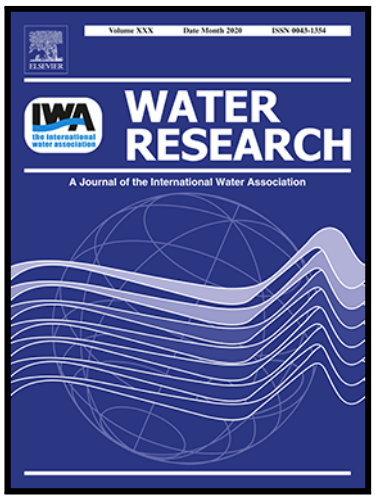


Intracellular Polyphosphate Length Characterization in Polyphosphate Accumulating Microorganisms (PAOs): Implications in PAO Phenotypic Diversity and Enhanced Biological Phosphorus Removal Performance

Dongqi Wang , Yueyun Li , Helen A. Cope , Xiaoxiao Li , Peisheng He , Cong Liu , Guangyu Li , Sheikh M. Rahman , Nicholas B. Tooker , Charles B. Bott , Annalisa Onnis-Hayden , Jyoti Singh , Alistair Elfick , Ricardo Marques , Henning J. Jessen , Adrian Oehmen , April Z. Gu

PII: S0043-1354(21)00920-9
DOI: https://doi.org/10.1016/j.watres.2021.117726
Reference: WR 117726

To appear in: *Water Research*

Received date: 13 June 2021
Revised date: 31 August 2021
Accepted date: 26 September 2021

Please cite this article as: Dongqi Wang , Yueyun Li , Helen A. Cope , Xiaoxiao Li , Peisheng He , Cong Liu , Guangyu Li , Sheikh M. Rahman , Nicholas B. Tooker , Charles B. Bott , Annalisa Onnis-Hayden , Jyoti Singh , Alistair Elfick , Ricardo Marques , Henning J. Jessen , Adrian Oehmen , April Z. Gu , Intracellular Polyphosphate Length Characterization in Polyphosphate Accumulating Microorganisms (PAOs): Implications in PAO Phenotypic Diversity and Enhanced Biological Phosphorus Removal Performance, *Water Research* (2021), doi: https://doi.org/10.1016/j.watres.2021.117726

Intracellular Polyphosphate Length Characterization in Polyphosphate Accumulating Microorganisms (PAOs): Implications in PAO Phenotypic Diversity and Enhanced Biological Phosphorus Removal Performance

Dongqi Wang[a,b,c,1], Yueyun Li[c,d,1], Helen A. Cope[e], Xiaoxiao Li[b], Peisheng He[f], Cong Liu[b], Guangyu Li[c,f], Sheikh M. Rahman[c,g], Nicholas B. Tooker[c,h], Charles B. Bott[i], Annalisa Onnis-Hayden[c], Jyoti Singh[j,k], Alistair Elfick[e], Ricardo Marques[l], Henning J. Jessen[j], Adrian Oehmen[l,m] and April Z. Gu[c,f,*,1] aprilgu@cornell.edu

[a]State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an, Shaanxi, 710048, China

[b]Department of Municipal and Environmental Engineering, Xi'an University of Technology, Xi'an, Shaanxi, 710048, China

[c]Department of Civil and Environmental Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA, 02115, United States

[d]Black and Veatch, 2999 Oak Road #490, Walnut Creek, CA, 94597, United States

[e]School of Engineering, Institute for Bioengineering, The University of Edinburgh, Edinburgh, United Kingdom

[f]School of Civil and Environmental Engineering, Cornell University, 220 Hollister Hall, Ithaca, NY, 14853, United States

[g]Department of Civil Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh

[h]Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Marston Hall, Amherst, MA, 01003, United States

[i]Hampton Roads Sanitation District, 1434 Air Rail Avenue, Virginia Beach, VA, 23454, United States

[j]Institute of Organic Chemistry, University of Freiburg, Albertstrasse 21, 79104 Freiburg, Germany

[k]Department of Chemistry, University College London, 20 Gordon St, Bloomsbury, London WC1H 0AJ, United Kingdom

[l]UCIBIO, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal

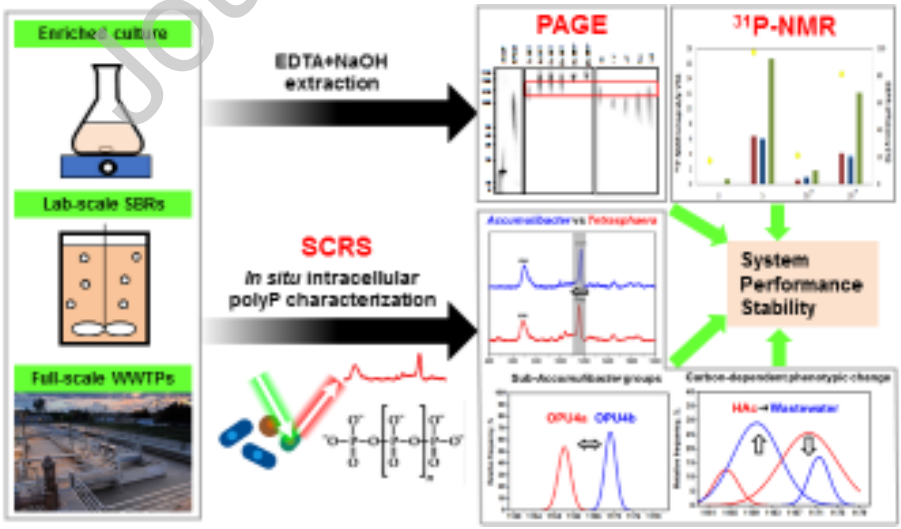
[m]School of Chemical Engineering, The University of Queensland, Brisbane, Queensland 4072, Australia

[*]Corresponding Author. School of Civil and Environmental Engineering, Cornell University, 220 Hollister Hall, Ithaca, NY, 14853, United States.
[1]These authors contributed equally to this study.

