Available online at www.sciencedirect.com
ScienceDirect
Geochimica et Cosmochimica Acta 75 (2011) 337–351
Geochimica et Cosmochimica Acta
www.elsevier.com/locate/gca

Compendium and synthesis of bacterial manganese reduction rates

Joel Z. Bandstra [a,*], Daniel E. Ross [b], Susan L. Brantley [b], William D. Burgos [b]

[a] Department of Chemistry, Mathematics, and Physical Sciences, Saint Francis University, 117 Evergreen Drive, P.O. Box 600, Loretto, PA 15940, USA
[b] Center for Environmental Kinetics Analysis, Pennsylvania State University, 2217 EES Building, University Park, PA 16802, USA

Received 10 May 2009; accepted in revised form 10 April 2010; available online 16 October 2010

Abstract

We have compiled time-series concentration data for the biological reduction of manganese(III/IV) published between 1985 and 2004 and fit these data with a simple hyperbolic rate expression or, when appropriate, one of its limiting forms. The compiled data and rate constants are available in Electronic Annex EA-1. The zero- and first-order rate constants appear to follow a log–normal distribution that could be used, for example, in predictive modeling of Mn-oxide reduction in a reactive transport scenario. We have also included details of the experimental procedures used to generate each time-series data-set in our compilation. These meta-data—mostly pertaining to the type and concentration of micro-organism, electron donor, and electron acceptor—enable us to examine the rate data for trends. We have computed a number of rudimentary, monovariate statistics on the compiled data with the hope of stimulating both more detailed statistical analyses of the data and new experiments to fill gaps in the existing data-set. We have also analyzed the data with parametric models based on the log–normal distribution and rate equations that are hyperbolic in the concentration of cells and Mn available for reduction. This parametric analysis allows us to provide best estimates of zero- and first-order rate constants both ignoring and accounting for the meta-data.

1. INTRODUCTION

Over the past few decades a number of studies have reported on the bacterially mediated reduction of Mn(III/IV) to Mn(II). Many of these studies have included time-series Mn concentration data collected in batch reactors such as the data shown in Fig. 1 (Burdige and Nealson, 1985; Lovley and Phillips, 1988; Myers and Nealson, 1988a,b; Rusin et al., 1991; Burdige et al., 1992; Kostka et al., 1995; Greene et al., 1997; Bratina et al., 1998; Kieft et al., 1999; Fredrickson et al., 2002, 2004). The conditions under which these experiments were conducted are summarized in Table 1. Several time-series data-sets are typically collected with an experimental condition varied systematically among them. Mechanistic conclusions are then inferred from observed differences (or lack thereof) in the reaction rate. In only a few of these studies, however, have the time-series data been analyzed with any sort of kinetic model (e.g., Burdige and Nealson, 1985; Burdige et al., 1992). Despite this lack of kinetic analysis within and between research studies, several geochemical models now provide algorithms that can be used to model the kinetics of dissimilatory Mn reduction in field settings.

While a great deal of information—especially concerning reaction mechanisms—has already been gained from qualitative analyses of rate data, it is likely that still more information can be gained from a quantitative approach. In particular, the time-series data should contain the information needed to estimate zero- and first-order rate constants which are needed for predictive modeling (e.g., with a reactive transport code) of bacterial Mn reduction

* Corresponding author. Tel.: +1 814 472 3854; fax: +1 814 472 3028.
E-mail address: jbandstra@francis.edu (J.Z. Bandstra).

0016-7037/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.gca.2010.04.069

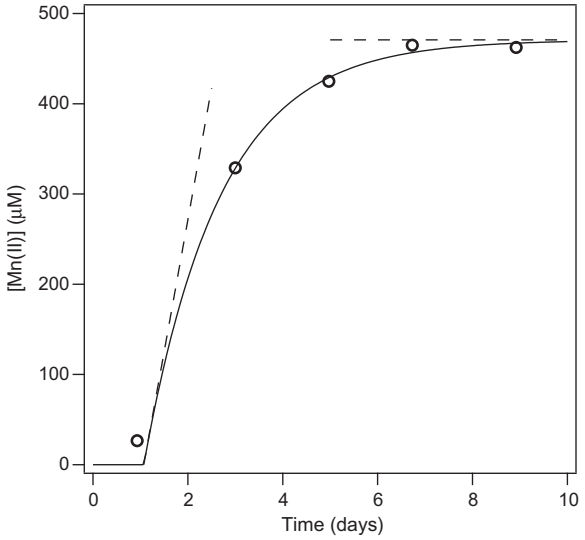

Fig. 1. An example concentration time-series data-set taken from Fig. 4 of Burdige et al. (1992).

rates in the environment. Furthermore, the presence or lack of a quantitatively established trend in rate data as a function of experimental inputs (e.g., concentration of cells, Mn available for reduction, electron donor concentration) should be useful in further establishing or challenging the conclusions of previous studies.

In order to perform the kind of comprehensive quantitative analysis described above, it is first necessary to gather the relevant time-series data in a consistent format. We have done this according to the procedure described in Section 2 and have provided these compiled data in Electronic Annex EA-1. We expect that the compilation will prove to be a particularly useful part of this work since it will enable other researchers to perform their own analyses and to add to the knowledge base for this system in a way that accounts for the existing data. It should be noted that several of the studies in the compilation used inhibitor compounds in some of their experiments. We have included these data in Electronic Annex, but have excluded them from all of the graphs and statistical analyses described below.

The problem presented by bacterial Mn reduction is prototypical of environmental kinetics as a whole in that existing (and generally high quality) laboratory based rate measurements are used to predict the behavior of natural systems. Recently, data compilation and "meta-analysis" have been employed as a means of solving this scaling problem (Olsen and Rimstidt, 2007; Brantley et al., 2008). It remains to be seen, however, if this approach can produce coherent results across the wide spectrum of environmental systems. In this light, the present work can be considered as a datum on the utility of data compilation.

2. METHODS

2.1. Data and meta-data harvesting

Our compilation includes data from studies of bacterial manganese reduction that reported metal concentration as a function of time. In most studies, time-series data were reported graphically. We used the extraction software GrabIt to digitize the data (Datatrend Software; Raleigh, NC). We have also included information about experimental conditions employed in each study, which are referred to henceforth as meta-data. We wish to stress that no tabular compilation of meta-data can replace journal length publications describing the data. The most important piece of meta-data, therefore, is the citation to the original journal article.

There are two kinds of meta-data in our compilation: categorical meta-data and quantitative meta-data. Categorical meta-data are those data that describe the type of experiment or type of reagent used. These data are descriptions rather than numerical values. Quantitative meta-data are those data that take numerical values. These data describe the degree of an experimental condition or the amount of reagent used.

The extracted time-series data and their relevant meta-data are provided in Table EA-1-1 of Electronic Annex.

2.2. Curve fitting

Almost all of the data in our compilation depict the appearance of reduced products over time and, therefore, we present our analysis methods in terms of the kinetics of Mn(II) appearance. In those few cases where data were reported that describe decreasing Mn(III/IV) concentrations, we followed the same basic procedures but with "disappearance" rate laws substituted for the "appearance" rate laws.

All of the data possessed one or more of the following features: (i) a lag period during which concentration remained constant; (ii) a linear increase in concentration over time; or (iii) an exponential approach to a horizontal asymptote. Considering a metal reduction reaction written generically as $M_{ox} \rightarrow M_{red}$ we can formulate a rate equation that mimics all three of these features as follows:

$$\frac{d[M_{red}]}{dt} = \begin{cases} 0 & t < t_{lag} \\ \frac{V_{max}[M_{ox}]}{K_{1/2}+[M_{ox}]} & t \geqslant t_{lag} \end{cases} \quad (1)$$

$$[M_{ox}] + [M_{red}] = [M_{ox}]_0 + [M_{red}]_0 = [M_{tot}] \quad (2)$$

where $[M_{red}]$ and $[M_{ox}]$ are, respectively, concentrations of the reduced and oxidized metal over time; $[M_{red}]_0$ and $[M_{ox}]_0$ are initial concentrations; $[M_{tot}]$ is a constant; V_{max} and $K_{1/2}$ are rate parameters with dimensions of concentration per time and concentration, respectively; and t_{lag} is the length of the lag period. We refer to Eq. (1) as the mixed-order model because it contains a transition from behavior that is zero-order in nature at high $[M_{ox}]$ to behavior that is first-order in nature at low $[M_{ox}]$.

Not all of the data exhibited each of the three features described above. Some of the data exhibited an induction period and a linear increase in concentration (features (i) and (ii)) but no exponential asymptote (feature (iii)). A rate equation mimicking these features can be written as:

$$\frac{d[M_{red}]}{dt} = \begin{cases} 0 & t < t_{lag} \\ k_{zo} & t \geqslant t_{lag} \end{cases} \quad (3)$$

Table 1
Summary of data-sets included in compilation.

