Environmental Science & Technology

Article

pubs.acs.org/est

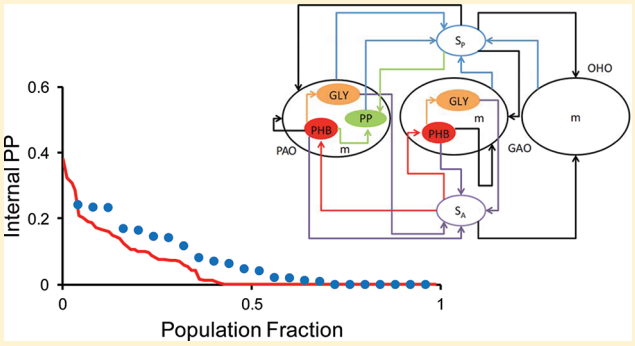
Heterogeneity of Intracellular Polymer Storage States in Enhanced Biological Phosphorus Removal (EBPR) – Observation and Modeling

Vanni Bucci,[†,‡,⊥] Nehreen Majed,[†,§,⊥] Ferdi L. Hellweger,*[,†] and April Z. Gu[†]

[†]Dept. of Civil and Environmental Engineering, Northeastern University, Boston Massachusetts 02115, United States .

Supporting Information

ABSTRACT: A number of agent-based models (ABMs) for biological wastewater treatment processes have been developed, but their skill in predicting heterogeneity of intracellular storage states has not been tested against observations due to the lack of analytical methods for measuring single-cell intracellular properties. Further, several mechanisms can produce and maintain heterogeneity (e.g., different histories, uneven division) and their relative importance has not been explored. This article presents an ABM for the enhanced biological phosphorus removal (EBPR) treatment process that resolves heterogeneity in three intracellular polymer storage compounds (i.e., polyphosphate, polyhydroxybutyrate, and glycogen) in three functional microbial populations (i.e., polyphosphate-accumulating, glycogen-accumulating, and ordinary heterotrophic organisms). Model predicted distributions were compared to those based on single-cell estimates obtained using a Raman microscopy method for a laboratory-scale sequencing batch reactor (SBR) system. The model can reproduce many features of the observed heterogeneity. Two methods for introducing heterogeneity were evaluated. First, biological variability in individual cell behavior was simulated by randomizing model parameters (e.g., maximum acetate uptake rate) at division. This method produced the best fit to the data. An optimization algorithm was used to determine the best variability (i.e., coefficient of variance) for each parameter, which suggests large variability in acetate uptake. Second, biological variability in individual cell states was simulated by randomizing state variables (e.g., internal nutrient) at division, which was not able to maintain heterogeneity because the memory in the internal states is too short. These results demonstrate the ability of ABM to predict heterogeneity and provide insights into the factors that contribute to it. Comparison of the ABM with an equivalent population-level model illustrates the effect of accounting for the heterogeneity in models.



1. INTRODUCTION

Enhanced biological phosphorus removal (EBPR) is a widely used process for removing phosphate from wastewater to avoid eutrophication of receiving water systems. The key microbes in the EBPR process are the polyphosphate-accumulating organisms (PAOs). Under anaerobic conditions, PAOs utilize intracellular polyphosphate (PP) and glycogen (GLY) as energy source and reducing power respectively to uptake and reduce volatile fatty acid (VFA) into polyhydroxyalkanoates (PHAs). Then, in a subsequent aerobic phase, PAOs can use their stored PHAs as energy source to support biomass growth, recover cellular glycogen, and restore PP accompanied by phosphate uptake. Net phosphorus (P) removal from the wastewater is achieved by removing biomass with high P content. Another functional group of population that affects EBPR process performance are the glycogen-accumulating organisms (GAOs). Under anaerobic conditions, GAOs also take-up VFAs for PHA formation while using glycogen as the primary source of energy (i.e., rather than PP like PAOs), therefore competing with PAOs for VFAs and not contributing to the P removal process. Both PAOs and GAOs rely on their intracellular polymers (i.e., PP, PHAs, and GLY) for metabolic functions and they play vital roles in the EBPR dynamics.

Early models for biological treatment processes, including EBPR, were developed using the traditional population-level modeling (PLM) approach.[1−6] In those models, bacteria populations are quantified using concentration state variables and population heterogeneity in intracellular storage states is ignored. It has been demonstrated that the averaging assumption of the PLM approach introduces an error when the model equations are nonlinear and there is heterogeneity in internal nutrients.[7] Specifically, the population growth rate based on the average nutrient content is higher than the average of the growth rates based on individual nutrient contents (i.e., Jensen's Inequality[8]). This is one of the main motivations for using agent-based modeling (ABM), which differs from PLM in that individual bacteria or groups of bacteria are simulated explicitly.

Received: June 29, 2010
Revised: February 23, 2012
Accepted: February 23, 2012
Published: February 23, 2012

A number of ABMs for wastewater treatment have been developed. In continuously stirred tank reactors (CSTRs), the hydraulic residence time (HRT) varies. That means the population in downstream reactors (or the recycle line) consists of cells with different life histories (i.e., they differ in their past exposure to nutrients). This results in heterogeneity in internal stored nutrients and consequently growth rates.[7,9,10] Comparison of equivalent PLM and ABM models showed that the effect of this heterogeneity is significant. Compared to a PLM, an ABM predicts lower population growth or requires higher maximum rate parameters to produce the same growth.[7,11] In models of SBR systems, where cells have the same HRT or identical histories, no heterogeneity in internal stored nutrients results from the system hydraulics.[7]

Recently, Majed et al. developed a Raman microscopy method to quantify functionally relevant intracellular polymers of single cells in microbial populations in the EBPR process.[12−14] Raman measurements showed for both laboratory- and full-scale EBPR systems that intracellular poly phosphate (PP), polyhydroxyalkanoates (e.g., poly hydroxybutyrates (PHB)), and glycogen (GLY) are indeed heterogeneous. The availability of this method opened the way for a direct validation of ABM predictions of heterogeneity.

