

# Quantitative image analysis of polyhydroxyalkanoates inclusions from microbial mixed cultures under different SBR operation strategies

António L. Amaral<sup>1,2</sup> · Hugo Abreu<sup>1</sup> · Cristiano Leal<sup>2</sup> · Daniela P. Mesquita<sup>2</sup> · Luís M. Castro<sup>1,3</sup> · Eugénio C. Ferreira<sup>2</sup>

Received: 30 August 2016 / Accepted: 27 April 2017 / Published online: 12 May 2017  
© Springer-Verlag Berlin Heidelberg 2017

**Abstract** Polyhydroxyalkanoates (PHAs) produced from mixed microbial cultures (MMC), regarded as potential substitutes of petrochemical plastics, can be found as intracellular granules in various microorganisms under limited nutrient conditions and excess of carbon source. PHA is traditionally quantified by laborious and time-consuming chromatography analysis, and a simpler and faster method to assess PHA contents from MMC, such as quantitative image analysis (QIA), is of great interest. The main purpose of the present work was to upgrade a previously developed QIA methodology (as reported by Mesquita et al. (*Anal Chim Acta* 770:36–44, 2013a, *Anal Chim Acta* 865:8–15, 2015)) for MMC intracellular PHA contents quantification, increase the studied intracellular PHA concentration range, and extend to different sequencing batch reactor (SBR) operation strategies. Therefore, the operation of a new aerobic dynamic feeding (ADF) SBR allowed further extending the studied operating conditions, dataset, and range of the MMC intracellular PHA contents from the previously reported anaerobic/aerobic cycle SBR. Nile Blue A (NBA) staining was employed for epifluorescence microscope visualization and image acquisition, further fed to a custom developed QIA. Data from each of the feast and famine cycles of both SBR were individually processed using

chemometrics analysis, obtaining the correspondent partial least squares (PLS) models. The PHA concentrations determined from PLS models were further plotted against the results obtained in the standard chromatographic method. For both SBR, the predicted ability was higher at the end of the feast stage than for the famine stage. Indeed, an independent feast and famine QIA data treatment was found to be fundamental to obtain the best prediction abilities. Furthermore, a promising overall correlation ( $R^2$  of 0.83) could be found combining the overall QIA data regarding the PHA prediction up to a concentration of  $1785.1 \text{ mg L}^{-1}$  (37.3 wt%). Thus, the results confirm that the presented QIA methodology can be seen as promising for estimating higher intracellular PHA concentrations for a larger reactors operation systems and further extending the prediction range of previous studies.

**Keywords** Sequencing batch reactors (SBR) · Mixed microbial cultures (MMC) · Polyhydroxyalkanoates (PHA) · Nile Blue A (NBA) staining · Quantitative image analysis (QIA) · Partial least squares (PLS)

## Introduction

Regarding wastewater treatment systems (WWT), the biomass usually found consists of mixed microbial cultures (MMC), including organisms able to store nutrient and energy sources intracellularly in granules or inclusions. This fact allows these organisms to use these sources when needed, conferring a competitive advantage over non-storing organisms (van Loosdrecht et al. 1997) in systems such as enhanced biological phosphorous removal (EBPR), among others. As such, under limited nutrient conditions and surplus of carbon sources, MMC are able to take up carbon sources like volatile fatty acids (VFA) and store them intracellularly as

Responsible editor: Gerald Thouand

✉ António L. Amaral  
lpamaral@isec.pt

<sup>1</sup> Instituto Politécnico de Coimbra, ISEC, DEQB, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal

<sup>2</sup> CEB - Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

<sup>3</sup> GERST/ CIEPQPF – Faculty of Sciences and Technology, Universidade de Coimbra, Pólo II, 3030-790 Coimbra, Portugal

polyhydroxyalkanoates (PHAs), whereas, when needed, the stored PHAs are used as carbon and energy sources for biomass growth (Oehmen et al. 2007).

Moreover, PHA production has also an economic drive since they are regarded as potential substitutes of petrochemical-based plastics, furthermore presenting the advantage of being biodegradable. Although, traditionally, PHA production has been performed by pure cultures, MMC are now regarded as an alternative way to produce them, further presenting the advantages of using renewable resources while not requiring sterile conditions (Pandiana et al. 2009; Reis et al. 2003; Carvalho et al. 2014; Duque et al. 2014; Chen et al. 2015a, 2015b; Samori et al. 2015).

Nowadays, PHA is usually quantified by gas chromatography (GC) analysis, with prior digestion, which is both labor-intensive and time-consuming and, therefore, not suitable for fast determinations (Smolders et al. 1994; Oehmen et al. 2005). Thus, an alternative simpler and faster monitoring procedure for PHA quantification is considered of the utmost interest. In this sense, quantitative image analysis (QIA) procedures can provide timely and valuable information for WWT processes, and coupled to microscopy techniques procedures have already been used to detect intracellular storage compounds, such as PHA, in pure cultures (Volova et al. 2013) and in MMC, by staining procedures, as previously reported by Serafim et al. (2002), Pandolfi et al. (2007), and Mesquita et al. (2013b, 2013c, 2015). Regarding PHA inclusions, Sudan Black B (SBB), Nile Blue A (NBA), and Nile Red dyes are commonly used for its selective staining (Ostle and Holt 1982; Redzwan et al. 1997; Alias and Tan 2005), as illustrated by the works of Mesquita et al. (2013a, 2013b, 2015) for PHA quantification by QIA.