Reference	No. of data-sets	Microbe	[*Cells*] (L^{-1})	Electron acceptor	$[Mn]_{total}$ (M)	Electron donor	[*Donor*] (mM)	Initial rate (mM h^{-1})
Bratina et al. (1998)	4	Microbes concentrated from water		Birnessite	0.0023	None, peptone and yeast extract		3.78e−05–0.00575
Burdige et al. (1992)	22	Chesapeake Bay enrichment culture	9×10^8	Vernadite, birnessite, pyrolusite	0.000115–0.0023	Acetate	0.015	1.78e−05–0.0288
Burdige and Nealson (1985)	19	Enrichment cultures		Vernadite	0.00115–0.00575	Lactate, succinate		0.000196–0.0463
Fredrickson et al. (2002)	11	*Shewanella putrefaciens* CN32	3×10^{11}	Pyrolusite, birnessite, sediments, mixed metals	0.02–0.0501	Hydrogen gas		0.00667–0.382
Fredrickson et al. (2004)	3	*Shewanella putrefaciens* CN32	8×10^{10}	Sediments	0.00417–0.0235	Lactate	0.01	0.0207–0.0239
Greene et al. (1997)	1	*Deferribacter thermophilus* BMA	4×10^7	Vernadite	0.015	Yeast extract		0.283
Kieft et al. (1999)	2	*Thermus* SA-01	1×10^{12}	Birnessite	0.01	Lactate, hydrogen gas	0.002	0.0230–0.0405
Kostka et al. (1995)	14	*Shewanella putrefaciens* MR-1	1×10^{10}	Aqueous Mn(III)	0.04	Formate, lactate, hydrogen gas	0.005–0.01	0.00418–0.695
Lovley and Phillips (1988)	1	*Geobacter metallireducens* GS-15	1×10^{10}	Birnessite	0.015	Acetate	0.001	0.236
Myers and Nealson (1988a)	19	*Shewanella oneidensis* MR-1	8×10^9	MnO_2, MnO_2 and Fe oxide	0.0002–0.000215	Succinate	0.015	0.0015–0.0346
Myers and Nealson (1988b)	10	*Shewanella oneidensis* MR-1	9×10^8–2×10^{10}	MnO_2	0.0002	Succinate	0.015	0.00142–0.00842
Rusin et al. (1991)	3	*Bacillus polymyxa* D1, *Shewanella oneidensis* MR-1, *Shewanella putrefaciens* sp. 200		Pyrolusite	0.04	Glucose, lactate	0.01	0.0136–0.148

where k_{zo} is a rate parameter with dimensions of concentration per time. We refer to Eq. (3) as the zero-order model. Other data exhibited an induction period and an asymptote (features (i) and (iii)) but no linear increase (feature (ii)); behavior which is mimicked by:

$$\frac{d[M_{red}]}{dt} = \begin{cases} 0 & t < t_{lag} \\ k_{fo}[M_{ox}] & t \geqslant t_{lag} \end{cases} \quad (4)$$

where k_{fo} is a rate parameter with dimensions of per time. We refer to Eq. (4) as the first-order model. Of course, some of the data did not exhibit an induction period. Such behavior can be mimicked with any of Eqs. (1), (3), or (4) by setting t_{lag} to zero.

We fit the integrated forms of the mixed-, zero-, and first-order models to the appropriate data by means of square error minimization either analytically—for linear parameters—or using the Levenberg–Marquardt algorithm (Press et al., 1992). Direct integration of the mixed-order model yields an implicit solution such that least-squares fitting requires the additional step of numerical inversion. For this we used Brent's method (Brent, 1973). Many of the data-sets could be fit with two or even all three of the models. Data that are purely first-order, for example, can always be fit with the mixed-order model by making appropriately large choices for V_{max} and $K_{1/2}$ (Agrawal et al., 2002). In such cases we chose the best model by calculating the reduced sum of square errors according to:

$$SSE_{red} = \frac{SSE}{N - M} \quad (5)$$

where SSE is the sum of square errors between the fitted model and the data, N is the number of data points, and M is the number of free parameters ($N - M$ is the degrees of freedom). We then chose the model with the least SSE_{red} as the best model for the data. We treated the existence of a lag period (i.e., whether to hold t_{lag} at zero) and the selection of initial conditions (i.e., whether to calculate $[M_{red}]_0$ *a priori* or to treat it as a fitting parameter) according to the same procedure. We estimated confidence intervals on fitted parameters from the covariance matrix assuming that the errors were normally distributed and homoscedastic (Seber and Wild, 2003).

We would ultimately like to calculate a rate that is consistent across data-sets even when the data were best fit by different models. Since the rate constants cannot be used for this purpose, we used the rate of the fitted model at the end of the lag period (i.e., the initial rate):

$$rate|_{t_{lag}} = \begin{cases} k_{zo} & \text{zero-order} \\ k_{fo}([M_{tot}] - [M_{red}]|_0) & \text{first-order} \\ \frac{V_{max}([M_{tot}]-[M_{red}]|_0)}{K_{1/2}+([M_{tot}]-[M_{red}]|_0)} & \text{mixed-order} \end{cases} \quad (6)$$

Confidence intervals for the rate evaluated at t_{lag} can be approximated by expanding Eq. (6) into the first two terms of a Taylor series around the mean values of the parameters and inserting this expression into the definition of variance. This gives an expression for the variance on the rate at t_{lag} in terms of the fitted parameters and their covariance matrix. In the zero-order case, the confidence interval on rate at time $= t_{lag}$ is simply equal to that on k_{zo}.

The resulting fitting parameters, covariance elements, and relevant meta-data have been compiled in Table EA-1-2 of Electronic Annex.

2.3. Statistical analyses

There are two kinds of output issuing from the curve fitting procedure: quantitative fitting parameters (together with associated uncertainties) and the model type that achieved the best fit. Each output variable, whether quantitative or categorical, possesses a distribution and these distributions are controlled, at least partially, by correlative distributions in the meta-data. In performing statistical analyses we are concerned, first, with describing the distributions of the output data and, second, with exploring correlations that exist between the output data and the meta-data. Since our meta-data also consist of both categorical and quantitative data, we require the means to summarize both types of data as well as measures of association between two categorical variables, between a categorical and a quantitative variable, and between two quantitative variables. Throughout our statistical analyses we have excluded those data where an inhibiting species was present.

The distribution of a categorical variable is fully described by the number of instances falling into each category. A convenient way to summarize this type of distribution is the contingency table—such as Tables 2 and 4—where the rows and columns are labeled by two different categorical variables and the entries give the number of observed events for each row/column combination. The contingency table also forms the basis for our method of assessing the degree of association between two categorical variables. If there is no association, the contingency table will be populated at random. If there is an association, some pattern should be evident in the contingency table. The χ^2 statistic, given by Eq. (7), measures the statistical strength of the pattern in a contingency table.

$$\chi^2 = \sum_{i,j} \frac{(N_{i,j} - n_{i,j})^2}{n_{i,j}} \quad (7)$$

$N_{i,j}$ is the number of observations in row i and column j of the contingency table and $n_{i,j}$ is the expected number of observations at location i, j based on a null hypothesis of no association between the variables (i.e., $n_{i,j} = \sum_j N_{i,j} \times \sum_i N_{i,j} \div \sum_{i,j} N_{i,j}$). A larger value of χ^2 corresponds to a greater deviation from a random distribution of entries in the contingency table. The statistical significance of a given value of χ^2 can be found using the complement of the incomplete gamma function with degrees of freedom equal to $I \cdot J - I - J + 1$ where I and J are the number of rows and columns, respectively (Press et al., 1992). The statistical significance found in this way can be interpreted as the probability that χ^2 would be greater than that observed under random distribution of the entries. Small values indicate, therefore, a significant association between the category variables.

The χ^2 statistic measures the significance of an association but not the strength of that association. For this

Table 2
Contingencies between model type and categorical meta-data.

		Zero-order, no lag	Zero-order, with lag	First-order, no lag	First-order, with lag	Mixed-order, no lag	Total
Micro-organism	Shewanella	11	12	7	4	10	44
	Enrichment culture	2	11	0	9	0	22
	Other	2	1	2	0	0	5
Donor	Hydrogen	7	0	5	0	0	12
	Acetate	0	9	1	8	0	18
	Lactate	2	4	2	2	3	13
	Succinate	5	10	0	3	3	21
	Formate	0	0	0	0	2	2
	Other	1	1	1	0	2	5
Acceptor	Solid Mn	7	20	0	11	1	39
	Aqueous Mn	1	0	0	0	6	7
	Other + Mn	7	4	9	2	3	25
Extraction method	HCl	3	11	9	8	1	32
	$CuSO_4$	10	6	0	5	3	24
	Other	1	1	0	0	6	8
	None	1	6	0	0	0	7
Total		15	24	9	13	10	71

Table 3
χ^2 and mutual information calculations for several model-type/categorical meta-data pairs.

Meta-data	Partition of model type	χ^2	Stat. sig.	$U(y\|x)$
Micro-organism	Model	26.5	8.5×10^{-4}	0.14
	Lag	19.7	5.3×10^{-5}	0.23
	Order	7.62	1.1×10^{-1}	0.08
	Plateau	0.331	8.5×10^{-1}	0.003
Donor	Model	60.2	6.7×10^{-6}	0.30
	Lag	31.2	8.5×10^{-6}	0.40
	Order	25.7	4.1×10^{-3}	0.18
	Plateau	5.83	3.2×10^{-1}	0.07
Acceptor	Model	59.8	5.2×10^{-10}	0.25
	Lag	27.2	1.2×10^{-6}	0.32
	Order	37.0	1.8×10^{-7}	0.20
	Plateau	9.10	1.1×10^{-2}	0.10
Extraction method	Model	53.9	2.8×10^{-7}	0.23
	Lag	9.25	2.6×10^{-2}	0.10
	Order	40.6	3.5×10^{-7}	0.26
	Plateau	11.6	8.9×10^{-3}	0.15
Reference	Model	123	1.8×10^{-9}	0.55
	Lag	53.3	1.6×10^{-7}	0.72
	Order	55.6	9.8×10^{-5}	0.40
	Plateau	19.7	5.0×10^{-2}	0.26

purpose, we use the uncertainty coefficient, $U(y|x)$, calculated from information theoretic entropies. Let x and y denote categorical variables that head, respectively, the rows and columns of a contingency table. Letting N denote the total number of observations in the table, the following entropies can be defined:

$$H(y) = -\sum_j \frac{\sum_i N_{i,j}}{N} ln\left(\frac{\sum_i N_{i,j}}{N}\right) \quad (8)$$

$$H(y|x) = -\sum_{i,j} \frac{N_{i,j}}{N} ln\left(\frac{N_{i,j}}{\sum_j N_{i,j}}\right) \quad (9)$$

$H(y)$ is the entropy of the distribution on the category variable, y, without taking into account knowledge of the category variable, x, and $H(y|x)$ is the entropy with knowledge of x taken into account; therefore, $H(y)$ is always greater than $H(y|x)$. The uncertainty coefficient is defined as the fractional difference between $H(y)$ and $H(y|x)$:

$$U(y|x) = \frac{H(y|x) - H(y)}{H(y)} \quad (10)$$

$U(y|x)$ is the fraction of y's entropy that is lost when x is known. When $U(y|x) = 0$, none of y's entropy is lost so there is no relationship between x and y. When $U(y|x) = 1$, all of y's entropy is lost and, therefore, knowledge of x completely determines y (which is the strongest possible association between two variables).