In addition to the need for validation, there remains the question about the source of the heterogeneity, which can be due to a number of mechanisms. As discussed above, different histories of individuals in continuous flow reactors can lead to heterogeneity. However, data from SBR systems (presented later in this article) also show heterogeneity in internal stored nutrients, which cannot be attributed to this mechanism. Another potential mechanism is growth asynchrony. Cells grow in between divisions so the cell biomass (gCOD/cell) and absolute quotas (gP/cell) of an unsynchronized population are expected to be heterogeneous. However, the Raman data also show heterogeneity in biomass-based nutrients (gP/gCOD). So there has to be another mechanism. It is recognized that there is significant phylogenetic diversity of PAOs in full-scale EBPR processes, which indicates inherited phenotype heterogeneity.[15,16] PAOs and GAOs candidate strains have been identified (references in ref 15) and models that resolve functional subtypes (e.g., PAOI, PAOII) have been developed.[6,17] The increased understanding of the microbes and their function may eventually lead to models that resolve individual groups or even strains. Here, we explore heterogeneity at the level of PAO, GAO, and ordinary heterotrophic organisms (OHO) functional types. These models correspond to the state-of-the-practice, and for these simplified models it is still desirable to accurately predict heterogeneity in internal nutrient content, and there are several mechanisms for doing so. This includes introducing biological variability in individual behavior (i.e., cells have slight differences in kinetic rates) and biological variability in individual states (i.e., cell division introduces differences in size and stored nutrients).

The aim of this study is to further our understanding of heterogeneity in intracellular polymer storage states of functionally relevant microbial populations in the EBPR process. We set up a laboratory-scale SBR-EBPR and evaluated the heterogeneity of storage states within the population using Raman microscopy. We then developed an ABM and utilized it to predict the heterogeneity and compared the predictions to the observations. Simulations with different methods of introducing biological variability provided insights into the mechanisms underlying the heterogeneity. Comparison of the ABM model to an equivalent PLM illustrated the effect of accounting for heterogeneity.

2. MATERIALS AND METHODS

2.1. Laboratory Set-up. The laboratory scale SBR-EBPR system was seeded with sludge from a full-scale municipal EBPR wastewater treatment plant in the US, operated in a controlled temperature room at 19−22 °C and characterized by four six-hour cycles per day. Each cycle consisted of 7 min fill followed by 133 min of anaerobic phase, 183 min of aerobic phase, 30 min of settling, and then 7 min of withdrawing. The composition of synthetic wastewater feed was according to Schuler and Jenkins.[18] Phosphorus was added as 36 mg/L sodium phosphate monobasic ($NaH_2PO_4{\cdot}H_2O$) (8 mg-P/L). The organic portion of the feeding consisted of sodium acetate ($CH_3COONa{\cdot}3H_2O$) in the amount of 425 mg/L (200 mg COD/L) and 15 mg/L of casamino acids. Nitrogen was added as 5 mg-N/L of ammonium chloride (NH_4Cl). Anaerobic conditions (DO 0.00−0.06 mgO_2/L) were maintained by sparging nitrogen gas in the system and aerobic conditions (DO 4−6 mgO_2/L) were maintained by sparging air using bubble aerators. The SRT and HRT of the system were maintained at about 13 days and 12 h respectively. The SBR was operated for a duration of at least 3 times SRT (40−45 days) to achieve steady-state before subjecting to populations analysis and Raman analysis. The presence of PAOs and EBPR activity in the SBR was confirmed and quantified with phosphate removal performance evaluation, Neisser[19] (Figure S1 of the Supporting Information) and DAPI PP staining[20] and intracellular polymers quantification.[12−14]

After over 40 days of SBR acclimation, a phosphate release and uptake batch test was performed using the sludge taken from the SBR reactor at the end of the settling stage. The batch testing consisted of 45 min of anaerobic phase with acetate addition (80 mg-COD/L) followed by 195 min of aerobic phase. Samples were taken at 15−90 min intervals and were subjected to Raman analysis for quantifying cellular-level intracellular polymers content including polyphosphate, PHB, and glycogen (next section for details). In addition, bulk chemical analyses were performed for soluble orthophosphate, total phosphate, and acetate. Note that validation of the Raman analysis through comparison and correlation with conventional bulk measurements were demonstrated in our previous studies.[12−14] The filtered samples through 0.45 μm filter papers were analyzed for Orthophosphate (orthoP, $(PO_4^{3})^{-}$) and acetate (CH_3COO^{-}) using DX-120 ion chromatograph (Dionex Benelux, Belgium). All phosphorus fractions were measured according to the standard method (4500-P).[21]

2.2. Raman Microscopy Analysis of Intracellular Polymers. Raman spectra were acquired using a WITec, Inc. (Ulm, Germany) Model CRM 2000 confocal Raman microscope. Excitation (ca. 30 mW at 633 nm) was provided by a Helium/Neon laser (Melles Griot, Carlsbad, CA). Spectra were collected at a 0.33 μm grid, with a dwell time of 1 s. Relative quantity of PP content in each individual cell was evaluated based on the Raman intensity (i.e., peak height in the unit of Charged Coupled Device (CCD) counts) of the PO_2^{-} stretching band occurring around 1168−1175 cm^{-1} wavenumber region after background correction.[12,13] The C═O stretching band of ester linkage occurring around 1734 cm^{-1} and glycogen vibration occurring around 480 cm^{-1} were used for quantification of PHB and glycogen content, respectively.[14] Amide I band occurring around 1665 cm^{-1} was used for

biomass quantification as an indicator of cellular protein content.[22] Since there was no commercially available polyhydroxyvalerate (PHV), we used PHB/PHV copolymer for observing signature spectrum for the presence of PHV. We found that PHV presence caused the increased intensity of the peak at 840 cm^{-1} and led to a higher ratio (>1) of the intensity of ν(CC) skeletal stretches at 840 cm^{-1} to the vibrations at 1734 cm^{-1} ($intensity_{860\ cm^{-1}}/intensity_{1734\ cm^{-1}}$), which is approximately ~1 for the spectrum for PHB only. The Raman method can detect relative inclusions of PHB and PHV for cells that do not contain glycogen, but for cells with glycogen it is difficult to differentiate PHB from PHB/PHV coexistence due to interference of one of the peaks pertaining to the glycogen molecule.[14] It has been shown that the relative PHB/PHV content depends on the substrate type. When fed with acetate, PAOs and GAOs produce mainly PHB. Previous studies indicated that PHV in PAOs accounts generally less than 10% of the total PHA and in GAOs, it can be approximately 25%.[23−25] Therefore, for our acetate-fed EBPR system, PHB is expected to be the dominant type of PHAs, which is consistent with the Raman analysis. The overlap of PHV and glycogen does not affect glycogen content detection because the corresponding peak is not used for glycogen quantification. More details on the description of the Raman Microscopy analysis can be found in Majed et al.[12] and Majed and Gu.[14]

3. MODEL DESCRIPTION

3.1. Biochemical Model. The conceptual biochemical model, shown in Figure 1, is based on the Activated Sludge

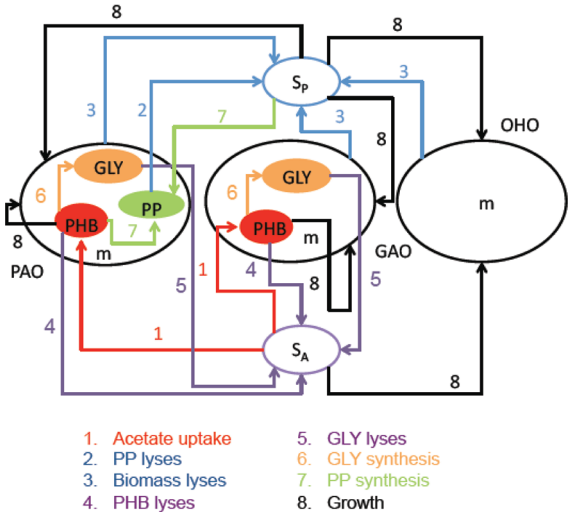


Figure 1. Schematic of the conceptual model. The model includes the three types of cells (PAOs, GAOs, OHOs) and their interactions with extracellular nutrients (S_A and S_P).