Other techniques employed to accurately determine the intracellular PHA contents, in a fast way, encompass the use of Fourier transform infrared (FTIR) spectroscopy, both for pure PHA-producing cultures (Porras et al. 2016) and for MMC cultures (Arcos-Hernandez et al. 2010; Cha et al. 2016; Isak et al. 2016). However, when compared to such methodologies, the QIA methodology has the potential, not only to estimate the PHA contents of a given MMC but also to describe the aggregated biomass structure (by simultaneous bright-field acquisition), which is quite valuable for any WWT system operation (Mesquita et al. 2013c).

Thus, with the purpose of developing an alternative simpler and faster monitoring procedure for PHA quantification, the current work expanded the previous studies of Mesquita et al. (2013a, 2015) of a QIA-based monitoring approach, using NBA to quantify the intracellular PHA produced from MMC, by studying a second SBR cycle strategy. Indeed, with the introduction of an ADF (aerobic feast/famine cycles) SBR in this work, it is intended to go beyond the previous findings,

by upgrading the developed QIA methodology and data treatment techniques, while expanding both the type of studied SBR cycle strategies (including aerobic feast/famine cycles to the existing studies of anaerobic feast/aerobic famine cycles), dataset (67% increase), and range (more than doubled) of PHA concentrations to be predicted. It should be stressed that in both studies, the biomass inoculum was provided by an EBPR system.

Furthermore, chemometric techniques, and namely principal components analysis (PCA), cross-correlation (CC) and partial least square (PLS) regression models, were further used to enlighten the relationships between the obtained QIA parameters, the operating strategies, and the feast/famine stages. Moreover, the estimated intracellular PHA concentrations were plotted against the PHA concentrations determined by the standard analytical method in order to determine the prediction ability of the proposed methodology. Finally, the PHA prediction ability at the end of the feast and famine cycles, for both studied SBR, was also compared.

## Materials and methods

### Experimental setup and synthetic medium

The aerobic feast/famine cycles SBR (system I) was composed by a 3.8 L reactor with a heating jacket and operated at constant 22 °C, in four consecutive periods attaining a 12-h cycle: feeding (15 min), aerobiosis (630 min), settling (60 min), and withdrawal (15 min). Compressed air was used to ensure the aerobic conditions, and the hydraulic retention time (HRT) was set at 22.8 h. Two liters of a synthetic medium containing mainly acetate (solution 1) and a trace metals solution was added in the first 15 min of the cycle. Solution 1 contained (per liter) 4.00 g  $C_2H_3O_2Na \cdot 3H_2O$ , 0.092 g  $K_2HPO_4$ , 0.045 g  $KH_2PO_4$ , 0.16 g  $NH_4Cl$ , 0.60 g  $MgSO_4 \cdot 7H_2O$ , 0.07 g  $CaCl_2 \cdot 2H_2O$ , 0.01 g allyl-N thiourea (inhibiting nitrification), 0.10 g EDTA, and 2 mL (per L) of a trace metals solution, described in detail in Mesquita et al. (2013a, 2015).

Regarding the anaerobic feast/aerobic famine cycles SBR (system II), described in detail in Mesquita et al. (2013a, 2015), consisted of a 4 L SBR operated at around 20 °C, in four consecutive periods attaining a 6 h cycle: anaerobiosis (120 min), aerobiosis (180 min), settling (55 min), and withdrawal (5 min). Two liters of a synthetic medium (Mesquita et al. 2013a, 2015) containing volatile fatty acids (VFA), 3.16 mL (per L) of a trace metals solution (similar to the SBR I) and a phosphate solution was added in the first 5 min of the anaerobic stage. Nitrogen was used to ensure anaerobic conditions and compressed air for aerobic conditions. The hydraulic retention time (HRT) was set at 12 h.

## Analytical procedures

The SBR I system was monitored for a period of 188 days, and mixed liquor samples were collected at approximately 20 min into the aerobiosis stage (corresponding to the feast stage given that the feed was immediately prior to the aerobiosis stage) and at the end of the aerobiosis stage (corresponding to the end of the famine stage). A volume of 15 mL of a 2% (v/v) aqueous solution of formaldehyde was added to 35 mL of the sample and then centrifuged for 15 min at 4000 rpm. Finally, the supernatant was removed and the pellet kept on a freezer at  $-20\text{ }^{\circ}\text{C}$  before further analysis for cell preservation purposes (Wallner et al. 1993). GC analysis allowed determining the polyhydroxyvalerate (PHV) and polyhydroxybutyrate (PHB) contents using the method developed by Smolders et al. (1994). Prior to the digestion, the samples were dehydrated for 48 h at  $60\text{ }^{\circ}\text{C}$  until constant weight. Dehydrated sludge samples were further transferred to glass vials, esterified with HCl/1-propanol (25:75 v/v) and the polymers extracted with dichloromethane with benzoic acid ( $1\text{ mg mL}^{-1}$ ) as the internal standard. This procedure was also performed for blank samples (without the dehydrated sludge). Suitable contact between the two phases was assured by stirring the vials, which were further digested at  $100\text{ }^{\circ}\text{C}$  for 3.5 h. After the digestion, the vial content was transferred by using 10 mL of ultra-pure water to additional vials, covered with a rubber seal, and contact between the two phases was promoted by placing the vials in an inverted position for 30 min. A volume of 1 mL of the organic phase was then collected, and  $\text{Na}_2\text{SO}_4$  was added until a white precipitate was formed, ensuring the complete dehydration of the organic phase.