The means of summarizing a quantitative variable are well known. In graphic terms, we have chosen to focus on cumulative distributions rather than histograms because the plot of a cumulative distribution is more sensitive to the goodness-of-fit between data and a proposed parametric distribution. Common summary statistics are the mean, standard deviation, and median which we denote as μ_{arith}, σ_{arith}, and x_{med}. It is important to distinguish arithmetic mean and standard deviation as arithmetic because we will presently make use of the log–normal distribution. The probability density and cumulative distribution functions for the log–normal are given as follows:

Table 4
Contingencies between model type and quantitative meta-data.

		Zero-order, no lag	Zero-order, with lag	First-order, no lag	First-order, with lag	Mixed-order, no lag	Total
[*Cells*] (L^{-1})	3.5×10^7–4×10^7	1	1	0	0	0	5
	9.33×10^8–9.33×10^8	0	4	0	2	0	6
	1.87×10^9–8×10^9	3	6	0	2	3	27
	1×10^{10}–8×10^{10}	2	0	3	0	7	19
	3×10^{11}–3×10^{11}	6	0	5	0	0	11
	1×10^{12}–1×10^{12}	2	0	0	0	0	2
[*Donor*] (mM)	0–0.005	2	0	1	0	3	7
	0.01–0.01	1	2	3	0	2	15
	0.015–0.015	4	16	0	10	3	51
$[Mn]_{total}$ (mM)	1.15×10^{-4}–2x10^{-4}	0	7	0	1	0	23
	2.15×10^{-4}–5.75×10^{-4}	0	7	0	9	3	26
	1.15×10^{-3}–5.75×10^{-3}	2	7	1	3	0	26
	0.01–0.0501	9	3	8	0	7	34
Surface area ($m^2\ L^{-1}$)	0.0085–4.51	0	7	5	6	0	23
	9.02–54	2	6	0	3	0	24
	83.346–3220	8	0	3	0	2	13

$$p(x) = \frac{1}{x \cdot \sigma_{ln} \cdot \sqrt{2\pi}} exp\left(-\frac{(ln(x) - \mu_{ln})^2}{\sigma_{ln}^2}\right) \tag{11}$$

$$p(X \leqslant x) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{ln(x) - \mu_{ln}}{\sigma_{ln} \cdot \sqrt{2}}\right) \tag{12}$$

where *erf* is the error function. Maximum likelihood estimates of μ_{ln} and σ_{ln} can be found from data by taking the arithmetic mean and standard deviation of the log-transformed data:

$$\hat{\mu}_{ln} = \frac{\sum_i ln(x_i)}{N} \tag{13}$$

$$\hat{\sigma}_{ln} = \sqrt{\frac{\sum_i (ln(x_i) - \hat{\mu}_{ln})^2}{N - 1}} \tag{14}$$

where the data are denoted by the x_i and N is the number of data points. For the log–normal distribution, 68.3% of the probability mass lies within the interval [$\exp(\mu_{ln}-\sigma_{ln})$, $\exp(\mu_{ln}+\sigma_{ln})$] and similar confidence intervals can be found at other levels (Limpert et al., 2001). A Monte-Carlo procedure was used to incorporate measurement error into estimates of the log–normal parameters (see below).

For the purpose of measuring goodness-of-fit (e.g., between a data-set and the maximum likelihood log–normal distribution) we chose the non-parametric Kolmogorov–Smirnov statistic which is defined as the maximum distance between two cumulative distributions. Further details can be found in Press et al. (1992).

We took a twofold approach to assess the degree of association between categorical variables and quantitative variables: first, we binned the quantitative variable (thus, forming a categorical variable) and performed a contingency table analysis. Second, we performed analysis of variance (ANOVA) (Snedecor and Cochran, 1989) on the quantitative data using the categorical data as the treatments. ANOVA is dependent on the assumption of homogeneous variance among the treatments; therefore, since we found our best-fit parameters to possess a structure of constant percent error, we performed ANOVA on log-transformed data.

The most familiar measure of association between two quantitative variables is Pearson's correlation coefficient. This measure has, however, some deficiencies for our purposes. First, calculation of the statistic references a parametric (usually, linear) model and we are not in a position to make such assumptions about the data in our compilation. Second, statements concerning the statistical significance of Pearson's correlation coefficient with a moderate number of data points (roughly $500 > N > 10$) require the assumption that the data are drawn from a binormal distribution, which, given the heteroscedasticity observed in our compilation, is an assumption that cannot be supported.

Instead of Pearson's correlation coefficient, we chose to use Spearman's rank-order correlation coefficient (Hollander and Wolfe, 1973). The rank-order correlation coefficient is calculated by ranking the data in ascending order (assigning ties the mean value of the ranks that they span) and then calculating the Pearson's correlation coefficient using the ranks in place of the original values of the data:

$$r_S = \frac{\sum_i (R_i - \overline{R})(S_i - \overline{S})}{\sqrt{\sum_i (R_i - \overline{R})^2 \sum_i (S_i - \overline{S})^2}} \tag{15}$$

where the R_i and S_i are the ranks and $\overline{R}$ and $\overline{S}$ denote the average ranks. Like Pearson's correlation coefficient, r_s ranges from negative one to positive one with negative one indicating a strong inverse association, zero indicating

no association and positive one indicating a strong direct association. Any association that is monotonically increasing or decreasing will be identified by r_s and, most importantly, the statistic:

$$t = r_S\sqrt{\frac{N-2}{1-r_S^2}} \tag{16}$$

is distributed approximately as a Student's distribution with $N-2$ degrees of freedom.

When two variables have a statistically significant rank-order correlation, it should be possible to fit the data with a parametric model. When doing so, it is necessary to take into account the error structure for the dependent variable. In the cases considered here, the dependent variable is always a rate parameter found by fitting the original time-series concentration data and the errors are the estimated uncertainties generated by the time-series fitting procedure. These errors are normally distributed but heteroscedastic (see Figs. 2B and 3B below). As mentioned above, the measurements themselves frequently follow a log–normal distribution. The log–normal distribution is not the error distribution. Rather, it represents the cumulative and deterministic effects of the underlying meta-data. The error—which follows a normal distribution—should be thought of as being superimposed on the log–normal distribution.

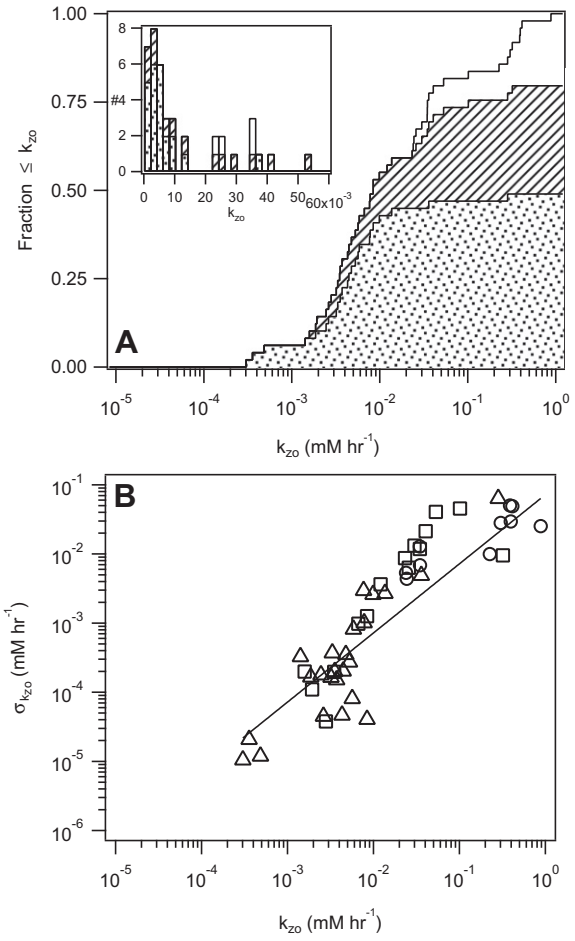

Fig. 2. (A) Cumulative distribution of k_{zo}: the dotted region represents zero-order data with a lag, the hashed region represents zero-order data without a lag, and the blank region represents mixed-order data. (B) Uncertainties on k_{zo} plotted against k_{zo}: (Δ) zero-order data with a lag; (□) zero-order data without a lag; (○) mixed-order data.

The complicated structure of the data presents a problem for parametric fitting. On one hand, the normally distributed and heteroscedastic measurement errors suggest weighted least-squares regression (Seber and Wild, 2003). On the other hand, the log–normal distribution of the measurements suggest un-weighted regression of log-transformed data, a procedure which is inconsistent with the normally distributed measurement errors (Miller, 1984). The most direct solution to this problem is a Monte-Carlo approach wherein un-weighted regression is performed on a number of log-transformed synthetic data-sets, where each synthetic data-set is drawn from a normal distribution based on the original data and the associated errors (Bandstra and Brantley, 2008). We used this Monte-Carlo based fitting approach with 100 iterations for estimating log–normal parameters (μ_{ln}, σ_{ln}) and for fitting rate data and rate constant data as a function of the concentration of cells, [*Cells*], and Mn available for reduction, $[Mn]_{total}$. The resulting parameter distributions were checked for normality visually and by calculating the skew. For all of the cases where the model fit the data well, the magnitude of the skew was less than 1.