Model number 2 (ASM2)[2] with the addition of GAOs[3−5] and GLY for PAOs.[9] An overview is presented below. More details, including stoichiometry table (Table S2 of the Supporting Information) and parameter table (Table S3 of the Supporting Information) are presented in the Supporting Information. The model is written in *MS Excel using Visual Basic for Applications* (*VBA*) and the source code is available from the corresponding author. The model simulates PAOs, GAOs and OHOs. Multiple internal nutrient quotas are simulated, including PP, PHB, and GLY for PAOs, PHB and GLY for GAOs, and none for OHOs. Only PHB is included because PHBs are expected to be the predominant type of PHAs and PHVs cannot be accurately measured by the Raman analysis yet as discussed above. The model therefore simulates PHB rather than PHAs. This approach was also used by Schuler.[9] The remaining part of the cell is referred to as biomass (m). Extracellular nutrients include acetate (S_A) and phosphorus (S_P). Growth of PAOs and GAOs is modeled based on internal (i.e., Droop kinetics) and external (i.e., Monód kinetics) nutrients, while the growth of OHOs is modeled as a function of external nutrient concentration only.

3.2. Agent-based Approach. The model simulates individual bacteria as agents and extracellular nutrients as concentration variables. To illustrate the approach, we present the equations for (a−b) PAOs cell growth, (c) the mass balance for PP in a PAO cell, and (d) the mass balance for extracellular phosphorus (S_P). The system of equations describing the dynamics of PAOs and OHOs is based on Schuler[9] and the dynamics of GAOs are based on Whang et al.[5] The model is implemented using the *iAlgae* framework.[26]

PAOs grow under aerobic conditions using the PHB accumulated during the anaerobic and aerobic phases (acetate uptake)[5,27] and the extracellular phosphorus S_P as substrate. For consistency with previously developed agent-based models,[9,10] we use a direct-biomass growth process[27] rather than considering growth as a result of the difference between storage and utilization of intracellular compounds.[28−30] Growth is therefore simulated with a Droop cell quota approach for PHB[9,10] modified using a Monod saturation term for S_P.[5] The growth formulation for a PAO cell is shown in eq a and that for the cell size (m) is shown in eq b.

$$\mu = \mu_{max}\left(\frac{PHB - PHB_0}{K_{PHB} + PHB - PHB_0}\right)\left(\frac{S_P}{S_P + K_P}\right)m \quad (a)$$

$$\frac{dm}{dt} = (\alpha\mu - R_m) \quad (b)$$

μ is the growth rate (gCOD $cell^{-1}$ day^{-1}), μ_{MAX} is the maximum growth rate (day^{-1}), K_{PHB} (gCOD $gCOD^{-1}$) is the half-saturation constant for PHB, K_P (gP m^{-3}) is the half-saturation constant for S_P, R_m is the biomass excretion rate (gCOD $cell^{-1}$ day^{-1}) computed as $R_m = b_m m$, and α is a switch function, which assumes the value of 0 in anaerobic condition and 1 in aerobic conditions. For cell division, the model uses the "deterministic cell size division approach".[26] That is, the cell divides when $m \geq 2m_0$, where m_0 (gCOD $cell^{-1}$) is the cell size after division.

In presence of volatile fatty acids such as acetate, PP is hydrolyzed and released by PAOs to produce energy for PHB storage (i.e., PP cleavage). This process occurs at a lower rate under aerobic conditions due to the lower acetate concentration. Once the system switches to the aerobic phase, PP is synthesized at a rate-dependent on the extracellular SP, the internal PHB and the maximum allowable PP. Maintenance energy-related P release by PAOs occurs during both anaerobic and aerobic conditions.[31−34] The process is termed "lysis" by Schuler and Jassby,[11] and indicates the endogenous degradation of storage products, which produces maintenance energy and

dx.doi.org/10.1021/es204052p | *Environ. Sci. Technol.* 2012, 46, 3244−3252

secondary nutrient sources. The mass balance for PP in a PAO cell is shown in eq c.

$$\frac{dPP}{dt} = -Y_{PREL}\frac{V_A}{m} + \alpha\frac{V_{PP}}{m} - \frac{R_{PP}}{m} - \alpha\frac{\mu}{m}PP + \frac{R_m}{m}PP \quad (c)$$

Y_{PREL} (gP $gCOD^{-1}$) is the stochiometric parameter for phosphorus release during acetate uptake, V_A (gCOD $cell^{-1}$ day^{-1}) is the acetate uptake rate, V_{PP} (gP $cell^{-1}$ day^{-1}) is the PP synthesis rate, R_{PP} (gP $cell^{-1}$ day^{-1}) is the PP lysis rate. The last two terms in eq c account for the change in biomass-based concentration (internal stored compounds are quantified on a biomass basis) due to the change in biomass (i.e., growth-dilution).

S_P (gP m^{-3}) increases due to PP release during acetate uptake and decreases in the aerobic phase due to PP synthesis. The mass balance for S_P is shown in eq d.

$$\frac{dS_P}{dt} = \frac{1}{V}[(Q_{in}S_{Pin} - Q_{out}S_P) + \sum_{PAOs} Y_{PREL}V_AS_R - \sum_{PAOs} \alpha V_{PP}S_R + \sum_{PAOs} R_{PP}S_R - \sum_{PAOs,\ GAOs,\ OHOs} Y_{PBIO}(\alpha\mu - R_m)S_R] \quad (d)$$

V (m^3) is the reactor volume, Q_{in} and Q_{out} are the inflow and outflow (m^3 day^{-1}) respectively, S_{Pin} (gP m^{-3}) is the inflow extracellular phosphorus concentration, Y_{PBIO} (gP $gCOD^{-1}$) is the phosphorus requirement for biomass growth and S_R is the superindividual number (below). Note that, in contrast to the traditional PLM approach, where cells are represented by a continuous concentration variable, the ABM approach accounts for discrete individuals and their effect on the extracellular concentration is based on the sum across the population.