PHA quantification was then performed through a *Clarus 480* GC system (PerkinElmer, Waltham, USA) outfitted with a flame ionization detector. PHB and PHV were separated using an *Elite Wax* capillary column (PerkinElmer, Waltham, USA), with helium as carrier gas. The split injection and detector temperatures were of 220 and  $250\text{ }^{\circ}\text{C}$ , respectively, whereas the oven temperature profile was of  $50\text{ }^{\circ}\text{C}$  for the first 2 min, with a  $15\text{ }^{\circ}\text{C min}^{-1}$  ramp to  $225\text{ }^{\circ}\text{C}$  (from 2 to 13.67 min) and then maintained until 20.67 min. Calibration curves were also obtained for PHB and PHV with the same internal standard and the *TotalChrom* software (PerkinElmer, Waltham, USA) allowed for the GC data analysis. PHA was then determined as the sum of PHB and PHV.

Regarding the SBR II system, it was monitored for a period of 87 days and mixed liquor samples were taken at the end of the anaerobic feast and aerobic famine stages. The procedure for sampling, fixation, and storage was quite similar to the SBR I, with the exception that the samples were lyophilized instead of dehydrated. The detailed procedure for the SBR II system can be found in Mesquita et al. (2013a, 2015).

## Staining procedure

As for the analytical procedure for PHA quantification, biomass samples were collected at the end of the feast and famine stages also for the NBA staining procedure. Regarding the SBR I system, 15 mL of the samples were centrifuged for 15 min at 4000 rpm and the pellet (after the removal of the supernatant) was re-suspended in 7.5 mL of phosphate buffer saline solution (PBS). This procedure was repeated twice with 2-min centrifugations. Finally, 2.5 mL of PBS solution and 7.5 mL of a 4% (v/v) aqueous solution of formaldehyde were added to the pellet and incubated at  $4\text{ }^{\circ}\text{C}$  for 2 h. The first centrifugation procedure was, once again, employed and the resulting pellet was stored in 7.5 mL of a 50% PBS-50% ethanol (96%) solution at  $-20\text{ }^{\circ}\text{C}$  before further analysis for cell preservation purposes (Wallner et al. 1993). The staining procedure was performed according to Mesquita et al. (2013a, 2015). Regarding the staining procedure for the SBR II system, the procedure was quite similar to the SBR I and can be found in Mesquita et al. (2013a, 2015).

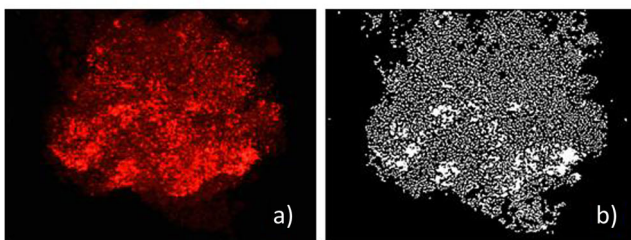
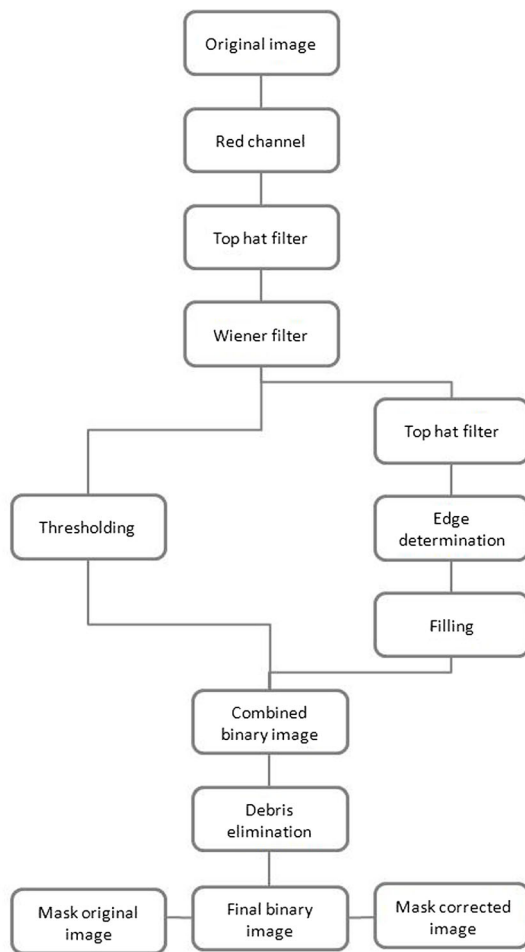
## Image visualization, acquisition, and QIA procedure

For both the SBR I and SBR II systems, the image visualization and acquisition procedure was similar and is described in detail in Mesquita et al. (2013a, 2015).

The QIA methodology focused on the quantification of PHA inclusions regions using a purposely developed program in MATLAB 7.8.0 (The MathWorks, Natick, MA), encompassing four main steps: image preprocessing, background correction, two-way segmentation, and post-processing. A detailed description of the QIA procedure is presented in Mesquita et al. (2013a, 2015), and a schematic representation is shown in Fig. 1 displaying also the original and final binary images. The subsequent morphological parameters determined are shown in Table 1. The overall QIA methodology was also upgraded, and in addition to the use of the raw QIA data, an exponential transformation ( $\text{fExp}_ = x^{3/2}$ ) was employed for the PHA inclusion total area and total intensities in order to correct the acquired images two-dimensional representation of the three-dimensional inclusions. Also, a linear transformation ( $\text{fLin} = x - 0.015$ ) of the PHA inclusions average intensities was now employed to account for the limit threshold value used to determine these inclusions.

## Chemometrics analysis

Cross-correlation (CC) analysis allows determining the correlation factor between each variables pair in a given set of variables. In the present case, the obtained image analysis variables dataset for each SBR (I and II) and each stage (famine and feast) was processed in order to determine the



**Fig. 1** Schematic representation of the QIA procedure for PHA inclusions determination (a original image; b final binary image)

correlation set for all variables. Microsoft Office Excel 2007 (Microsoft, Redmond, WA) was used to that intent.