3. RESULTS AND DISCUSSION

3.1. Guide to compiled data

Tables EA-1-1 and EA-1-2 in Electronic Annex contain both the raw time-series data and the best-fit rate parameters. Each entry in Table EA-1-1 consists of a reference to a data-set, the time-series, the best-fit parameters, and elements of the corresponding covariance matrix. The reference, which points to the bibliography at the end of Table EA-1-2, includes that paper, figure, and another identifying note allowing the source of each data-set to be uniquely identified. The reference line also indicates which model was used to fit the data. Further instructions for extracting and using the data are provided in Electronic Annex.

3.2. Distribution of model type

The model type that best fit each data-set is a category variable. With three possible reaction orders (zero, first, and mixed) and the possibility of a lag, there are six possible categories. Table 2 shows that the number of data-sets falling into each category decreases from zero-order to first-order to mixed-order. It is interesting that both the zero- and first-order data-sets are characterized by both lag and no-lag behavior; however, there were no mixed-order data-sets exhibiting a lag. Almost the same number of data-sets exhibit an asymptote in reduced Mn concentration as not (32 with an asymptote, 39 without).

The preponderance of data-sets exhibiting some sort of zero-order region (i.e., the zero-order data-sets plus the

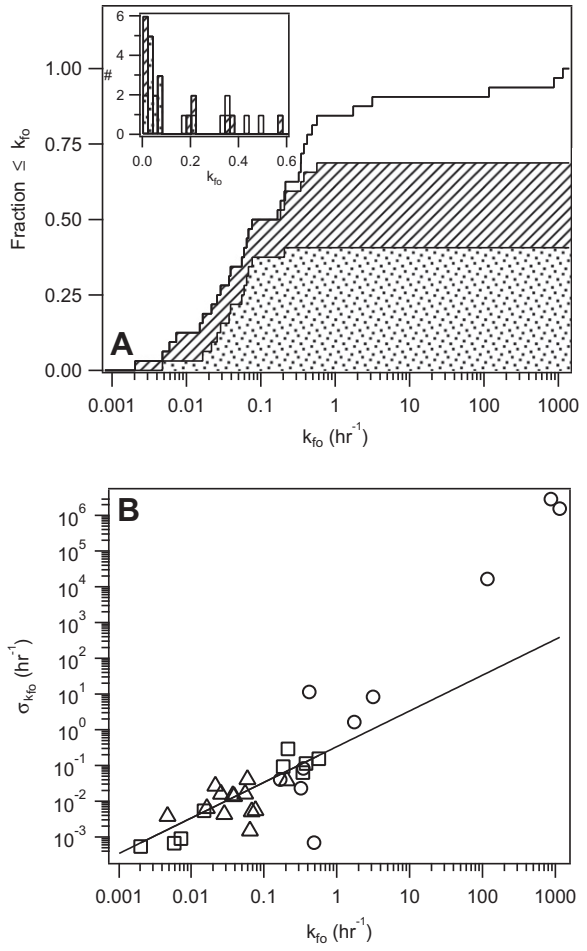

Fig. 3. (A) Cumulative distribution of k_{fo}: the dotted region represents first-order data with a lag, the hashed region represents first-order data without a lag, and the blank region represents mixed-order data. (B) Uncertainties on k_{fo} plotted against k_{fo}: (Δ) first-order data with a lag; (□) first-order data without a lag; (○) mixed-order data.

mixed-order data-sets) indicates that the conditions favored by experimentalists generally supply an excess of both electron donor and acceptor at the beginning of the experiment (e.g., Gaudy et al., 1971; Lovley and Phillips, 1988; Myers and Nealson, 1988b; Kieft et al., 1999; Fredrickson et al., 2002). It is tempting to interpret the greater number of data-sets exhibiting an asymptote as indicative of experiments often being terminated prior to or exactly at the point of exhaustion of an essential ingredient (e.g., before all the Mn is reduced); however, we did not analyze this possibility.

The lag in our kinetic model is not derived from any sort of reaction mechanism, so the interpretation of the rough equivalence between the numbers of data-sets exhibiting a lag and those not exhibiting a lag is unclear. It is possible, however, that the lags are frequently caused by the culture preparation method, in which case the data presented here would indicate that experimentalists are largely unsettled with respect to the method of growing and harvesting DMRBs for use in kinetics experiments. To our knowledge, a systematic study of the effect of culture preparation methods on Mn reaction kinetics has not appeared in the literature; several studies have, however, examined the effects of culture preparation on protein expression (Beliaev et al., 2002, 2005; Tang et al., 2007; Elias et al., 2008).

3.3. Correlations between model type and meta-data

The observations described above hold, with a few exceptions, when the model-type data are partitioned by either categorical (Table 2) or quantitative meta-data (Table 4). The exceptions can be identified most readily by examining the correlation statistics in Tables 3 and 5. Table 3 shows χ^2, the associated statistical significance, and $U(y|x)$ for model-type data partitioned by categorical meta-data (i.e., the data shown in Table 2). Table 5 shows the same statistics for model-type data partitioned by quantitative meta-data (i.e., the data shown in Table 4). In both cases we have computed the statistics based on the five distinct model types indicated in Tables 2 and 4 and, also, with three possible ways of lumping the model types: (i) with/without a lag; (ii) reaction order; and (iii) with/without an asymptote.

As discussed in Section 2, χ^2 is a measure of association between category variables with large values of χ^2 indicating a strong association. Most of the meta-data/model-type pairs show a statistically significant correlation at greater than the 99% level. This is consistent with the conclusion that the type of model that best fits each data-set is dependent on the way in which the experiment was conducted. An exception is when the model types were lumped according to the existence of an asymptote. In this case, the relationships with the meta-data typically show lower significance levels (between 15% and 99%). This indicates that the existence of an asymptote is largely independent of the meta-data; an unsurprising result since any closed system will eventually approach an asymptote as equilibrium is established.

Table 5
χ^2 and mutual information calculations for several model-type/quantitative meta-data pairs.

Meta-data	Partition of model type	χ^2	Stat. sig.	$U(y\|x)$
[*Cells*]	Model	51.3	1.4×10^{-4}	0.41
	Lag	29.0	2.4×10^{-5}	0.63
	Order	21.0	2.1×10^{-2}	0.26
	Plateau	11.1	5.0×10^{-2}	0.20
[*Donor*]	Model	25.9	1.1×10^{-3}	0.20
	Lag	17.9	1.3×10^{-4}	0.32
	Order	7.02	1.3×10^{-1}	0.07
	Plateau	2.47	2.9×10^{-1}	0.04
$[Mn]_{total}$	Model	57.5	1.4×10^{-6}	0.33
	Lag	39.4	5.9×10^{-8}	0.50
	Order	14.3	7.3×10^{-2}	0.14
	Plateau	11.7	1.9×10^{-2}	0.15
Surface area	Model	30.1	2.0×10^{-4}	0.32
	Lag	21.0	2.8×10^{-5}	0.45
	Order	9.39	5.2×10^{-2}	0.14
	Plateau	3.50	1.7×10^{-1}	0.06

Although model type is found to correlate with most of the meta-data, the relationships are weak. This can be seen in the values of $U(y|x)$, whose absolute value varies from zero to one. The strongest relationship (with $U = 0.63$) is between the existence of a lag and the concentration of cells (see Table 5). The reason for this relatively strong relationship can be seen in Table 4, where no lags occur for $[Cells] \geqslant 10^{10}$ cells L.

It is interesting to note that for all the meta-data fields except one (the extraction method), the strongest correlation is found with the existence of a lag. This observation indicates that the existence of a lag is the kinetic feature that is most sensitive to the manner in which an experiment was performed. When H_2 was used as the electron donor, for example, no lags were observed, while lags were observed in most of the experiments where acetate was used as the electron donor. Similarly, when the total Mn concentration was large, few lags were observed relative to when $[Mn]_{total}$ was small.

The result of many statistically significant but weak correlations between the model type and the meta-data is consistent with the multi-factorial character of this data-set. No single meta-data field uniquely determines the type of kinetic behavior that is observed. This is amplified by the correlations found between model type and the paper in which the data were published (see Table 3) which were among the strongest correlations that we found. Experimentalists typically maintain a highly consistent set of meta-data (i.e., experimental conditions) throughout a paper. Each reference is, therefore, a de-facto lumped meta-data field that accounts for at least some of observed variability.

3.4. Distribution of model parameters

There are three parameters in the zero-order model (k_{zo}, t_{lag}, and $[Mn]_0$), four in the first-order model (k_{fo}, t_{lag}, $[Mn]_0$, and $[Mn]_\infty$), and five in the mixed-order model (V_{max}, $K_{1/2}$, t_{lag}, $[Mn]_0$, and $[Mn]_\infty$). The cumulative distributions of k_{zo} and k_{fo} are shown in Figs. 2A and 3A, respectively, along with a histogram of the parameter values. Note that the abscissa is in log format. We have also represented the mixed-order data in these figures by calculating an equivalent k_{zo} and k_{fo} from V_{max} and $K_{1/2}$ (see Section 2).