Explicit simulation of each individual cell would require about 0.5 petabytes of RAM and 150 000 years computer time (2.8 GHz processor, for one simulation) with the present model, which is not feasible. The model therefore uses superindividuals that behave like individual cells but are representative of a number of cells (S_R).[35,8] This is common practice in microbe ABMs and does not constitute a problem as long as the number of superindividuals is sufficiently large to resolve the intrapopulation variability in cell properties.[36] This is the case here, as was verified by increasing the number of superindividuals until the simulation results are no longer sensitive to it.

3.3. Biological Variability. Two approaches for introducing biological variability (i.e., differences between individual cells) are evaluated. The first approach mimics phenotypic differentiation.[37,38] Variability in individual behavior (e.g., nutrient uptake rate) is included at division by randomization of model parameters (e.g., maximum uptake rate) by drawing values from a normal distribution with specified mean and coefficient of variation (CV), truncated at ±2 standard deviations and ≥0. Values are drawn from a global distribution independent of the parameter value of the mother cell. This is done to prevent the evolution of cells with cumulative changes, which would not be meaningful unless properly constrained (i.e., "Darwinian Demon" problem[39]). This approach was originally proposed by Kreft et al.[40] This method causes heterogeneity in individual behavior, which may eventually lead to heterogeneity in individual states. The second approach mimics uneven stochastic partitioning of internal quotas among daughter cells. Variability in individual states (e.g., internal nutrient) is included at division by randomization of the state variables (i.e., biomass and internal storage quotas) by drawing the split fractions from normal distributions with a mean of 0.5 and specified CV truncated at ±2 standard deviations and ≥0. This method requires less computer memory because parameter values do not have to be stored for each agent, although this was not a limitation in the present study. This method directly causes heterogeneity in individual states. Note that both methods lead to stochastic differences between individual cells. However, there are many bacteria (or agents) and these differences are averaged out leading to stable and reproducible population-level behavior (e.g., bulk nutrient uptake rate).

3.4. Relating Raman Spectroscopy Intensities and Model State Variables. Conversion factors between Raman intensities, biomass and polymeric inclusions for bacteria in wastewater treatment applications are not available in the literature, so we developed them based on our data as follows. First, the PP/Amide I ratio of Raman intensities is determined for each cell. Then the average over all cells and times is computed. The same is done for the model output for the times corresponding to the data. The ratio of these quantities (model/data) is used as a global conversion factor. Conversion factors for PHB, m, and GLY were developed in a similar manner. As a result of this, the absolute magnitudes of the model predictions are not independent of the data. However, the purpose of this study is to explore patterns of heterogeneity, and, because a global conversion factor is used, this constitutes an independent and valid model-data comparison. The resulting conversion factors for cell size (m) (pgCOD/Cell) and the Amide I band (CCD) is 0.024 (pgCOD/Cell)/CCD, for PP (gP/gCOD) and PP/Amide I ratio (CCD/CCD) it is 0.015 (gP/gCOD)/(CCD/CCD), for PHB (gCOD/gCOD) and PHB/Amide I ratio (CCD/CCD) it is 0.23 (gCOD/gCOD)/(CCD/CCD), and for GLY (gCOD/gCOD) and GLY/Amide I ratio (CCD/CCD) it is 0.50.

3.5. Quantifying Model Performance. We evaluate the model's ability to predict heterogeneity by comparing the CV of the predicted and observed distributions. This is an appropriate metric, because it does not include the absolute magnitude, which is not independent (section 3.4). For example, the observed PP distribution in part E of Figure 3 has a CV of 1.27. The modeled CV for the simulation with parameters randomization is 1.15. The relative error is therefore 9.5%. The overall error is the relative root-mean-square error (RMSE) of these values for all times for all state variables.

3.6. Model Set-Up and Calibration. In order to compare the model predictions with laboratory data from the phosphorus uptake-release test (above), we first simulate a start-up period of 40 days of SBR cycles (4 cycles per day) used to dissipate the initial conditions, and then we perform the batch test as done for the laboratory-scale system (see section 2.1).

Most parameter values are generally taken from the literature (Supporting Information), with some exceptions that were adjusted to fit our data (Supporting Information for model parameters and literature ranges, Table S3 of the Supporting Information).

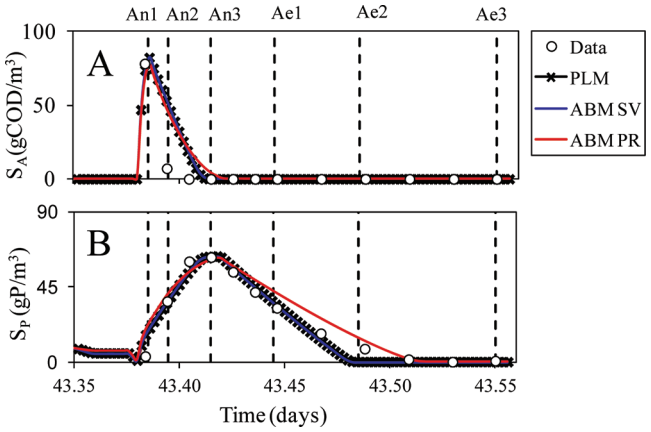


Figure 2. Model-data comparison for extracellular nutrients. (a) Extracellular acetate concentration (gCOD m^{-3}) vs time. (b) Extracellular phosphorus concentration (gP m^{-3}) vs time. Symbols are data. Lines are output from three models: red (ABM PR) is agent-based model with parameters randomization, blue (ABM SV) is agent-based model with state variables randomization, and black (PLM) is for the equivalent population-level model. Black vertical lines indicate the time points at which the Raman intensities were measured (e.g., An2 corresponds to parts b, h, n, and t of Figure 3).

In addition to the population mean of the model parameters, the coefficient of variation (CV) has to be specified for the simulation with parameters randomization. Our first approximation was to use the same CV for all parameters. Then the CV values were refined for each parameter using a genetic algorithm (no recombination) as an optimization tool.[41,42] The algorithm is initialized with the same CV of 0.20 for all parameters. It randomly orders the list of parameters. Then it randomly adjusts the CV of one parameter by ±0.05 and performs the simulation. The performance of the model is quantified by calculating the root-mean-square error (RMSE) between model predictions and data. If the performance of the model is improved (smaller RMSE), the change in CV is retained. Otherwise it is discarded. Then the algorithm moves on to the next parameter. Once the model has gone through all parameters, it starts a new iteration by reshuffling the list. This is done to minimize biases due to parameters ordering. The algorithm stops when we have convergence at a minimum RMSE (Figure S3 of the Supporting Information).

