using wavelet and entropy approaches. *Environ. Model. Softw.* 85, 27–41. doi: 10.1016/j.envsoft.2016.08.005

Arora, B., Mohanty, B. P., McGuire, J. T., and Cozzarelli, I. M. (2013). Temporal dynamics of biogeochemical processes at the Norman Landfill site. *Water Resour. Res.* 49, 6909–6926. doi: 10.1002/wrcr.20484

Bælum, J., Chambon, J. C., Scheutz, C., Binning, P. J., Laier, T., Bjerg, P. L., et al. (2013). A conceptual model linking functional gene expression and reductive dechlorination rates of chlorinated ethenes in clay rich groundwater sediment. *Water Res.* 47, 2467–2478. doi: 10.1016/j.watres.2013.02.016

Bælum, J., Nicolaisen, M. H., Holben, W. E., Strobel, B. W., Sørensen, J., and Jacobsen, C. S. (2008). Direct analysis of *tfdA* gene expression by indigenous bacteria in phenoxy acid amended agricultural soil. *ISME J.* 2, 677–687. doi: 10.1038/ismej.2008.21

Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao S., and Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9697–9702. doi: 10.1073/pnas.112318199

Blaszczyk, M. (1993). Effect of medium composition on the denitrification of nitrate by *Paracoccus denitrificans*. *Appl. Environ. Microbiol.* 59, 3951–3953. doi: 10.1128/aem.59.11.3951-3953.1993

Boogerd, F. C., van Verseveld H. W., Torenvliet, D., Braster, M., and Stouthamer, A. H. (1984). Reconsideration of the efficiency of energy transduction in *Paracoccus denitrificans* during growth under a variety of culture conditions. *Arch. Microbiol.* 139, 344–350. doi: 10.1007/BF00408377

Bouchez, T., Blieux, A. L., Dequiedt, S., Domaizon, I., Dufresne, A., Ferreira, S., et al. (2016). Molecular microbiology methods for environmental diagnosis. *Environ. Chem. Lett.* 14, 423–441. doi: 10.1007/s10311-016-0581-3

Bowen, J. L., Babbin, A. R., Kearns, P. J., and Ward, B. B. (2014). Connecting the dots: Linking nitrogen cycle gene expression to nitrogen fluxes in marine sediment mesocosms. *Front. Microbiol.* 5:429. doi: 10.3389/fmicb.2014.00429

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. Royal Stat. Soc., Ser. B* 26, 211–252. doi: 10.1111/j.2517-6161.1964.tb00553.x

Brow, C. N., O'Brien Johnson, R., Johnson, R. L., and Simon, H. M. (2013). Assessment of anaerobic toluene biodegradation activity by *bssA* transcript/gene ratios. *Appl. Environ. Microbiol.* 79, 5338–5344. doi: 10.1128/AEM.01031-13

Byrne, G. D., and Hindmarsh, A. C. (1975). A polyalgorithm for the numerical solution of ordinary differential equations. *ACM Trans. Math. Softw.* 1, 71–96. doi: 10.1145/355626.355636

Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., and Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* 2, 1–6. doi: 10.1101/043372

Chavez Rodriguez, L., Ingalls, B., Schwarz, E., Streck, T., Uksa, M., and Pagel, H. (2020). Gene-centric model approaches for accurate prediction of pesticide biodegradation in soils. *Environ. Sci. Technol.* 54, 13638–13650. doi: 10.1021/acs.est.0c03315

Crack, J. C., Hutchings, M. I., Thomson, A. J., and Le Brun, N. E. (2016). Biochemical properties of *Paracoccus denitrificans* FnrP: Reactions with molecular oxygen and nitric oxide. *J. Biol. Inorg. Chem.* 21, 71–82. doi: 10.1007/s00775-015-1326-7

Dong, H., Li, L., Lu, Y., Cheng, Y., Wang, Y., Ning, Q., et al. (2019). Integration of nanoscale zero-valent iron and functional anaerobic bacteria for groundwater remediation: a review. *Environ. Int.* 124, 265–277. doi: 10.1016/j.envint.2019.01.030

Foreman-Mackey D. (2020). *Exoplanet-Dev/Pymc3-Ext*. Version 0.0.2. Available online at: https://github.com/exoplanet-dev/pymc3-ext.