The cumulative distributions of both k_{zo} and k_{fo} follow a sigmoidal curve when plotted in semi-log format. This behavior is characteristic of distributions with support only on the positive numbers and a right-hand tail that decays at an exponential rate. Some well-known distributions of this type include the Gamma, Weibull, and log–normal distributions. It is notable, given the frequency with which power-law distributions are applied to biogeochemical data (Birdi, 1993; Clark et al., 1995; Turcotte, 2002; Hondzo et al., 2005), that the observed behavior is inconsistent with this type of distribution. Maximum likelihood estimates of the parameters for the log–normal distribution are given in Table 6. In addition to the parameters for the rate constant distributions, we have included parameters for the distributions of t_{lag}, $[Mn]_\infty$, and the initial rate, all of which have cumulative distributions shaped similarly to those of the rate constants (see Figs. EA-1-1 and EA-1-2 in Electronic Annex). In all cases, the maximum likelihood log–normal distribution generates a good fit to the data as measured by the Kolmogorov–Smirnov statistic.

What can be said, then, concerning the best estimate of the rate of bacterial manganese reduction (the best estimate, that is, without any knowledge of the meta-data)? First, the arithmetic mean is not a good measure of central tendency. Note the order of magnitude difference between the means and medians in Table 6. This is due to the positive skew of the distributions. Second, since the log–normal distribution accurately describes the data, $e^{\mu_{ln}}$ should be a good measure of central tendency. In fact it is; the value of $e^{\mu_{ln}}$ closely corresponds with the median value for each of the parameters in Table 6. The best estimate is, therefore, the median of the log–normal distribution, $e^{\mu_{ln}}$. In a similar manner, these results indicate that confidence intervals for rate parameters should be drawn from the log–normal distribution rather than calculated from the sample variance. (Confidence intervals for the log–normal distribution can be found by exponentiation, see Section 2.)

Fitting models to time-series data generates both estimates of the model parameters and estimates of the uncertainty on each model parameter. It is important to understand the error structure associated with estimated model parameters in order to use the parameters correctly. The uncertainties for rate data commonly scale with the rate (or rate constant). This form of heteroscedasticity is evident in Figs. 2B and 3B where the estimated standard errors of k_{zo} and k_{fo}, respectively, are seen to increase in direct proportion to the rate constant. The average percent errors for k_{zo} and k_{fo} as well as for other model parameters (which also exhibited this form of heteroscedasticity, see Figs. EA-1-1 and EA-1-2 in Electronic Annex) are reported in Table 6. The percent errors, ranging from 7% to 34%, are small relative to ranges of the parameter distributions. The

Table 6
Descriptive statistics for distributions on fitted model parameters.

Data	μ_{arith}	σ_{arith}	x_{med}	μ_{ln}	σ_{ln}	K–S (stat. sig.)	Avg. percent error
k_{zo} (mM h^{-1})	0.0778	0.165	0.00842	−4.94	2.25	0.119 (0.467)	0.0720
k_{fo} (h^{-1})	67.0	251	0.0682	−2.29	3.26	0.201 (0.132)	0.335[a]
t_{lag}[b] (h)	22.5	20.4	19.0	2.50	1.45	0.145 (0.392)	0.226
$[Mn]_{inf}$ (mM)	3.08	6.12	0.486	−0.552	1.75	0.211 (0.189)	0.0707
Init. rate (mM h^{-1})	0.0679	0.128	0.0123	−4.27	1.96	0.0627 (0.935)	0.112

[a] Excluded the three points with $k_{fo} > 10\ h^{-1}$ to avoid leveraging problems caused by these points.
[b] Excluded all zeros from the calculations.

parameter uncertainties are, therefore, not important when estimating the rate of bacterial manganese reduction; when analyses are performed on the fitted parameters (e.g., examining for correlations between rates and meta-data as we do below), however, the error structure must be taken into account.

It is interesting to note differences in the rate distributions based solely on the best-fit model type. For both k_{zo} and k_{fo}, rate constants from data that did not exhibit a lag were greater than rate constants from data that did exhibit a lag (though, the statistical significance of the difference is fairly weak in both cases: for the zero- and first-order cases, respectively, the *K–S* statistics are 0.47 and 0.48 with significance levels of 0.02 and 0.12). The difference is unsurprising since we had previously found the existence of a lag to be linked with lower cell concentrations and one would expect these conditions to produce slower rates. In addition, the rate constants estimated from the mixed-order model were greater than those obtained using the zero- and first-order models, respectively. (The differences in this case are statistically significant: *K–S* statistics of 0.612 and 0.53 with significance levels of 2×10^{-3} and 0.018.) It is not intuitive why this should be the case; though, since there are only 10 mixed-order cases in our compilation, it seems unlikely that any hypotheses could be tested with this data-set.

3.5. Correlations between initial rates and meta-data

The distributions of model parameters described above are, presumably, due to the differences in the way each underlying experiment was performed. ANOVA calculations performed on log-transformed initial rates parsed according to the same meta-data categories used in Table 2 showed significant treatment effects (at levels >99%) for each of the four categories. Tests for pairwise treatment effects (Tukey HSD and Newman–Keuls) within each meta-data category likewise indicate differences in the rate due to the meta-data with a few notable exceptions: (i) the data with H_2, lactate, and formate as the electron donor are not significantly different from one another; (ii) the data with acetate as the donor are not significantly different from the data with succinate as the donor; and (iii) the data with HCl extraction, $CuSO_4$ extraction, and no extraction are not significantly different from each other. The full ANOVA reports are available in Table EA-1-3 in Electronic Annex.

The ANOVA results demonstrate that the type of microorganism and type of electron acceptor strongly influence the rate of reaction, while the donor type and extraction method have, at most, a weak influence on rates. In fact, the higher median rates for data with either H_2, formate, or lactate as the donor, as compared to data with either acetate or succinate as the donor, coincides with the former data being dominated by relatively high $[Cells]$ and $[Mn]_{total}$ and the latter being dominated by relatively low values of those variables. The lack of effect due to the type of donor is consistent with the concentration of donor being present in excess for most experimental designs (e.g., Lovley and Phillips, 1988; Myers and Nealson, 1988b; Kieft et al., 1999; Fredrickson et al., 2002).

Table 7 shows Spearman rank-order correlation coefficients between initial rates and four of the quantitative meta-data variables ($[Cells]$, $[Donor]$, $[Mn]_{total}$, and Surface Area), with statistically significant relationships underlined. The top column in Table 7 is for all the data treated together. In this case, statistically significant, but weak correlations exist between initial rate and cell concentration, donor concentration, and total Mn. That rates should increase with $[Cells]$ and $[Mn]_{total}$ is intuitive (Liu et al., 2001a; Burgos et al., 2003); our results also indicate, however, that rates are statistically independent of Surface Area and that they decrease with $[Donor]$. We discuss the reason for this below.

The subsequent columns in Table 7 show correlation coefficients for initial rates parsed by the reaction order that best described the data. The correlations of initial rate with $[Cells]$ and $[Mn]_{total}$ are preserved when the data are so parsed. These correlations are also broadly preserved when the initial rate data are parsed by categorical meta-data (see Table EA-1-4). Plotting initial rates vs. $[Cells]$ and $[Mn]_{total}$ (see Fig. 4A and B, respectively) suggests patterns that are consistent with hyperbolic rate equations (Liu et al., 2001b; Roden, 2008). Monte-Carlo based log–normal regression of such an equation gives the curves shown in Fig. 4A and B:

$$init.\ rate = \frac{(0.0129 \pm 0.007)\frac{\mathrm{mM}}{h}[Cells]}{(3.57 \times 10^{10} \pm 2.5 \times 10^{10})\frac{1}{\mathrm{L}} + [Cells]}$$

$$init.\ rate = \frac{(0.0318 \pm 0.0106)\frac{\mathrm{mM}}{h}[Mn]_{total}}{(0.00196 \times 0.0001)\mathrm{M} + [Mn]_{total}}$$

It is noteworthy that most of the rates fall above the *Shewanella* + solid Mn data (circles with dots in Fig. 4) indicating that these experiments represent a lower bound for the rate of bacterial Mn reduction.

As noted above, correlation analyses on the lumped initial rate data indicate an inverse relationship between rate and $[Donor]$. Parsing the data by reaction order in Table 7 partially diminishes this relationship; furthermore, when the initial rate data are parsed by the type of donor, significant correlations are not found (see Table EA-1-4). This is

Table 7
Maximum likelihood estimates of log–normal parameters and Spearman's rank-order correlations (statistical significance in parentheses) for initial rate data parsed by the reaction order best describing the data. Units for initial rates are mM h^{-1}.

	μ_{ln}	σ_{ln}	$[Cells]$	$[Donor]$	$[Mn]_{total}$	Surface area
All data	−4.27	1.96	0.677 (1.3×10^{-10})	−0.553 (3.9×10^{-7})	0.583 (2.9×10^{-11})	0.292 (0.025)
Zero-order data	−5.76	2.00	0.697 (1.1×10^{-4})	−0.441 (0.027)	0.522 (6.5×10^{-4})	0.638 (1.1×10^{-3})
First-order data	−4.36	2.11	0.774 (3.1×10^{-3})	−0.709 (4.5×10^{-3})	0.678 (5.3×10^{-4})	0.399 (0.11)
Mixed-order data	−2.67	1.72	0.454 (0.19)	−0.709 (0.049)	0.557 (0.094)	N/A