3.7. Comparison with Equivalent Population-Level Model. Predictions of the ABM are compared to an equivalent population-level model, as was done in previous studies.[7,9] That model is based on the same processes as the ABM. However, because the model produced significantly different results, some

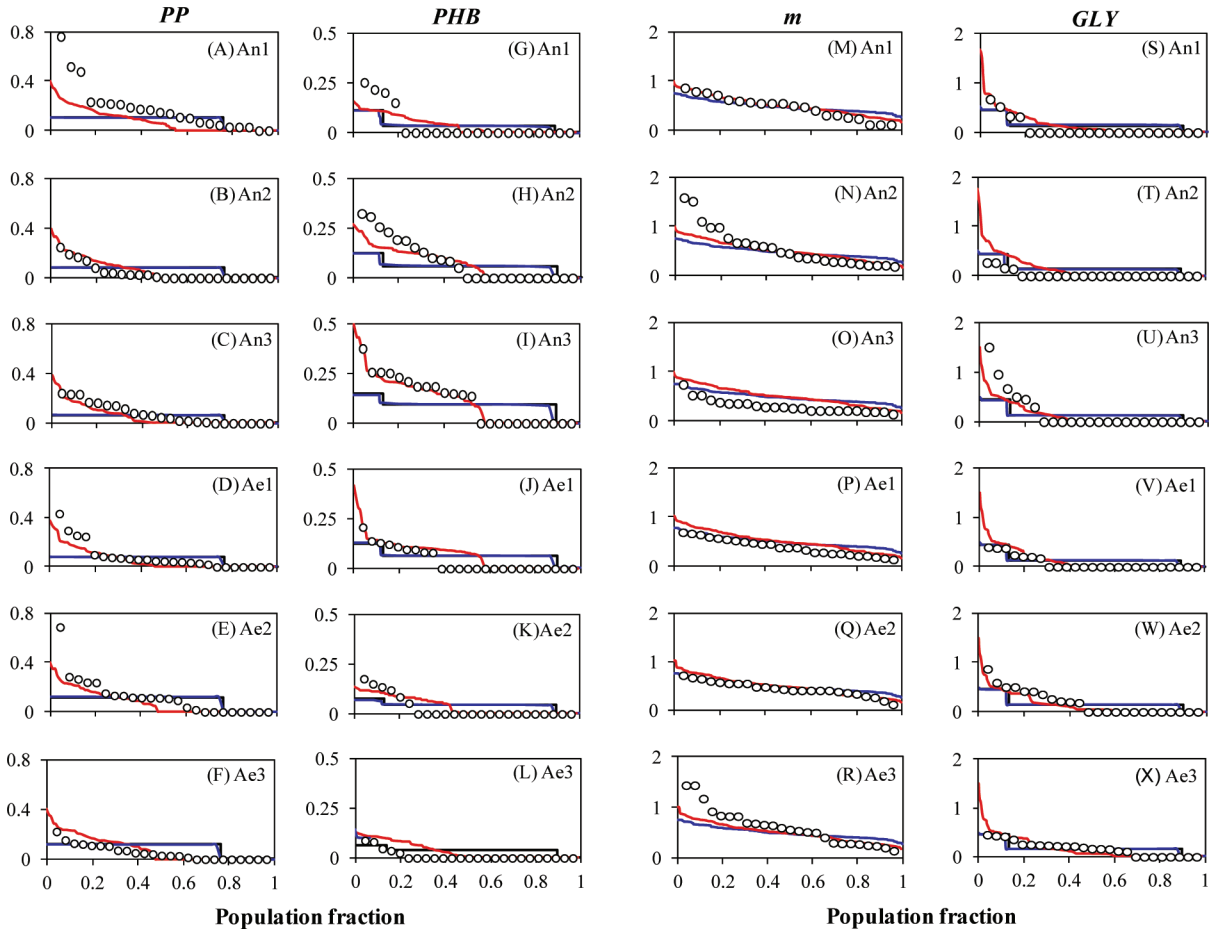


Figure 3. Model-data comparison for intracellular compounds distributions. Distributions for: (a)−(f) internal PP (gP/gCOD), (g)−(l) internal PHB (gCOD/gCOD), cell size *m* (m)−(r)(pgCOD/cell), (s)−(x) internal GLY (gCOD/gCOD). Symbols are data. Lines are output from three different models: red is agent-based model with parameters randomization, blue is agent-based model with state variables randomization, and black equivalent population-level model (also caption of Figure 2).

dx.doi.org/10.1021/es204052p | *Environ. Sci. Technol.* 2012, 46, 3244−3252

parameters were adjusted (Table S3 of the Supporting Information).

4. RESULTS AND DISCUSSION

4.1. Observations. The time series of observed extracellular nutrient concentrations is presented in Figure 2 (symbols, lines are model predictions discussed in the next section). The system shows the characteristic behavior of an EBPR-SBR. Acetate is taken up by PAOs and GAOs in the anaerobic phase (part A of Figure 2). The extracellular phosphorus concentration increases in the anaerobic phase due to PAOs PP release and decreases in the aerobic phase due to cellular growth and PAOs PP synthesis (part B of Figure 2).

Distributions of cell state variables for the times indicated in Figure 2 are presented in Figure 3. The data show a decrease in PP content and fraction of the population with measurable PP during the anaerobic phase (PP release during acetate uptake) and an increase during the aerobic phase (PP synthesis) (parts A−F of Figure 3). The PP decrease during the anaerobic phase and increase during the aerobic phase is a central hallmark of EBPR. PHB distributions shows a pattern opposite to that observed for PP (parts G−L of Figure 3). There is an increase during the anaerobic phase (accumulation due to strong acetate uptake) and a decrease during the aerobic phase (growth and PP/GLY synthesis). Cell size/biomass data are distributed by a factor of 2, which is consistent with a cell size doubling during the cell cycle. This observation is consistent with our choice of using Amide I as a measure of the cell size. GLY data show an oscillatory behavior (parts S−X of Figure 3), which is not consistent with the traditional knowledge of anaerobic phase decrease followed by aerobic phase increase (next section). This behavior was not observed in our past studies[14] and the reasons for it is unclear, possibly a result of experimental variability.