Principal components analysis (PCA) is a technique that allows the transformation of a given dataset of correlated variables to a new set of uncorrelated (orthogonal) variables, called principal components (PC), in order to reduce the data dimensionality. The principal components (PC1, PC2, ..., PCn) are obtained in order of their importance in explaining the input data, that is, decreasing in their ability to explain the original dataset. In the case of PCY, a given parameter is considered as the output (Y) data, whereas the remaining parameters are considered the input (X) data, allowing to explain (predict) the output parameter. In the present case, the

intracellular PHA concentration was used as the Y parameter, whereas the obtained image analysis variables dataset, for each SBR (I and II), was used as the input (X) data. Simca-P 8.0 (Umetrics AB, Umea, Sweden) was used to perform the PCY analysis.

In PLS regression, the decomposition of the variables (X) data matrix and response (Y) data array is performed iteratively. In each step, the information exchange between the two blocks allows for the rotation of the latent variables (LVs) of the X matrix improving the predictive power regarding the Y array. A more detailed explanation of the PLS algorithm can be found in Einax et al. (1997).

In the present study, the variables (X) data matrix, listed in Table 1, was preprocessed using the standard normal variate method to remove undesirable variations and further used to estimate the response (Y) data array. It is known that, in PLS, the number of LVs is of critical importance and, therefore, the methodology presented in Mesquita et al. (2015) was employed. Accordingly, the data was randomly divided into a training set (67% of the observations) to train the model and a validation set (33% of the observations) to validate the model. Furthermore, this procedure was repeated 20,000 times, obtaining 20,000 different, and randomly selected, datasets to select the optimal number of LVs. The studied number of LVs ranged from one to half the number of X variables in the model in order to prevent over fitting problems. Then, the coefficient of determination ( $R^2$ ) between the observed and model-generated (predicted) Y values was determined for the training, validation, and overall (comprehending both training and validation) datasets. This later was used as the selection criteria for the best model selection.

MATLAB™ 7.8.0 (The MathWorks, Natick, MA) was used to perform the PLS analysis, regarding the PHA prediction, using 122 combined samples from SBR I and SBR II. From these, 49 samples (23 feast (FST) and 26 famine (FMN)) were from the SBR I dataset and the remaining 73 samples (36 FST and 37 FMN) from the SBR II dataset. The obtained QIA results were used in four studies, using different datasets, for each SBR dataset. In all cases, a first PLS model was performed using the merged (further designated as global) feast and famine QIA results, a second PLS model used solely the feast (FST) data, and a third PLS model used solely the famine (FMN) data. Finally, a fourth study was undertaken (further designated as FST + FMN), where the independent FST and FMN data PLS analyses were combined together.

## Results and discussion

In this work, a second set of experiments was added to the previous experiments used in Mesquita et al. (2013a, 2015) extending the range of intracellular storage PHA prediction ability by using QIA data from NBA staining and PLS

**Table 1** Image analysis variables used for the PLS analysis

Parameter	Description
Predictor ( $X$ )	$TI_{ORG}$ – Total intensity of the original image
	$TI_{BKG}$ – Total intensity of the background corrected image
	TA – PHA inclusions total area per image
	fExp (TA) – fExp (PHA inclusions total area per image)
	$TIPHA_{ORG}$ – PHA inclusions total intensity calculated from the original image
	fExp ( $TIPHA_{ORG}$ ) – fExp (PHA inclusions total intensity calculated from the original image)
	$AvIPHA_{ORG}$ – PHA inclusions average intensity calculated from the original image
	fLin ( $AvIPHA_{ORG}$ ) – fLin (PHA inclusions average intensity calculated from the original image)
	$TIPHA_{BKG}$ – PHA inclusions total intensity calculated from the background corrected image
	fExp ( $TIPHA_{BKG}$ ) – fExp (PHA inclusions total intensity calculated from the background corrected image)
Response ( $Y$ )	$AvIPHA_{BKG}$ – PHA inclusions average intensity calculated from the background corrected image
	fLin ( $AvIPHA_{BKG}$ ) – fLin (PHA inclusions average intensity calculated from the background corrected image)
	Polyhydroxyalkanoates concentration (mg PHA L <sup>-1</sup> )

fExp represents the exponential function  $x^{3/2}$ , and fLin represents the linear function  $x - 0.015$

analysis. In fact, as shown in Table 2, while the previous studies allowed predicting PHA concentrations up to 754.4 mg L<sup>-1</sup> (and 35.4 in wt% mg PHA mg<sup>-1</sup> dry weight), the new data expands the predicted range up to 1785.1 mg L<sup>-1</sup> (and 37.3 in wt%). Thus, two datasets, using the QIA information listed in Table 1, were used, respectively, from SBR I regarding the aerobic FST/FMN cycles and from SBR II regarding the (anaerobic FST)/(aerobic FMN) cycles. Furthermore, the intracellular PHA concentration (referred as observed PHA) was also determined by a standard analytical methodology, as previously reported in “Analytical procedures” section. For both SBR, the samples were obtained at the end of the FST and FMN stages.