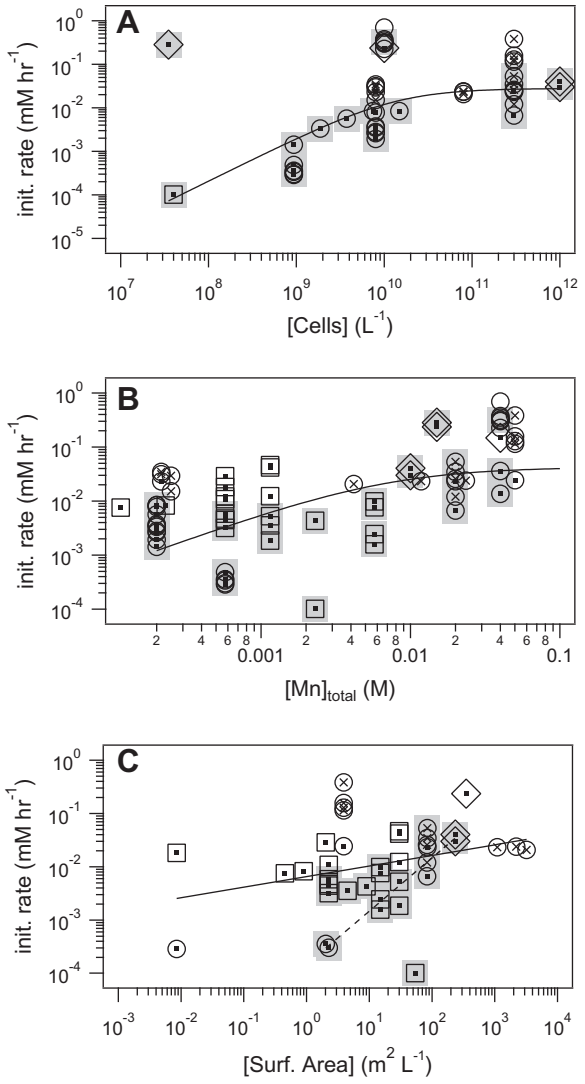

Fig. 4. Initial rate data plotted against (A) the concentration of cells, (B) the concentration of Mn available for reduction, and (C) the surface area of Mn solid per liter of solution. Micro-organism type is represented as: (○) *Shewanella*; (□) enrichment culture; and (◇) other. Electron acceptor type is represented as: (■) solid Mn; (×) Mn plus some other electron acceptor; and open symbol aqueous Mn. Data best fit by the zero-order model are highlighted in gray.

because each type of donor was used almost exclusively at a single concentration and, therefore, these data cannot be used to establish a trend between rate and [*Donor*]. (The relationship, or lack thereof, can be examined further in Fig. EA-1-3 in Electronic Annex.)

The lack of a statistically significant relationship between initial rate and surface area in Table 7 is surprising since heterogeneous reactions rates typically scale directly with available reaction sites on the surface, and since surface area normalized behavior has been well documented for bacterial iron reduction (Roden, 2008). Parsing initial rates by reaction order gives a significant correlation with surface area for the zero-order data, but not for the first-order data. This result can also be seen in Fig. 4C where initial rate is plotted against surface area. Fitting all the data together with a power law shows only a mild upward trend (the solid line in Fig. 4C), while the zero-order data—highlighted in gray—are consistent with a surface area normalized rate of 1.47×10^{-4} (mmol h^{-1} m^{-2}) depicted by the dashed line.

3.6. Best estimates of zero- and first-order rate constants

We have already noted that, when no recourse is made to the meta-data, best estimates—and corresponding confidence intervals—for the zero- and first-order rate constants can be found by log–normal analysis (see Table 6). We should, however, be able to improve our estimates of the log–normal parameters using knowledge of the meta-data.

As discussed above, ANOVA results with (log-transformed) initial rates indicate significant treatment effects due to the type of micro-organism and type of electron acceptor used. It is, therefore, reasonable to expect the distributions of k_{zo} and k_{fo} to depend on these categorical meta-data. The first two columns of Tables 8 and 9 include log–normal parameters for, respectively, k_{zo} and k_{fo} data that have been parsed according to the type of micro-organism and electron acceptor. These results further demonstrate the conclusions drawn from the initial rate ANOVA calculations so that, for example, if it is known that solid Mn is serving as the sole electron acceptor in a system then the best estimate of k_{zo}—based on the data in our compilation—would be a log–normally distributed random variable with $\mu_{ln} = -6.27$ and $\sigma_{ln} = 2.11$.

The last four rows of both Tables 8 and 9 pertain to rate constant data that have been parsed by both the type of micro-organism and the type of electron acceptor (NB: the enrichment culture data were all collected with solid Mn as the electron acceptor). If it is known, for example, that solid Mn is the sole electron acceptor and that the micro-organisms are a strain of *Shewanella*, then the appropriate log–normal parameters are $\mu_{ln} = -6.24$ and $\sigma_{ln} = 1.7$. It is interesting that the results obtained by parsing the data according to both of these categorical meta-data are not significantly different from the corresponding data parsed only by the type of electron acceptor (based on *F*-tests with log-transformed rate constants). This is a trivial result in the case of either aqueous Mn or Mn + Other as the electron acceptor since these acceptors were only used with *Shewanella* as the micro-organism. In the case of solid Mn as the electron acceptor, however, the result indicates that knowledge of the type of electron acceptor is more useful for predicting rates than knowledge of the type of micro-organism—at least, when comparing strains of *Shewanella* to organisms that have been cultured from sediments and soils (Liu et al., 2002). It would be interesting to make similar comparisons with other well defined microbial strains; unfortunately, Mn reduction data with other micro-organisms are not available in sufficient numbers to make such a comparison.

The correlation analyses of initial rate data discussed above indicate that the rates depend on [*Cells*] and $[Mn]_{total}$. These quantitative meta-data, therefore, should also inform

Table 8
Best estimates of the zero-order rate constant as a function of the meta-data.

Data-set	Constant		[*Cells*] dependent			$[Mn]_{\text{total}}$ dependent			Surf. area dependent		
	μ_{ln}	σ_{ln}	V (mM h^{-1})	K (cells L^{-1})	σ_{ln}	V (mM h^{-1})	K (M)	σ_{ln}	A	Pow	σ_{ln}
Total (49)	−4.94	2.25	*0.0105 ± 0.0020*	*(5.91 ± 43) × 10^{5}*	*3.84*	**0.0465 ± 0.023**	**(9.29 ± 6.9) × 10^{-3}**	**1.92**	(6.70 ± 3.0) × 10^{-4}	0.586 ± 0.13	2.40
Shewanella (33)	−4.75	2.39	**0.0256 ± 0.0067**	**(1.61 ± 0.65) × 10^{10}**	**2.33**	0.0538 ± 0.021	(6.79 ± 4.6) × 10^{-3}	1.93	(3.07 ± 2.7) × 10^{-4}	0.752 ± 0.17	2.34
Enrichment culture (13)	−5.90	1.44				(4.49 ± 2.6) × 10^{-3}	(7.79 ± 16) × 10^{-4}	1.45	(1.65 ± 1.1) × 10^{-3}	0.326 ± 0.29	2.75
Solid Mn (35)	−5.78	1.80	*(3.67 ± 0.58) × 10^{-3}*	*(2.28 ± 20) × 10^{-9}*	*4.19*	*(8.14 ± 4.9) × 10^{-3}*	*(1.23 ± 1.5) × 10^{-3}*	*1.71*	(6.74 ± 3.2) × 10^{-3}	0.579 ± 0.19	2.68
Aqueous Mn (7)	−1.61	1.12									
Other + Mn (7)	−4.20	1.12	0.0192 ± 0.0097	(6.36 ± 16) × 10^{9}	1.04	0.0183 ± 0.0063	(1.34 ± 3.4) × 10^{-3}	1.04	0.0790 ± 0.26	0.00104 ± 0.35	2.03
She. + solid (19)	−6.14	1.75	(9.79 ± 6.3) × 10^{-3}	(2.27 ± 2.1) × 10^{10}	1.92	(7.32 ± 5.6) × 10^{-3}	(9.92 ± 15) × 10^{-4}	1.69	(1.46 ± 0.96) × 10^{-4}	0.914 ± 0.27	2.84
She. + aq. (7)	−1.57	1.12									
She. + oth. (7)	−4.21	1.12	0.0179 ± 0.0074	(2.16 ± 5.1) × 10^{9}	1.09	0.0181 ± 0.0069	(1.33 ± 3.6) × 10^{-4}	1.07	0.157 ± 0.80	−0.00043 ± 0.36	2.06
Enrich. + solid (13)	−5.90	1.50				(4.37 ± 4.7) × 10^{-3}	(1.28 ± 5.3) × 10^{-3}	1.47	(1.56 ± 0.91) × 10^{-3}	0.334 ± 0.25	2.74

Table 9
Best estimates of the first-order rate constant as a function of the meta-data.

Data-set	Constant		[*Cells*] dependent			$[Mn]_{total}$ dependent			Surf. area dependent		
	μ_{ln}	σ_{ln}	V (h^{-1})	K (cells L^{-1})	σ_{ln}	V (h^{-1})	K (M L^{-1})	σ_{ln}	A	Pow	σ_{ln}
Total (32)	**−2.29**	**3.26**	**0.297 ± 0.065**	**(3.31 ± 1.4) × 10^{9}**	**3.76**	*0.114 ± 0.032*	*(7.41 ± 22) × 10^{-5}*	*3.23*	0.0387 ± 0.012	−0.0723 ± 0.087	1.74
Shewanella (21)	−1.37	3.64	0.349 ± 0.091	(3.24 ± 1.1) × 10^{9}	3.61	*0.257 ± 0.068*	*(1.08 ± 5.0) × 10^{-5}*	*3.60*	0.0706 ± 0.027	−0.135 ± 0.090	1.69
Enrichment culture (9)	−3.76	1.07				*2.90 ± 22*	*0.0878 ± 0.62*	*0.995*	0.0257 ± 0.010	0.0605 ± 0.15	1.81
Solid Mn (15)	−3.81	1.81	0.224 ± 0.16	(2.54 ± 3.3) × 10^{10}	3.74	*0.0240 ± 0.0062*	*(1.00 ± 5.6) × 10^{-5}*	*1.85*	0.0235 ± 0.0074	0.0288 ± 0.10	1.88
Aqueous Mn (6)	0.793	4.38									
Mn + other (11)	−2.02	3.04	*0.132 ± 0.045*	*(6.62 ± 29) × 10^{-7}*	*3.05*	*0.138 ± 0.043*	*(−4.28 ± 9.5) × 10^{-20}*	*3.01*	0.423 ± 0.24	−0.478 ± 0.15	1.64
She. + solid (4)	−3.14	3.05	*0.396 ± 0.22*	*(2.71 ± 3.0) × 10^{10}*	*2.09*						
She. + aq. (6)	0.909	4.31									
She. + oth. (11)	−2.03	3.02	*0.154 ± 0.037*	*(−7.84 ± 26) × 10^{-7}*	*3.01*				0.408 ± 0.22	−0.483 ± 0.13	1.64
Enrich. + solid (9)	−3.84	1.17							0.0256 ± 0.010	0.0362 ± 0.14	1.81