4.2. Model-Data Comparison. Model results for extracellular nutrients and intracellular distributions are presented in Figures 2 and 3. Three lines are presented corresponding to the different models. The red line is the ABM with biological variability in individual behavior (i.e., parameters randomization), the blue line is the ABM with biological variability in individual states (i.e, state variables randomization), and the black line is the equivalent population-level model (i.e, no heterogeneity). We will discuss first the results from the model with parameters randomization (red line in figures), which produced the best fit for the intracellular distributions (Figure 3). We will compare and discuss the difference between the two ABMs and between the ABM and PLM in the following sections.

For extracellular nutrients, the model captures the main features of the acetate and phosphorus time series (Figure 2). For intracellular distributions, the model captures much of the observed dynamics (Figure 3, red line), including the PP decrease during the anaerobic phase and increase during the aerobic phase. However, the model underpredicts the PP content at the beginning of the anaerobic phase (part A of Figure 3) and slightly overpredicts it at the end of the aerobic phase (part F of Figure 3). For PHB, the model reproduces the increase during the anaerobic phase and decrease during the aerobic phase. However, it overpredicts the fraction of cells with detectable PHB content (compared to Raman analysis) at the beginning of the anaerobic phase (part G of Figure 3) and at the end of the aerobic phase (parts K and L of Figure 3). The reason for the difference between the model and data for PP and PHB at the beginning of the anaerobic and end of aerobic phases is unknown, and it may be due to experimental variability.

The model also captures well the biomass or cell size (m) distribution in the reactor (Figure 3, third column). The biomass generally varies by a factor of 2, consistent with the cell size doubling during the cell cycle (as pointed out above). The GLY distribution predicted by the model is shown in the fourth column of Figure 3. The theoretical knowledge of the glycogen dynamics during this type of test predicts its decrease during the anaerobic phase because glycolysis occurs to generate reducing power for PHB formation followed by its replenishment in the aerobic phase.[15] However, this behavior is not seen in the data where, instead, a high-low oscillatory pattern is observed. The model does not include a mechanism that would produce such a pattern. It predicts a relatively constant distribution with a shape similar to observations, but the magnitude lies in between the higher and lower observations.

Overall, for the intracellular polymer storage states, the model captures the main patterns of the magnitude of PP and PHB but not the cyclic behavior in GLY. Additional applications of the model to different data sets are needed to determine if these discrepancies are due to experimental variability or a systematic problem or limitation in the model. For the heterogeneity, the model captures the main patterns observed in the data. This can be seen qualitatively in the model-data comparison presented in Figure 3. We quantify the model performance by comparing the observed and modeled CVs, which compare with an overall relative RMSE of 35% (PP = 18%, PHB = 45%, m = 24%, GLY = 45%). A similar study focusing on phytoplankton had a relative RMSE of 30% for intracellular P.[43] Therefore, this model does slightly better than the phytoplankton model for PP but PHB and GLY are worse.

4.3. Mechanisms that Cause Heterogeneity. The Raman data are from an SBR system and are presented on a biomass-basis, meaning the heterogeneity cannot be attributed to different HRTs or growth asynchrony (see Introduction). The model was used to explore biological variability in individual behavior (i.e, parameters randomization) and individual states (i.e., state variables randomization) as the mechanism underlying the observed heterogeneity. The parameters randomization method (red line) result in the best fit with the data (relative RMSE = 35%). The optimum value for each parameter's CV was determined with an optimization routine (genetic algorithm), which minimized the RMSE (Figure S3 of the Supporting Information). The highest variability is assigned to the maximum acetate uptake rate ($V_{MAX,A}$) for PAOs, which results in a CV value of 0.60 (expected due to the significant variability in the PP data) as well as to the cell-size at division for all functional groups ($m_{0,PAO}$, $m_{0,GAO}$, $m_{0,OHO}$), which is consistent with the high variability in cell size observed in the Amide 1 estimates. Similarly, high variability is predicted for the PAOs stochiometric coefficient for PHB production and consumption (Y_{PHBS}, Y_{PHBP}) glycogen formation (Y_{PHBGLY}) (CV = 0.5) and phosphorus release (Y_{PREL}) (CV = 0.3). Results for all parameters are presented in Table S3 of the Supporting Information. The allocation of variability by the optimization routine may point to the real mechanism of individual variability. Could it be that the most significant differences within the PAO population are related to the variation in phenotypes related to biochemical pathways and mechanism of uptake?

The second ABM explores biological variability in individual states by randomizing the state variables at division. Even though similar results in the extracellular nutrients concentration are obtained with both models (Figure 2, blue vs red lines), the internal nutrient distributions are better reproduced using parameters randomization (Figure 3, red lines) rather than state variables randomization (Figure 3, blue lines). This is also evident in the RMSE where the simulation with state variables randomization results in a worse fit to the data (relative RMSE = 58%). Specifically, it predicted minimal heterogeneity within the bacteria groups, which is evident as a steplike distribution (e.g., Figure 3).

To understand the reason behind the different results predicted by the two methods, the time-series of PP content of a single cell through the division event is presented in Figure 4.

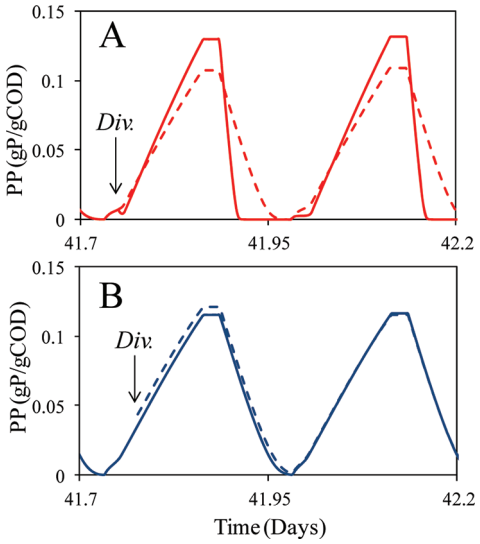


Figure 4. Tracking PP (gP/gCOD) vs time for a mother cell before division (solid line) and two daughter cells after division (solid and dashed lines). Lines are output from two agent-based models: (A) parameters randomization, (B) state variables randomization.

The parameters randomization method (part A of Figure 4) allows for the production and maintenance of heterogeneity. Starting right after the division event, a difference between the two daughter cells develops and is maintained. In contrast, the state variables randomization method (part B of Figure 4) abruptly produces heterogeneity, which is however dissipated in a time shorter than one SBR cycle. The dynamics of the cells are such that any memory introduced by varying internal states is dissipated relatively rapidly. In other words, the concentration of intracellular compounds responds and adjusts rapidly to the environment. The memory in internal states is relatively short and cannot maintain population heterogeneity.