As in Mesquita et al. (2013a, 2015) studies, the first PLS models employed the global (merged) FST and FMN QIA datasets. The PLS analysis results (coefficients of determination and linear regression equations) for the PHA prediction in each SBR, as well as for the ensemble SBR I + II, regarding the training, validation, and overall (training + validation) datasets are depicted in Table 3. Each coefficient of determination ( $R^2$ ) and regression equation corresponds to the linear

correlation found by the least squares method. Analyzing Table 3, it could be retrieved that the prediction results from SBR II were somewhat better than SBR I, possibly due to the narrower range of PHA concentrations (maximum 754.4 mg L<sup>-1</sup> in comparison with a maximum of 1785.1 mg L<sup>-1</sup> for SBR I). In fact, it could be expected that, for quite larger PHA concentrations, the acquired images could suffer from some bleaching, thus leading to the need of modifying the fluorescence intensity or the image acquisition time. This is in accordance with the results obtained from the CC analysis, which showed a higher correlation, between the PHA concentration and the PHA inclusions total intensity parameters, for SBR II (0.65 and 0.67 correlation factors) than for SBR I (0.16 and 0.22 correlation factors). As expected, analyzing Table 4, it is also evident that the variable importance (VIP) of the PHA inclusions total intensity parameters in SBR I is much smaller than the average intensity parameters, regarding the PHA concentration prediction. Such difference infers that a fluorescence bleaching problem could be affecting the results on the higher PHA concentrations obtained on SBR I.

Furthermore, when the PLS analysis was conducted with the ensemble SBR I + II, the obtained prediction results were slightly better (taking into account the  $R^2$  and slope values) than for any of the individual SBR. However, these results remained somewhat distant from 1 (coefficient of determination— $R^2$  of 0.69 for the training + validation sets) with an average error of 172.7 mg L<sup>-1</sup> (9.7% of the studied range), and thus, the prediction ability could be considered as questionable.

In order to improve the intracellular PHA concentration prediction ability, the FST and FMN cycles were modeled independently by the PLS analysis, and the obtained results were further merged to obtain the combined (FST + FMN)

**Table 2** Minimum and maximum values for the PHA concentration (in mg PHA L<sup>-1</sup> and in wt% mg PHA mg<sup>-1</sup> dry weight) in the feast and famine cycles of SBR I and SBR II

		[PHA] (mg L <sup>-1</sup> )		[PHA] (wt%)	
		Min.	Max.	Min.	Max.
SBR I	Feast	54.8	1676.7	1.2	35.0
	Famine	86.4	1785.1	2.5	37.3
SBR II	Feast	17.6	754.4	2.0	32.2
	Famine	11.6	751.1	0.9	35.4

**Table 3** PHA concentrations prediction results for SBR I, SBR II, and the ensemble SBR I + II using the global (merged) feast and famine QIA data and PLS

	SBR I	SBR II	SBR I + II
$R^2$ training set	0.53	0.73	0.72
Linear regression (training set)	$y = 0.95x$	$y = 0.94x$	$y = 0.94x$
$R^2$ validation set	0.49	0.66	0.66
Linear regression (validation set)	$y = 1.06x$	$y = 1.01x$	$y = 1.04x$
$R^2$ overall (training + validation) set	0.51	0.70	0.69
Linear regression (training + validation set)	$y = 0.99x$	$y = 0.96x$	$y = 0.98x$

PHA concentration predictions. Again, this procedure is in line with the one implemented in Mesquita et al. (2013a, 2015) studies (anaerobic feast and aerobic famine). Accordingly, the PLS analysis results (coefficients of determination and linear regression equations) for the PHA prediction in each SBR, as well as for the ensemble SBR I + II, regarding the training, validation, and overall (training + validation) datasets are depicted in Table 5. Again, each coefficient of determination and regression equation corresponds to the linear correlation found by the least squares method.

Comparing the PHA prediction ability from the FST and FMN stages calculated independently, the PLS results demonstrated that it was possible to predict with greater accuracy the PHA concentration at the end of the FST stage ( $R^2$  of 0.85) with respect to the end of the FMN stage ( $R^2$  of 0.81), possibly due to the different PhaP protein levels in these stages. Furthermore, the results obtained from the CC analysis showed a much higher correlation, at the end of the FST stage, between the PHA concentration and both the PHA inclusions total intensity parameters (0.47 and 0.51 correlation factors for the end of the FST stage and 0.17 and 0.21 for the end of the FMN stage) and the PHA inclusions average intensity parameters (0.73 and 0.74 correlation factors for the end of the FST

stage and 0.64 and 0.67 for the end of the FMN stage). It was also noticed that this effect was slightly more prominent in SBR I than in SBR II, both in terms of the coefficient of determination ( $R^2$  of 0.79 and 0.81 at the end of the FST stage against 0.62 and 0.78 at the end of the FMN stage for SBR I and SBR II, respectively) and on the CC analysis. With respect to the later, a clear distinction could be found in the correlation factors of SBR I, at the end of the FST and FMN stages, between the PHA concentration and both the PHA inclusions total intensity parameters (0.62 and 0.64 correlation factors for the end of the FST stage and 0.20 and 0.19 for the end of the FMN stage) and the PHA inclusions average intensity parameters (0.69 and 0.71 correlation factors for the end of the FST stage and 0.40 and 0.41 for the end of the FMN stage).

A PCY analysis further corroborated these findings, with a slightly clearer distinction between the end of the FMN stage and the end of the FST stage, for the two first principal components (PC) in SBR I, as shown in Fig. 2. In effect, the results of the PCY analysis clearly demonstrate the clear difference in the data collected at the end of each stage, but only a slight difference between SBR I and II, with 62 out of 73 (83.6%) data points in SBR II and 42 out of 49 (85.7%) data points in SBR I, able to be separated.

It is known that PhaP proteins have regulatory functions regarding the granules number and size, and its expression can be induced by PHA accumulation (York et al. 2001). It could be expected that, in the FST stage, higher intracellular PHA concentrations result in larger PhaP concentrations and large numbers of smaller PHA granules, whereas in the FMN stage, lower intracellular PHA concentrations result in smaller PhaP levels and bigger PHA granules. Further enlightenment of whether or not this is the cause for the observed differences should also be sought on future studies.