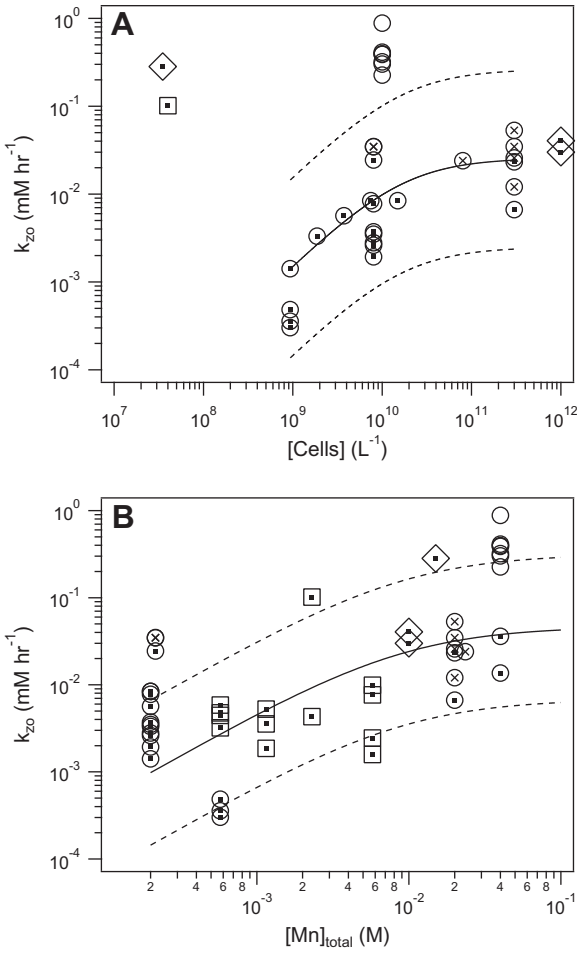

Fig. 5. Zero-order rate constants plotted against (A) the concentration of cells and (B) the concentration of Mn available for reduction. Micro-organism type is represented as: (○) *Shewanella*; (□) enrichment culture; and (◇) other. Electron acceptor type is represented as: (■) solid Mn; (×) Mn plus some other electron acceptor; and open symbol aqueous Mn.

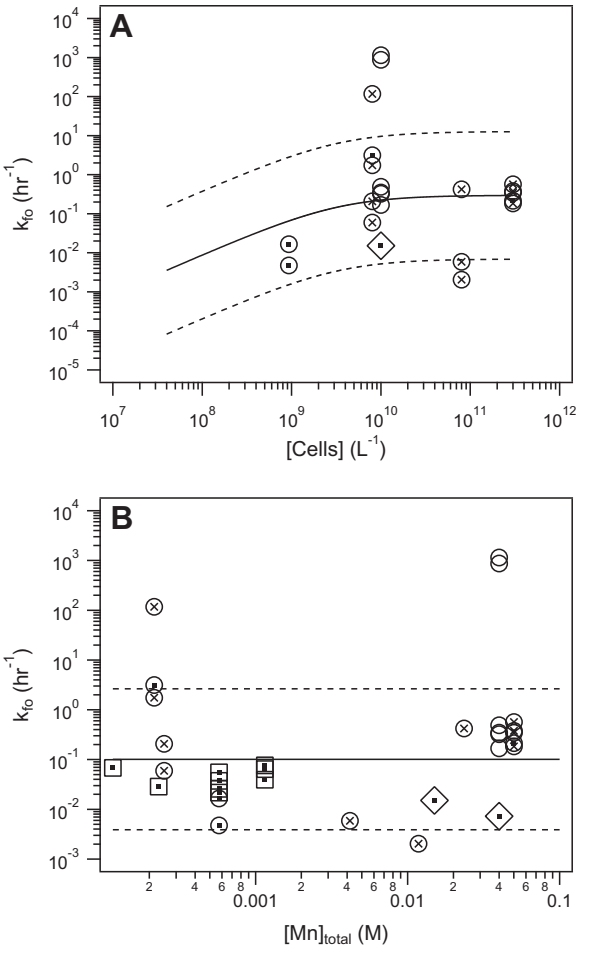

Fig. 6. First-order rate constants plotted against (A) the concentration of cells and (B) the concentration of Mn available for reduction. Micro-organism type is represented as: (○) *Shewanella*; (□) enrichment culture; and (◇) other. Electron acceptor type is represented as: (■) solid Mn; (×) Mn plus some other electron acceptor; and open symbol aqueous Mn.

our best estimates of the rate constants. Log–normal analysis can be adapted to this end by treating $\exp(\mu_{ln})$ as a function of $[Cells]$ and/or $[Mn]_{total}$. Tables 8 and 9 show the results of such an analysis using hyperbolic rate equations:

$$k_x = V_{max}\frac{[\cdot]}{K_{1/2} + [\cdot]}$$

where k_x can be either the zero- or the first-order rate constant and $[\cdot]$ represents either $[Cells]$ or $[Mn]_{total}$.

For some of the data-sets, there are too few points to perform reliable analyses; these have been left blank in Tables 8 and 9. For other data-sets, the analysis returned a value of $K_{1/2}$ that was smaller than the measured $[Cells]$ or $[Mn]_{total}$; these cases have been printed in italics. A value of $K_{1/2}$ that is so small indicates that the observed rate constants do not show a trend with the relevant concentration variable. For both of these cases, the best possible estimate of the rate constant is the log–normal distribution with no concentration dependence. The case of k_{fo} vs. $[Mn]_{total}$ bears special mention in this regard since no trend was identified for these data regardless of the meta-data. This is consistent with the fact that the first-order model already takes account of the concentration of reducible Mn.

For the remaining data-sets, the effect of the concentration variables could be reliably extracted from the rate constant data. If $[Cells]$ or $[Mn]_{total}$ data are, therefore, known for a system—or if a simulation is performed that treats these concentrations dynamically—then the best estimate of the rate constant should be evaluated taking these concentration data into account using a log–normal distribution with the concentration-dependent model parameters given in Tables 8 and 9.

Tables 8 and 9 also show the results of log–normal Monte-Carlo analysis with surface area as the independent variable. Here, μ_{ln} has been modeled with a power law dependence on surface area ($k_x = A[Surf.\ Area]^{pow}$). The results are entirely consistent with our analysis of initial rates in that k_{zo} shows a broad dependence on surface area while k_{fo} does not. In general we have found $[Mn]_{total}$ to be a

stronger predictor variable than surface area. These variables are, however, strongly covariant; therefore, our result is likely due to the greater abundance and quality of $[Mn]_{total}$ data rather than any mechanistic feature of the system.

Figs. 5 and 6 show, respectively, k_{zo} and k_{fo} plotted against both $[Cells]$ and $[Mn]_{total}$. Also plotted on these graphs are curves representing the median and outer limits of the 68% confidence interval based on the model that we judge to best represent the data as a whole (whose parameters are printed in bold in Tables 8 and 9). The median value is calculated as $\exp(\mu_{ln})$ and the 68% confidence interval is calculated as $[\exp(\mu_{ln} - \sigma_{ln}),\ \exp(\mu_{ln} + \sigma_{ln})]$. With one exception, we find treating all the data together to be a robust method for generating rate equations even though there are statistically significant differences based on the categorical meta-data. Even in the exception (the case of k_{zo} vs. $[Cells]$), it is only the two experiments performed at very low $[Cells]$ that prevent an accurate fit that includes all the data. We find, therefore, that even though the particular details of an experiment do affect reduction rates—even in a narrowly defined statistical sense—it is still possible to observe trends in the rate data at a broad level of aggregation. The compilation of kinetics data and analyses of these data in aggregate, then, offers a means for predicting rates of environmentally relevant reactions.

The compilation of kinetics data also offers an avenue for identifying new areas for research. Our results point to gaps in the knowledge of factors that control the reaction order and the existence of lag. We also see that the effect of donor concentration on reduction rate has not been systematically examined. In addition, the data are sparse for micro-organisms other than strains of *Shewanella*—and especially so for experiments with measured concentration of cells. New experiments would be warranted, therefore, at lower donor concentrations and with a wider array of micro-organisms.

Finally, the compilation of laboratory rates of Mn reduction provide an opportunity for comparison with observations in natural systems. Techniques have been proposed for the determination of bacterial Mn reduction rates in both groundwater (Istok et al., 1997) and marine sediments (Wang et al., 2008). Since these techniques require knowledge not only of Mn concentrations but also the concentration of other electron acceptors and flow conditions, the rates so determined contain several additional sources of variability. Nevertheless, the surface area normalized Mn reduction rates reported by D'Hondt et al. (2004)—which vary between 2.5×10^{-8} (mmol m^{-2} h^{-1}) and 6.96×10^{-5} (mmol m^{-2} h^{-1})—are generally consistent with both our finding of the log–normal distribution of rates and the average surface area normalized rate (for zero-order data) in our compilation of 1.47×10^{-4} (mmol m^{-2} h^{-1}). Further comparisons of lab rates with field data are, however, clearly needed.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CHE-0431328. J.Z.B. thanks Ms. Beatriz Sancho for assistance with Monte-Carlo analyses. We also acknowledge the helpful comments of the anonymous reviewers.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gca.2010.04.069.