4.4. Effect of Heterogeneity. To explore the importance of accounting for heterogeneity, the ABM with parameters randomization is compared to the equivalent PLM, which ignores heterogeneity (black line in Figures 2 and 3). For extracellular nutrients, the PLM can produce essentially the same predictions as the ABM (Figure 2). However, it requires an adjustment in three parameters. Specifically, the calibrated maximum acetate uptake rate for PAOs of the PLM (3.6 gCOD $gCOD^{-1}$ day^{-1}) is 38% lower than that of the ABM with parameters randomization (5.8 gCOD $gCOD^{-1}$ day^{-1}) while the maximum growth rate and the growth yield for PAOs of the PLM are respectively 55% and 33% lower (2.3 vs 6.0 day^{-1} and 0.6 vs 0.9 gCOD/gCOD). This is consistent with previous ABM/PLM comparison studies,[7,9] where models accounting for heterogeneity predict lower growth or require higher maximum rate/yield parameters, which is due to the nonlinearity of the model equations (i.e., Jensen's Inequality, see Introduction). Interestingly, the parameter values for the PLM are identical to those of the ABM with state variables randomization, which fails to predict the heterogeneity. The PLM does not predict heterogeneity in intracellular states within functional groups (Figure 3). This ABM/PLM comparison shows that accounting for heterogeneity leads to lower growth, and that this can be counteracted with an adjustment of parameters.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Details on: modeled processes, additional figures and analysis, stochiometric table, table with model parameter values, table with system's operational aspects and acronym definition table. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +1(617)373-3992, Fax: +1(617)373-4419, E-mail: ferdi@coe.neu.edu.

Present Addresses

‡Program in Computational Biology, Memorial Sloan-Kettering Cancer Center, New York, 10044, USA.

§Department of Environmental Science and Management, North South University, Bashundhara, Dhaka-1229, Bangladesh.

Author Contributions

⊥These two authors equally contributed to the outcome of the study and should be considered cofirst author.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (NSF Grant 0730239, NSF Grant 932665) and by NU ITRI scholarship sponsored by Anox-Kaldnes Inc., Sweden. The authors thank the advice and assistance in Raman Microscopy provided by Professor Max Diem and Dr. Christian Matthäus in Chemistry and Chemical Biology department at Northeastern University. The authors also thank Dr. Joao Xavier from the Computational Biology Program at Memorial Sloan-Kettering Cancer Center for constructive criticism. Two anonymous reviewers provided valuable feedback.

■ REFERENCES

(1) Henze, M.; Grady, Jr., C. P. L.; Gujer, W.; Marais, G. R.; Matsuo, T. Activated sludge model no. 1. Scientific and Technical Report No. 1, IAWPRC: London, 1987.

(2) Gujer, W.; Henze, M.; Mino, T.; Matsuo, T.; Wentzel, M. C.; Marais, G. v. R. The activated sludge model no. 2: Biological Phosphorus Removal. *Water Sci. Technol.* **1995**, *31* (2), 1−11.

(3) Manga, J.; Ferrer, J.; Garcia-Usach, F.; Seco, A. A modification to the activated sludge model No. 2 based on the competition between phosphorus-accumulating organisms and glycogen-accumulating organisms. *Water Sci. Technol.* **2001**, *43* (11), 161−171.