These results highlight the importance of performing independent analyses on the data collected from the FST and FMN stages, corroborating the findings of the previous (Mesquita et al. 2013a, 2015) studies. Furthermore, the results for the overall (training + validation) coefficient of determination ( $R^2$ ) for all cases (SBR I, SBR II and SBR I + II) improved markedly regarding the prior global analysis (from 0.51 to 0.72, from 0.70 to 0.80, and from 0.69 to 0.83, respectively). It should also be stressed that the previous works allowed

**Table 4** Variable importance (VIP) for the image analysis variables for SBR I and SBR II PHA concentration prediction using the global (merged) feast and famine QIA data and PLS

	VIP	
	SBR I	SBR II
TI <sub>ORG</sub>	0.98	0.98
TI <sub>BKG</sub>	0.97	0.92
TA	0.88	0.89
fExp (TA)	0.92	0.90
TIPHA <sub>ORG</sub>	0.74	1.09
fExp (TIPHA <sub>ORG</sub> )	0.77	1.03
AvIPHA <sub>ORG</sub>	1.21	1.01
fLin (AvIPHA <sub>ORG</sub> )	1.21	1.01
TIPHA <sub>BKG</sub>	0.85	1.07
fExp (TIPHA <sub>BKG</sub> )	0.86	1.09
AvIPHA <sub>BKG</sub>	1.22	0.99
fLin (AvIPHA <sub>BKG</sub> )	1.22	0.99

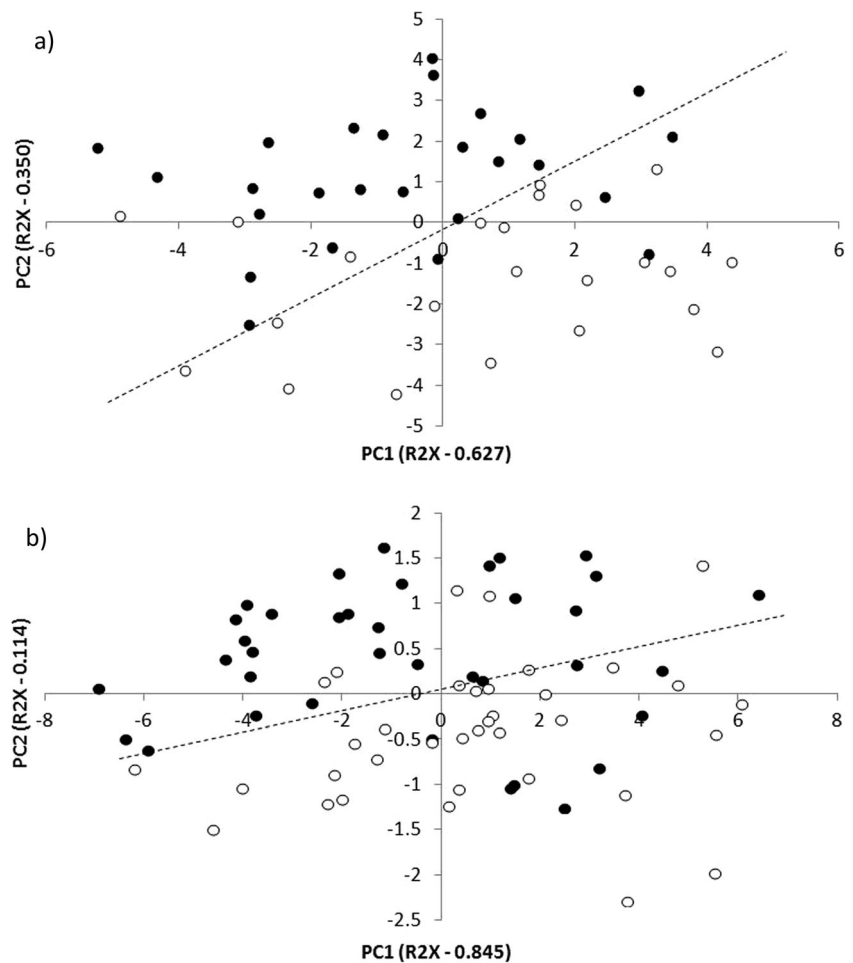
**Table 5** PHA concentrations prediction results for SBR I, SBR II, and the ensemble SBR I + II using independent feast (FST) and famine (FMN) QIA data and PLS

	SBR I		SBR II		SBR I + II	
	FST	FMN	FST	FMN	FST	FMN
$R^2$ training set	0.80	0.73	0.79	0.80	0.86	0.87
Linear regression (training set)	$y = 0.95x$	$y = 0.99x$	$y = 0.94x$	$y = 0.97x$	$y = 0.95x$	$y = 0.99x$
$R^2$ validation set	0.86	0.46	0.84	0.75	0.87	0.71
Linear regression (validation set)	$y = 1.21x$	$y = 1.12x$	$y = 1.03x$	$y = 0.98x$	$y = 1.15x$	$y = 1.06x$
$R^2$ overall (training + validation) set	0.79	0.62	0.81	0.78	0.85	0.81
Linear regression (training + validation set)	$y = 1.02x$	$y = 1.03x$	$y = 0.97x$	$y = 0.97x$	$y = 1.01x$	$y = 1.01x$
(FST + FMN) $R^2$ (training + validation set)	0.72		0.80		0.83	
Linear regression (FST + FMN)	$y = 1.02x$		$y = 0.97x$		$y = 1.01x$	