REFERENCES

Agrawal A., Ferguson W. J., Christ J. A., Bandstra J. Z. and Tratnyek P. G. (2002) Effects of carbonate species on the kinetics of dechlorination of 1,1,1-trichloroethane by zero-valent iron. *Environ. Sci. Technol.* **36**, 4326–4333.

Bandstra J. Z. and Brantley S. L. (2008) Data fitting techniques with applications to mineral dissolution kinetics. In *Kinetics of Water–Rock Interaction* (eds. S. L. Brantley, J. D. Kubicki and A. F. White). Springer, New York, pp. 211–257.

Beliaev A. S., Thompson D. K., Khare T., Lim H., Brandt C. C., Li G., Murray A. E., Heidelberg J. F., Giometti C. S., Yates J., Nealson K. H., Tiedje J. M. and Zhou J. (2002) Gene and protein expression profiles of *Shewanella oneidensis* during anaerobic growth with different electron acceptors. *OMICS: J. Integr. Biol.* **6**, 39–60.

Beliaev A. S., Klingeman D. M., Klappenbach J. A., Wu L., Romine M. F., Tiedje J. M., Nealson K. H., Fredrickson J. K. and Zhou J. (2005) Global transcriptome analysis of *Shewanella oneidensis* MR-1 exposed to different terminal electron acceptors. *J. Bacteriol.* **187**, 7138–7145.

Birdi K. S. (1993) *Fractals in Chemistry, Geochemistry, and Biophysics: An Introduction*. Springer, New York.

Brantley S. L., Kubicki J. D. and White A. F. (2008) *Kinetics of Water–Rock Interaction*. Springer, New York.

Bratina B. J., Stevenson B. S., Green W. J. and Schmidt T. M. (1998) Manganese reduction by microbes from oxic regions of the lake vanda (Antarctica) water column. *Appl. Environ. Microbiol.* **64**, 3791–3797.

Brent R. P. (1973) *Algorithms for Minimization Without Derivatives*. Prentice-Hall, Englewood Cliffs, NJ.

Burdige D. J. and Nealson K. H. (1985) Microbial manganese reduction of enrichment cultures from coastal marine sediments. *Appl. Environ. Microbiol.* **50**, 491–497.

Burdige D. J., Dhakar S. P. and Nealson K. H. (1992) Effects of manganese oxide mineralogy on microbial and chemical manganese reduction. *Geomicrobiol. J.* **10**, 27–48.

Burgos W. D., Fang Y. L., Royer R. A., Yeh G. T., Stone J. J., Jeon B. H. and Dempsey B. A. (2003) Reaction-based modeling of quinone-mediated bacterial iron(III) reduction. *Geochim. Cosmochim. Acta* **67**, 2735–2748.

Clark M. B., Brantley S. L. and Fisher D. M. (1995) Power-law vein-thickness distributions and positive feedback in vein growth. *Geology* **23**, 975–978.

D'Hondt S., Jørgensen B. B., Miller D. J., Batzke A., Blake R., Cragg B. A., Cypionka H., Dickens G. R., Ferdelman T., Hinrichs K.-U., Holm N. G., Mitterer R., Spivack A., Wang G., Bekins B., Engelen B., Ford K., Gettemy G., Rutherford S. D., Sass H., Skilbeck C. G., Aiello I. W., Guèrin G., House C. H., Inagaki F., Meister P., Naehr T., Niitsuma S., Parkes R. J., Schippers A., Smith D. C., Teske A., Wiegel J., Padilla C. N. and Acosta J. L. S. (2004) Distributions of microbial activities in deep subseafloor sediments. *Science* **306**, 2216–2221.

Elias D. A., Tollaksen S. L., Kennedy D. W., Mottaz H. M., Giometti C. S., McLean J. S., Hill E. A., Pinchuk G. E., Lipton

M. S., Fredrickson J. K. and Gorby Y. A. () The inxuence of cultivation methods on *Shewanella oneidensis* physiology and proteome expression. *Arch. Microbiol.* **189**, 313–324.

Fredrickson J. K., Zachara J. M., Kennedy D. W., Liu C., Duff M. C., Hunter D. B. and Dohnalkova A. (2002) Influence of Mn oxides on the reduction of uranium(VI) by the metal-reducing bacterium *Shewanella putrefaciens*. *Geochim. Cosmochim. Acta* **66**, 3247–3262.

Fredrickson J. K., Zachara J. M., Kennedy D. W., Kukkadapu R. K., McKinley J. P., Heald S. M., Liu C. and Plymale A. E. (2004) Reduction of TcO_4^- by sediment-associated biogenic Fe(II). *Geochim. Cosmochim. Acta* **68**, 3171–3187.

Gaudy, Jr., A. F., Obayashi A. and Gaudy E. T. (1971) Control of growth rate by initial substrate concentration at values below maximum rate. *Appl. Environ. Microbiol.* **22**, 1041–1047.

Greene A. C., Patel B. K. C. and Sheehy A. J. (1997) *Deferribacter thermophilus* gen. nov., sp. nov., a novel thermophilic manganese- and iron-reducing bacterium isolated from a petroleum reservoir. *Int. J. Syst. Bacteriol.* **47**, 505–509.

Hollander M. and Wolfe D. A. (1973) *Nonparametric Statistical Methods*. Wiley, New York.

Hondzo M., Feyaerts T., Donovan R. and O'Connor B. L. (2005) Universal scaling of dissolved oxygen distribution at the sediment–water interface: a power law. *Limnol. Oceanogr.* **50**, 1667–1676.

Istok J. D., Humphrey M. D., Schroth M. H., Hyman M. R. and O'Reilly K. T. (1997) Single-well, "push–pull" test for in situ determination of microbial activities. *Ground Water* **35**, 619–631.

Kieft T. L., Fredrickson J. K., Onstott T. C., Gorby Y. A., Kostandarithes H. M., Bailey T. J., Kennedy D. W., Li S. W., Plymale A. E., Spandoni C. M. and Gray M. S. (1999) Dissimilatory reduction of Fe(III) and other electron acceptors by *Thermus* isolate. *Appl. Environ. Microbiol.* **65**, 1214–1221.

Kostka J. E., Luther G. W. and Nealson, III, K. H. (1995) Chemical and biological reduction of Mn(III)–pyrophosphate complexes: potential importance of dissolved Mn(III) as an environmental oxidant. *Geochim. Cosmochim. Acta* **59**, 885–894.

Limpert E., Stahel W. A. and Abbt M. (2001) Log–normal distributions across the sciences: keys and clues. *Bioscience* **51**, 341–352.

Liu C., Zachara J. M., Gorby Y. A., Szecsody J. E. and Brown C. F. (2001a) Microbial reduction of Fe(III) and sorption/precipitation of Fe(II) on *Shewanella putrefaciens* strain CN32. *Environ. Sci. Technol.* **35**, 1385–1393.

Liu C., Kota S., Zachara J. M., Fredrickson J. K. and Brinkman C. K. (2001b) Kinetic analysis of the bacterial reduction of geothite. *Environ. Sci. Technol.* **35**.

Liu C., Gorby Y. A., Zachara J. M., Fredrickson J. K. and Brown C. F. (2002) Reduction kinetics of Fe(III), Co(III), U(VI), Cr(VI), and Tc(VII) in cultures of dissimilatory metal-reducing bacteria. *Biotechnol. Bioeng.* **80**, 637–649.

Lovley D. R. and Phillips E. J. P. (1988) Novel mode of microbial energy metabolism: organic carbon oxidation coupled to dissimilatory reduction of iron or manganese. *Appl. Environ. Microbiol.* **54**, 1472–1480.

Miller D. C. (1984) Reducing transformation bias in curve fitting. *Am. Stat.* **38**, 124–126.

Myers C. R. and Nealson K. H. (1988a) Microbial reduction of manganese oxides: interactions with iron and sulfur. *Geochim. Cosmochim. Acta* **52**, 2727–2732.

Myers C. R. and Nealson K. H. (1988b) Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor. *Science* **240**, 1319–1321.

Olsen A. A. and Rimstidt J. D. (2007) Using a mineral lifetime diagram to evaluate the persistence of olivine on mars. *Am. Mineral.* **92**, 598–602.

Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P. (1992) *Numerical Recipes in C*. Cambridge University Press, Cambridge.

Roden E. A. (2008) Microbiological controls on geochemical kinetics: 1. Fundamentals and case study on microbial Fe(III) oxide reduction. In *Kinetics of Water–Rock Interaction* (eds. S. L. Brantley, J. D. Kubicki and A. F. White). Springer, New York.

Rusin P. A., Quintana L., Sinclair N. A., Arnold R. G. and Oden K. L. (1991) Physiology and kinetics of manganese-reducing *Bacillus polymyxa* strain D1 isolated from manganiferous silver ore. *Geomicrobiol. J.* **9**, 13–25.

Seber G. A. F. and Wild C. J. (2003) *Nonlinear Regression*. John Wiley & Sons, New York.

Snedecor G. W. and Cochran W. G. (1989) *Statistical Methods*. Iowa State University Press, Ames, IA.

Tang Y. J., Hwang J. S., Wemmer D. E. and Keasling J. D. (2007) *Shewanella oneidensis* MR-1 fluxome under various oxygen conditions. *Appl. Environ. Microbiol.* **73**, 718–729.

Turcotte D. L. (2002) Fractals in petrology. *Lithos* **65**, 261–271.

Wang G., Spivack A. J., Rutherford S., Manor U. and D'Hondt S. (2008) Quantification of co-occurring reaction rates in deep subseafloor sediments. *Geochim. Cosmochim. Acta* **72**, 3479–3488.

Associate editor: Susan Glasauer