Freitag, T. E., and Prosser, J. I. (2009). Correlation of methane production and functional gene transcriptional activity in a peat soil. *Appl. Environ. Microbiol.* 75, 6679–6687. doi: 10.1128/AEM.01021-09

Gaimster, H., Alston, M., Richardson, D. J., Gates, A. J., and Rowley, G. (2018). Transcriptional and environmental control of bacterial denitrification and N_2O emissions. *FEMS Microbiol. Lett.* 365:fnx277. doi: 10.1093/femsle/fnx277

Gesztelyi, R., Zsuga, J., Kemeny-Beke A., Varga, B., Juhasz, B., and Tosaki, A. (2012). The Hill equation and the origin of quantitative pharmacology. *Arch. Hist. Exact Sci.* 66, 427–438. doi: 10.1007/s00407-012-0098-5

Giannopoulos, G., Sullivan, M. J., Hartop, K. R., Rowley, G., Gates, A. J., Watmough, N. J., et al. (2017). Tuning the modular *Paracoccus denitrificans* respirome to adapt from aerobic respiration to anaerobic denitrification. *Environ. Microbiol.* 19, 4953–4964. doi: 10.1111/1462-2920.13974

Gutiérrez, M., Biagioni, R. N., Teresa Alarcón-Herrera M., and Rivas-Lucero B. A. (2018). An overview of nitrate sources and operating processes in arid and semiarid aquifer systems. *Sci. Total Environ.* 624, 1513–1522. doi: 10.1016/j.scitotenv.2017.12.252

Härtig, E., and Zumft, W. G. (1999). Kinetics of *nirS* expression (cytochrome cd_1 nitrite reductase) in *Pseudomonas stutzeri* during the transition from aerobic respiration to denitrification: evidence for a denitrification-specific nitrate- and nitrite-responsive regulatory system. *J. Bacteriol.* 181, 161–166. doi: 10.1128/JB.181.1.161-166.1999

Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., et al. (2005). SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Trans. Math. Softw.* 31, 363–396. doi: 10.1145/1089014.1089020

Hoffman, M. D., and Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 15, 1593–1623. Available online at: https://dl.acm.org/doi/10.5555/2627435.2638586

Howarth, R. W. (2008). Coastal nitrogen pollution: a review of sources and trends globally and regionally. *Harmful Algae* 8, 14–20. doi: 10.1016/j.hal.2008.08.015

Ingalls, B. P. (2013). *Mathematical Modeling in Systems Biology: An Introduction.* Cambridge, MA: MIT Press.

Jackson, K. R., and Sacks-Davis R. (1980). An alternative implementation of variable step-size multistep formulas for stiff ODEs. *ACM Trans. Math. Softw.* 6, 295–318. doi: 10.1145/355900.355903

Kinzelbach, W., Schäfer, W., and Herzer, J. (1991). Numerical modeling of natural and enhanced denitrification processes in aquifers. *Water Resour. Res.* 27, 1123–1135. doi: 10.1029/91WR00474

Koutinas, M., Kiparissides, A., Silva-Rocha R., Lam, M.-C., Martins dos Santos V. A. P., de Lorenzo V., et al. (2011). Linking genes to microbial growth kinetics—An integrated biochemical systems engineering approach. *Metab. Eng.* 13, 401–413. doi: 10.1016/j.ymben.2011.02.001

Kumar, R., Carroll, C., Hartikainen, A., and Martin, O. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *J. Open Source Softw.* 4:1143. doi: 10.21105/joss.01143

Kunz, J. V., Hensley, R., Brase, L., Borchardt, D., and Rode, M. (2017). High frequency measurements of reach scale nitrogen uptake in a fourth order river with contrasting hydromorphology and variable water chemistry (Weiße Elster Germany). *Water Resour. Res.* 53, 328–343. doi: 10.1002/2016WR019355

Kuypers, M. M. M., Marchant, H. K., and Kartal, B. (2018). The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* 16, 263–276. doi: 10.1038/nrmicro.2018.9

Lee, Y.-Y., Shearer, N., and Spiro, S. (2006). Transcription factor NNR from *Paracoccus denitrificans* is a sensor of both nitric oxide and oxygen: Isolation of *nnr** alleles encoding effector-independent proteins and evidence for a haem-based sensing mechanism. *Microbiology* 152, 1461–1470. doi: 10.1099/mic.0.28796-0

Li, L., Maher, K., Navarre-Sitchler A., Druhan, J., Meile, C., Lawrence, C., et al. (2017a). Expanding the role of reactive transport models in critical zone processes. *Earth Sci. Rev.* 165, 280–301. doi: 10.1016/j.earscirev.2016.09.001