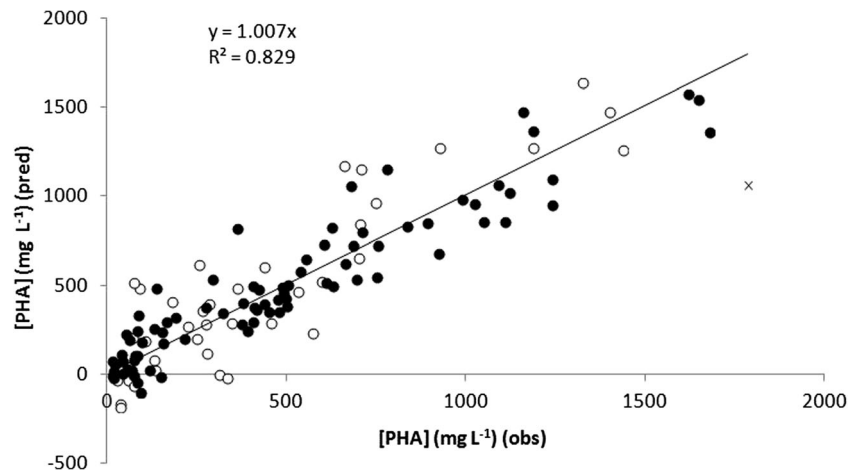
inferring that independent aerobic and anaerobic QIA data treatment should be undertaken for greater accuracy regarding the PHA concentration prediction. However, it was not clear if the aerobic/anaerobic conditions were responsible for the found differences or were rather the feast/famine conditions. With this work, it could be concluded that, even with aerobic

cycles (ADF), the differences remain when comparing the end of the feast and famine stages. As such, the independent feast and famine QIA data treatment, rather than a plain aerobic and anaerobic QIA data treatment, was found to be fundamental to obtain the best prediction abilities regarding the intracellular PHA quantification.

**Fig. 2** Plot of the first two components (PC1 and PC2) of the PCY analysis regarding **a** SBR I and **b** SBR II. Filled circles (●) correspond to the end of the famine stage and open circles (○) to the end of the feast stage



**Fig. 3** Correlations between the observed and predicted intracellular PHA concentrations combining the two independent analyses (FST + FMN) for both SBR systems. Filled circles (•) correspond to training data, open circles (◦) to validation data, and crosses (×) to discarded data



Again, better results were obtained for SBR II in comparison with SBR I, due to the reasons previously stated. Furthermore, when the PLS analysis was conducted with the ensemble SBR I + II, the obtained prediction results, presented in Table 5 and Fig. 3, were better than for any of the individual SBR. Given the obtained results and the fact that combining both SBR results allowed for an increased PHA prediction, it could be inferred that it was possible to assess the intracellular PHA inclusions, by the proposed QIA procedure, for two different SBR operating strategies even taking into consideration the resulting physiological differences, as long as the feast and famine stages are individually addressed. Therefore, it seems licit to conclude that the proposed QIA methodology was successfully proven to be relatively robust regarding different SBR operating strategies. And, although still somewhat far from perfect, the obtained prediction ability (coefficient of determination— $R^2$  of 0.83 for the training + validation sets), with an average error of  $132.9 \text{ mg L}^{-1}$  (7.4% of the studied range), can be seen as promising, within the studied intracellular PHA concentration range, taking into consideration the greater simplicity and speediness of the proposed method with respect to the standard analytical techniques.

### Conclusions

In this work, a QIA methodology was used and upgraded, to quantify MMC intracellular PHA contents of two different SBR cycle strategies. The operation of an ADF SBR allowed further extending the studied SBR operating conditions, dataset, and intracellular PHA range from the previous anaerobic/aerobic cycle SBR studies (Mesquita et al. 2013a, 2015). Besides the standard analytical methodologies to determine the intracellular PHA concentrations, NBA staining and QIA methodologies were also employed. The performed PCA and CC analyses allowed to enlighten the relationships between the obtained

QIA parameters, operating strategies, and feast/famine stages. The QIA data was further processed using PLS analysis to allow for a promising estimation (coefficient of determination— $R^2$  of 0.83 and average error of  $132.9 \text{ mg L}^{-1}$ , i.e., 7.4% of the studied range) of the PHA concentrations in a wider range (up to a concentration of  $1785.1 \text{ mg L}^{-1}$ ) than the previous studies. The obtained results also pointed to a greater accuracy regarding the PHA concentration prediction at the end of the feast stage in comparison to the end of the famine stage. Moreover, the independent feast and famine QIA data treatment was found to be fundamental to obtain the best prediction abilities. The underlying reason for these differences should also be sought on future studies.

**Acknowledgements** This study was supported by the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2013 unit and COMPETE 2020 (POCI-01-0145-FEDER-006684) and BioTecNorte operation (NORTE-01-0145-FEDER-000004) funded by European Regional Development Fund under the scope of Norte2020 - Programa Operacional Regional do Norte. The authors also acknowledge the financial support to Cristiano S. Leal (PTDC/EBB-EBI/103147/2008, FCOMP-01-0124-FEDER-009704) and Daniela P. Mesquita through the FCT postdoctoral grant (SFRH/BPD/82558/2011).

### References

Alias Z, Tan IKP (2005) Isolation of palm oil-utilising, polyhydroxyalkanoate (PHA)-producing bacteria by an enrichment technique. *Bioresour Technol* 96:1229–1234

Arcos-Hernandez MV, Gurieffa N, Pratt S, Magnusson P, Werker A, Vargas A, Lant P (2010) Rapid quantification of intracellular PHA using infrared spectroscopy: an application in mixed cultures. *J Biotechnol* 150:372–379

Carvalho G, Oehmen A, Albuquerque MGE, Reis MAM (2014) The relationship between mixed microbial culture composition and PHA production performance from fermented molasses. *New Biotechnol* 31:257–263