(4) Yagci, N.; Insel, G.; Artan, N.; Orhon., D. Modelling and calibration of phosphate and glycogen accumulating organism competition for acetate uptake in a sequencing batch reactor. *Water Sci. Technol.* **2004**, *50* (6), 41−50.
(5) Whang, L. M.; Filipe, C. D. M.; Park, J. K. Model-based evaluation of competition between polyphosphate- and glycogen-accumulating organisms. *Water Res.* **2007**, *41* (6), 1312−1324.
(6) Lopez-Vazquez, C. M.; Oehmen, A.; Hooijmans, C. M.; Brdjanovic, D.; Gijzena, H. J.; Yuan, Z.; van Loosdrecht, M. C. M. Modeling the PAO−GAO competition: Effects of carbon source, pH and temperature. *Water Res.* **2009**, *43* (2), 450−462.
(7) Gujer, W. Microscopic versus macroscopic biomass models in activated sludge systems. *Water Sci. Technol.* **2002**, *45* (6), 1−11.
(8) Hellweger, F. L.; Bucci, V. A bunch of tiny individuals − Individual-based modeling for microbes (review paper). *Ecol. Model.* **2009**, *220* (1), 8−22.
(9) Schuler, A. J. Diversity matters: dynamic simulation of distributed bacterial states in suspended growth biological wastewater treatment systems. *Biotechnol. Bioeng.* **2005**, *91* (1), 62−74.
(10) Schuler, A. J. Distributed microbial state effects on competition in enhanced biological phosphorus removal systems. *Water Sci. Technol.* **2006**, *54* (1), 199−207.
(11) Schuler, A. J.; Jassby, D. Distributed State Simulation of Endogenous Processes in Biological Wastewater Treatment. *Biotechnol. Bioeng.* **2007**, *97* (5), 1087−1097.
(12) Majed, N.; Matthaus, C.; Diem, M.; Gu, A. Z. Evaluation of Intracellular Polyphosphate Dynamics in Enhanced Biological Phosphorus Removal Process Using Raman Microscopy. *Environ. Sci. Technol.* **2009**, *43* (11), 5436−5442.
(13) Majed, N.; Matthäus, C.; Diem, M.; Gu, A. Z. Application of Raman Microscopy in the Evaluation of Poly-phosphate Granules in PAOs, *IWA- Specialized conference on microbial population dynamics in biological wastewater treatment ASPD 5*; Aalborg, Denmark, May 24−27, 2009.
(14) Majed, N.; Gu, A. Z. Application of Raman Microscopy for Simultaneous and Quantitative Evaluation of Multiple Intracellular Polymers Dynamics Functionally Relevant to Enhanced Biological Phosphorus Removal Processes. *Environ. Sci. Technol.* **2010**, *44* (22), 8601−8608.
(15) Oehmen, A.; Lemos, P. C.; Carvalho, G.; Yuan, Z.; Keller, J.; Blackall, L. L.; Reis, M. A. M. Advances in enhanced biological phosphorus removal: from micro to macro scale. *Water Res.* **2007**, *41* (11), 2271−2300.
(16) Gu, A. Z.; Saunders, A.; Neethling, J. B.; Stensel, H. D.; Blackall, L. L. Functionally relevant microorganisms to enhanced biological phosphorus removal performance at full-scale wastewater treatment plants in the United States. *Water Envir. Res.* **2008**, *80* (8), 688−698.
(17) Oehmen, A.; Lopez-Vazquez, C. M.; Carvalho, C.; Reis, M. A. M.; van Loosdrecht, M. C. M. Modelling the population dynamics and metabolic diversity of organisms relevant in anaerobic/anoxic/aerobic enhanced biological phosphorus removal processes. *Water Res.* **2010**, *44* (15), 4473−4486.
(18) Schuler, A. J.; Jenkins, D. Enhanced biological phosphorus removal from wastewater by biomass with different phosphorus contents, part I: Experimental results and comparison with metabolic models. *Water Envir. Res.* **2003**, *75* (6), 485−498.
(19) Jenkins, D.; Richards, M. G.; Daigger, G. T. *Manual on the Causes and Control of Activated Sludge Bulking and Foaming*; 2nd ed.; Lewis: London, 1993.
(20) Streichan, M.; Golecki, J. R.; Schon, G. Polyphosphate-accumulating bacteria from sewage plant with different processes of biological phosphorus removal. *FEMS Microbiol. Ecol.* **1990**, *73*, 113−124.
(21) *Standard Methods for the Examination of water and wastewater*, American Public Health Association: Washington, DC, 1998.
(22) Ferraro, J. R.; Nakamoto, K.; Brown, C. W. *Introductory Raman Spectroscopy*. 2nd ed.; Academic Press: San Diego, CA, 2003.
(23) Mino, T.; Van Loosdrecht, M. C. M.; Heijnen, J. J. Microbiology and biochemistry of the enhanced biological phosphate removal process. *Water Res.* **1998**, *32* (11), 3193−3207.
(24) Satoh, H.; Mino, T.; Matsuo, T. Uptake of organic substrates and accumulation of polyhydroxyalkanoates linked with glycolysis of intracellular carbohydrates under anaerobic conditions in the biological excess phosphate removal processes. *Water Sci. Technol.* **1992**, *26*, 933−942.
(25) Filipe, C. D. M.; Daigger, G. T.; Grady, C. P. L. A metabolic model for acetate uptake under anaerobic conditions by glycogen accumulating organisms: Stoichiometry, kinetics, and the effect of pH. *Biotechnol. Bioeng.* **2001**, *76* (1), 17−31.
(26) Hellweger, F. L.; Kianirad, E. Accounting for intra-population variability in biogeochemical models using agent-based methods. *Environ. Sci. Technol.* **2007**, *41* (8), 2855−2860.
(27) Xavier, J. B.; de Kreuk, M. K.; Picioreanu, C.; van Loosdrecht, M. C. M. Multi-scale individual-based model of microbial and bioconversion dynamics in aerobic granular sludge. *Environ. Sci. Technol.* **2007**, *41* (18), 6410−6417.
(28) Murnleitner, E.; Kuba, T.; van Loosdrecht, M. C. M.; Heijnen, J. J. An integrated metabolic model for the aerobic and denitrifying biological phosphorus removal. *Biotechnol. Bioeng.* **1997**, *54* (5), 434−450.
(29) Houweling, D.; Comeau, Y.; Takács, I.; Dold, P. Comment on "Modelling the PAO-GAO competition: Effects of carbon source, pH and temperature" by Lopez-Vazquez, C. M.; Oehmen, A.; Hooijmans, C. M.; Brdjanovic, D.; Gijzen, H. J.; Yuan, Z.; van Loosdrecht, M. C. M. *Water Res.* **2009**, *43* (11), 2947−2949.
(30) Van Loosdrecht, M. C. M.; Oehmen, A.; Hooijmans, C. M.; Brdjanovic, D.; Gijzen, H. J.; Yuan, Z.; Lopez-Vazquez, C. M. Response to the comment on "Modelling the PAO-GAO competition: Effects of carbon source, pH and temperature" by Dwight Houweling et al. *Water Res.* **2009**, *43* (11), 2950−2951.
(31) Ahn, J.; Daidou, T.; Tsuneda, S.; Hirata, A. Transformation of phosphorus and relevant intracellular compounds by a phosphorus-accumulating enrichment culture in the presence of both the electron acceptor and electron donor. *Biotechnol. Bioeng.* **2002**, *79*, 83−93.
(32) Bond, P. L.; Erhart, R.; Wagner, M.; Keller, J.; Blackall, L. L. Identification of some of the major groups of bacteria in efficient and nonefficient biological phosphorus removal activated sludge systems. *Appl. Environ. Microbiol.* **1999**, *65*, 4077−4084.
(33) Kong, Y. H.; Beer, M.; Rees, G. N.; Seviour, R. J. Functional analysis of microbial communities in aerobic-anaerobic sequencing batch reactors fed with different phosphorus/carbon (P/C) ratios. *Microbiology-Sgm* **2002**, *148*, 2299−2307.
(34) Pijuan, M.; Saunders, A. M.; Guisasola, A.; Baeza, J. A.; Casas, C.; Blackall, L. L. Enhanced Biological Phosphorus Removal in a Sequencing Batch Reactor Using Propionate as the Sole Carbon Source. *Biotechnol. Bioeng.* **2004**, *85* (1), 56−67.
(35) Scheffer, M.; Baveco, J. M.; DeAngelis, D. L.; Rose, K. A.; Van Nes, E. H. Super-individuals a simple solution for modeling large populations on an individual basis. *Ecol. Model.* **1995**, *80* (2−3), 161−170.
(36) Hellweger, F. L. Spatially explicit individual-based modeling using a fixed super-individual density. *Comput. Geosci.* **2008**, *34* (2), 144−152.
(37) Perkins, T. J.; Swain, P. S. S. Strategies for cellular decision-making. *Molecular Systems Biology* **2009**, *5* (326), 1−15.
(38) Feinerman, O.; Veiga, J.; Dorfman, J. R.; Germain, R. N.; Altan-Bonnet, G. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. *Science* **2008**, *31* (5892), 1081−1084.
(39) Law, R. Optimal life histories under age-specific predation. *The American Naturalist* **1979**, *114* (3), 399−417.
(40) Kreft, J. U.; Booth, G.; Wimpenny, J. W. T. BacSim, a simulator for individual-based modelling of bacterial colony growth. *Microbiology* **1998**, *144*, 3275−3287.
(41) Forrest, S. Genetic algorithm: principle of natural selection applied to computation. *Science* **1993**, *261* (5123), 872−878.

(42) Forrest, S. Genetic algorithms. *AMC Computing Surveys* **1996**, *26* (1), 77−80.
(43) Bucci, V.; Nunez-Milland, D.; Twining, B.; Hellweger, F. L. Microscale patchiness leads to large and important intraspecific internal nutrient heterogeneity in phytoplankton. *Aquatic Ecology* **2012**, *46* (1), 101−118.