Li, M., Gao, Y., Qian, W.-J., Shi, L., Liu, Y., Nelson, W. C., et al. (2017b). Targeted quantification of functional enzyme dynamics in environmental samples for microbially mediated biogeochemical processes. *Environ. Microbiol. Rep.* 9, 512–521. doi: 10.1111/1758-2229.12558

Li, M., Qian, W.-J., Gao, Y., Shi, L., and Liu, C. (2017c). Functional enzyme-based approach for linking microbial community functions with biogeochemical process kinetics. *Environ. Sci. Technol.* 51, 11848–11857. doi: 10.1021/acs.est.7b03158

Liu, C., and Zachara, J. M. (2001). Uncertainties of Monod kinetic parameters nonlinearly estimated from batch experiments. *Environ. Sci. Technol.* 35, 133–141. doi: 10.1021/es001261b

Liu, L., and Greaver, T. L. (2009). A review of nitrogen enrichment effects on three biogenic GHGs: The CO_2 sink may be largely offset by stimulated N_2O and CH_4 emission. *Ecol. Lett.* 12, 1103–1117. doi: 10.1111/j.1461-0248.2009.01351.x

Louca, S., Hawley, A. K., Katsev, S., Torres-Beltran M., Bhatia, M. P., Kheirandish, S., et al. (2016). Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5925–E5933. doi: 10.1073/pnas.1602897113

Maier, T., Schmidt, A., Güell, M., Kühner, S., Gavin, A.-C., Aebersold, R., et al. (2011). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* 7:511. doi: 10.1038/msb.2011.38

Mellage, A., Smeaton, C. M., Furman, A., Atekwana, E. A., Rezanezhad, F., and Cappellen, P. V. (2019). Bacterial Stern layer diffusion: experimental determination with spectral induced polarization and sensitivity to nitrite toxicity. *Near Surf. Geophys.* 17, 623–635. doi: 10.1002/nsg.12058

Monard, C., Martin-Laurent F., Lima, O., Devers-Lamrani M., and Binet, F. (2013). Estimating the biodegradation of pesticide in soils by monitoring pesticide-degrading gene expression. *Biodegradation* 24, 203–213. doi: 10.1007/s10532-012-9574-5

Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L.-K., et al. (2013). Sizing up metatranscriptomics. *ISME J.* 7, 237–243. doi: 10.1038/ismej.2012.94

Nicolaisen, M. H., Bælum, J., Jacobsen, C. S., and Sørensen, J. (2008). Transcription dynamics of the functional *tfdA* gene during MCPA herbicide degradation by *Cupriavidus necator* AEO106 (pRO101) in agricultural soil. *Environ. Microbiol.* 10, 571–579. doi: 10.1111/j.1462-2920.2007.01476.x

Nokhal, T.-H., and Schlegel, H. G. (1983). Taxonomic study of *Paracoccus denitrificans*. *Int. J. Syst. Bacteriol.* 33, 26–37. doi: 10.1099/00207713-33-1-26

Pagel, H., Poll, C., Ingwersen, J., Kandeler, E., and Streck, T. (2016). Modeling coupled pesticide degradation and organic matter turnover: From gene abundance to process rates. *Soil Biol. Biochem.* 103, 349–364. doi: 10.1016/j.soilbio.2016.09.014

Pronk, G. J., Mellage, A., Milojevic, T., Smeaton, C. M., Engel, K., Neufeld, J. D., et al. (2020). Carbon turnover and microbial activity in an artificial soil under imposed cyclic drainage and imbibition. *Vadose Zone J.* 19, e20021. doi: 10.1002/vzj2.20021

Qu, Z., Bakken, L. R., Molstad, L., Frostegård, Å., and Bergaust, L. L. (2015). Transcriptional and metabolic regulation of denitrification in *Paracoccus denitrificans* allows low but significant activity of nitrous oxide reductase under oxic conditions. *Environ. Microbiol.* 18, 2951–2963. doi: 10.1111/1462-2920.13128

Reed, D. C., Algar, C. K., Huber, J. A., and Dick, G. J. (2014). Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc. Natl. Acad. Sci. U.S.A.* 111, 1879–1884. doi: 10.1073/pnas.1313713111

Reed, D. C., Breier, J. A., Jiang, H., Anantharaman, K., Klausmeier, C. A., Toner, B. M., et al. (2015). Predicting the response of the deep-ocean microbiome to geochemical perturbations by hydrothermal vents. *ISME J.* 9, 1857–1869. doi: 10.1038/ismej.2015.4

Rohe, L., Oppermann, T., Well, R., and Horn, M. A. (2020). Nitrite induced transcription of p450nor during denitrification by Fusarium oxysporum correlates with the production of N_2O with a high 15N site preference. *Soil Biol. Biochem.* 151:108043. doi: 10.1016/j.soilbio.2020.108043

Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Comput. Sci.* 2:e55. doi: 10.7717/peerj-cs.55

Sanz-Prat A., Lu, C., Finkel, M., and Cirpka, O. A. (2015). On the validity of travel-time based nonlinear bioreactive transport models in steady-state flow. *J. Contam. Hydrol.* 175-176:26–43. doi: 10.1016/j.jconhyd.2015.02.003

Scheibe, T. D., Mahadevan, R., Fang, Y., Garg, S., Long, P. E., and Lovley, D. R. (2009). Coupling a genome-scale metabolic model with a reactive transport model to describe *in situ* uranium bioremediation. *Microb. Biotechnol.* 2, 274–286. doi: 10.1111/j.1751-7915.2009.00087.x

Seyboldt, A. (2020). *Sunode*. Version 0.1.2. Zenodo. doi: 10.5281/zenodo.4058330

Song, H.-S., and Liu, C. (2015). Dynamic metabolic modeling of denitrifying bacterial growth: the cybernetic approach. *Ind. Eng. Chem. Res.* 54, 10221–10227. doi: 10.1021/acs.iecr.5b01615

Song, H.-S., Thomas, D. G., Stegen, J. C., Li, M., Liu, C., Song, X., et al. (2017). Regulation-structured dynamic metabolic model provides a potential mechanism for delayed enzyme response in denitrification process. *Front. Microbiol.* 8:1866. doi: 10.3389/fmicb.2017.01866

Spiro, S. (2017). "Regulation of Denitrification," in *Metalloenzymes in Denitrification: Applications and Environmental Impacts,* eds I. Moura, J. J. G. Moura, S. R. Pauleta, and L. B. Maia, *Number 9 in Metallobiology* (Cambridge: The Royal Society of Chemistry), 312–330. doi: 10.1039/9781782623762-00312

Starke, R., Jehmlich, N., and Bastida, F. (2019). Using proteins to study how microbes contribute to soil ecosystem services: the current state and future perspectives of soil metaproteomics. *J. Proteomics* 198, 50–58. doi: 10.1016/j.jprot.2018.11.011

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., et al. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science* 347:1259855. doi: 10.1126/science.1259855

Störiko, A., Pagel, H., Mellage, A., and Cirpka, O. (2021). *Nitrogene: Modelling Code and Data of an Enzyme-Based Denitrification Model (Version 0.0.2).* Zenodo. doi: 10.5281/zenodo.4620167

Stouthamer, A. H. (1980). Bioenergetic studies on *Paracoccus denitrificans*. *Trends Biochem. Sci.* 5, 164–166. doi: 10.1016/0968-0004(80)90015-8

van Verseveld H. W., Meijer, E. M., and Stouthamer, A. H. (1977). Energy conservation during nitrate respiration in *Paracoccus denitrificans*. *Arch. Microbiol.* 112, 17–23. doi: 10.1007/BF00446649

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2020). Rank-normalization, folding, and localization: an improved $\widehat{R}$ for assessing convergence of MCMC. *Bayesian Anal.* doi: 10.1214/20-BA1221. [Epub ahead of print].

Wegner, C.-E., Gaspar, M., Geesink, P., Herrmann, M., Marz, M., and Küsel, K. (2019). Biogeochemical regimes in shallow aquifers reflect the metabolic coupling of the elements nitrogen, sulfur, and carbon. *Appl. Environ. Microbiol.* 85:e02346-18. doi: 10.1128/AEM.02346-18

Wood, N. J., Alizadeh, T., Bennett, S., Pearce, J., Ferguson, S. J., Richardson, D. J., et al. (2001). Maximal expression of membrane-bound nitrate reductase in *Paracoccus* is induced by nitrate via a third FNR-like regulator named NarR. *J. Bacteriol.* 183, 3606–3613. doi: 10.1128/JB.183.12.3606-3613.2001

Yan, S., Liu, Y., Liu, C., Shi, L., Shang, J., Shan, H., et al. (2016). Nitrate bioreduction in redox-variable low permeability sediments. *Sci. Total Environ.* 539:185–195. doi: 10.1016/j.scitotenv.2015.08.122

Zhang, H., Fu, H., Wang, J., Sun, L., Jiang, Y., Zhang, L., et al. (2013). Impacts of nitrate and nitrite on physiology of *Shewanella oneidensis*. *PLoS ONE* 8:e62629. doi: 10.1371/journal.pone.0062629