- Cha SH, Son JH, Jamal Y, Zafar M, Park HS (2016) Characterization of polyhydroxyalkanoates extracted from wastewater sludge under different environmental conditions. *Biochem Eng J* 112:1–12
- Chen Z, Huang L, Wen Q, Guo Z (2015a) Efficient polyhydroxyalkanoate (PHA) accumulation by a new continuous feeding mode in three-stage mixed microbial culture (MMC) PHA production process. *J Biotechnol* 209:68–75
- Chen Z, Guo Z, Wen Q, Huang L, Bakke R, Du M (2015b) A new method for polyhydroxyalkanoate (PHA) accumulating bacteria selection under physical selective pressure. *Int J Biol Macromolec* 72: 1329–1334
- Duque AF, Oliveira CSS, Carmo ITD, Gouveia AR, Pardelha F, Ramos AM, Reis MAM (2014) Response of a three-stage process for PHA production by mixed microbial cultures to feedstock shift: impact on polymer composition. *New Biotechnol* 31:276–288
- Einax JW, Zwanziger HW, Geiss S (1997) *Chemometrics in Environmental Analysis*. VCH, Weinheim, Germany.
- Isak I, Patel M, Riddell M, West M, Bowers T, Wijeyekoon S, Lloyd J (2016) Quantification of polyhydroxyalkanoates in mixed and pure cultures biomass by Fourier transform infrared spectroscopy: comparison of different approaches. *Lett Appl Microbiol* 63:139–146
- Mesquita DP, Leal C, Cunha JR, Oehmen A, Amaral AL, Reis MAM, Ferreira EC (2013a) Prediction of intracellular storage polymers using quantitative image analysis in enhanced biological phosphorus removal systems. *Anal Chim Acta* 770:36–44
- Mesquita DP, Selvaggio G, Cunha JR, Leal C, Amaral AL, Ferreira EC (2013b) Image analysis for automatic characterization of polyhydroxyalkanoates granules. *Lecture Notes in Comput Sci* 7950:790–797
- Mesquita DP, Amaral AL, Ferreira EC (2013c) Activated sludge characterization through microscopy: a review on quantitative image analysis and chemometric techniques. *Anal Chim Acta* 802:14–28
- Mesquita DP, Amaral AL, Leal C, Oehmen A, Reis MAM, Ferreira EC (2015) Polyhydroxyalkanoate granules quantification in mixed microbial cultures using image analysis: Sudan Black B versus Nile Blue A staining. *Anal Chim Acta* 865:8–15
- Oehmen AB, Keller-Lehmann RJ, Zeng Z, Yuan ZJ, Keller J (2005) Optimisation of poly-beta-hydroxyalkanoate analysis using gas chromatography for enhanced biological phosphorus removal systems. *J Chromatography A* 1070:131–136
- Oehmen A, Lemos PC, Carvalho G, Yuan ZG, Keller J, Blackall LL, Reis MAM (2007) Advances in enhanced biological phosphorus removal: from micro to macro scale. *Water Res* 41:2271–2300
- Ostle AG, Holt JG (1982) Nile blue a as a fluorescent stain for poly-beta-hydroxybutyrate. *Appl Environ Microb* 44:238–241
- Pandiana SRK, Kalishwaralal DV, Muniyandi J, Rameshkumar N, Gurunathan S (2009) Synthesis of PHB nanoparticles from optimized medium utilizing dairy industrial waste using *Brevibacterium casei* SRKP2: a green chemistry approach. *Colloids Surf B* 74:266–273
- Pandolfi D, Pons MN, da Motta M (2007) Characterization of PHB storage in activated sludge extended filamentous bacteria by automated colour image analysis. *Biotechnol Lett* 29:1263–1269
- Porras MA, Cubitto MA, Villar MA (2016) A new way of quantifying the production of poly(hydroxyalkanoate)s using FTIR. *J Chem Technol Biotechnol* 91:1240–1249
- Redzwan G, Gan SN, Tan IKP (1997) Isolation of polyhydroxyalkanoate-producing bacteria from an integrated-farming pond and palm-oil mill effluent ponds. *World J Microb Biotechnol* 13:707–709
- Reis MAM, Serafim LS, Lemos PC, Ramos AM, Aguiar FR, Van Loosdrecht MCM (2003) Production of polyhydroxyalkanoates by mixed microbial cultures. *Bioprocess Biosyst Eng* 25:377–385
- Samori C, Abbondanzi F, Galletti P, Giorgini L, Mazzocchetti L, Torri C, Tagliavini E (2015) Extraction of polyhydroxyalkanoates from mixed microbial cultures: impact on polymer quality and recovery. *Bioresour Technol* 189:195–202
- Serafim LS, Lemos PC, Levantesi C, Tandoi V, Santos H, Reis MAM (2002) Methods for detection and visualization of intracellular polymers stored by polyphosphate-accumulating microorganisms. *J Microbiol Methods* 51:1–18
- Smolders GJF, van de Meij J, van Loosdrecht MCM, Heijnen JJ (1994) Model of the anaerobic metabolism of the biological phosphorus removal process: stoichiometry and pH influence. *Biotechnol Bioeng* 43:461–470
- van Loosdrecht MCM, Pot MA, Heijnen JJ (1997) Importance of bacterial storage polymers in bioprocesses. *Water Sci Technol* 35:41–47
- Volova TG, Zhila NO, Kalacheva GS, Brigham CJ, Sinskey AJ (2013) Effects of intracellular poly(3-hydroxybutyrate) reserves on physiological-biochemical properties and growth of *Ralstonia eutropha*. *Res Microbiol* 164:164–171
- Wallner G, Amann R, Beisker W (1993) Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* 14:136–143
- York GM, Stubbe J, Sinskey AJ (2001) New insight into the role of the PhaP phasin of *Ralstonia eutropha* in promoting synthesis of polyhydroxybutyrate. *J Bacteriol* 173:2394–2397