Highlight:

- Distinct polyP characteristics were firstly revealed among different EBPR systems
- Linkage between specific PAO phenotype with EBPR performance/stability are observed
- SCRS allows differentiation of PAO phenotypes at the microdiversity level
- SCRS captures phenotypic changes of PAOs in response to system conditions
- SCRS enables *in situ* intracellular polyP analysis compared to other methods

Graphical Abstract



Abstract

Polyphosphate (polyP) accumulating organisms (PAOs) are the key agent to perform enhanced biological phosphorus removal (EBPR) activity, and intracellular polyP plays a key role in this process. Potential associations between EBPR performance and the polyP structure have been suggested, but are yet to be extensively investigated, mainly due to the lack of established methods for polyP characterization in the EBPR system. In this study, we explored and demonstrated that single-cell Raman spectroscopy (SCRS) can be employed for characterizing intracellular polyPs of PAOs in complex environmental samples such as EBPR systems. The results, for the first time, revealed distinct distribution patterns of polyP length (as Raman peak position) in PAOs in lab-scale EBPR reactors that were dominated with different PAO types, as well as among different full-scale EBPR systems with varying configurations. Furthermore, SCRS revealed distinctive polyP composition/features among PAO phenotypic sub-groups, which are likely associated with phylogenetic and/or phenotypic diversity in EBPR communities, highlighting the possible resolving power of SCRS at the microdiversity level. To validate the observed polyP length variations via SCRS, we also performed and compared bulk polyP length characteristics in EBPR biomass using conventional polyacrylamide gel electrophoresis (PAGE) and solution ^{31}P nuclear magnetic resonance (^{31}P-NMR) methods. The results are consistent with the SCRS findings and confirmed the variations in the polyP lengths among different EBPR systems. Compared to conventional methods, SCRS exhibited advantages as compared to conventional methods, including the ability to characterize *in situ* the intracellular polyPs at subcellular resolution in a label-free and non-destructive way, and the capability to capture subtle and detailed biochemical fingerprints of cells for phenotypic classification. SCRS also has recognized limitations in comparison with ^{31}P-NMR and PAGE, such as the inability to quantitatively detect the average polyP chain length and its distribution. The results provided initial evidence for the potential of SCRS-enabled polyP characterization as an alternative and complementary microbial community

phenotyping method to facilitate the phenotype-function (performance) relationship deduction in EBPR systems.

Keywords

EBPR; polyphosphate chain length; NMR; polyacrylamide gel electrophoresis; single-cell Raman spectroscopy

1. Introduction

The discharge of phosphorus (P) into water bodies has been identified as one of the major factors causing eutrophication (Selman and Greenhalgh 2009). P, as one of the most critical elements for food production, is still almost entirely obtained from the limited and non-renewable phosphate rock resource (Cordell et al. 2009). It has been estimated that 15-20% of world demand could theoretically be satisfied by recovering P from domestic waste streams (Yuan et al. 2012). The enhanced biological phosphorus removal (EBPR) process has been recommended as a promising strategy to achieve sustainable wastewater P removal and also facilitate subsequent recovery (Yuan et al. 2012), with economic and environmental advantages over traditional chemical P removal and other tertiary treatment techniques (e.g., ionic exchange, filtration, sorption) in full-scale wastewater treatment plants (WWTPs) (Gu et al. 2016, Gu et al. 2019). However, many facilities still suffer from unpredictable EBPR performance upsets for reasons that are unclear (Gu et al. 2008, Oehmen et al. 2007). The advantages of EBPR are thus often offset in practice by the need to have chemicals as a standby option to ensure reliable and consistent P removal performance to meet compliance. The underlying mechanisms governing EBPR

performance and stability are still not fully understood (Neethling et al. 2006, Oehmen et al. 2007, Onnis-Hayden et al. 2020).

The enrichment of polyphosphate (polyP) accumulating organisms (PAOs) in activated sludge is considered to be responsible for achieving successful EBPR operation because of their unique ability of luxury phosphate uptake (Oehmen et al. 2007). Typically, PAOs (e.g., *Candidatus* Accumulibacter (hereafter *Accumulibacter*)) can anaerobically use the energy (i.e., adenosine triphosphate (ATP)) generated from the cleavage of polyP to assimilate readily biodegradable carbon substrates (e.g., volatile fatty acids (VFAs)) and store them as intracellular metabolites or polymers (e.g., polyhydroxyalkanoates (PHAs)). Under the subsequent aerobic/anoxic conditions, PAOs utilize the stored substrates as carbon and energy sources for biomass growth, maintenance, and polyP synthesis accompanied by enhanced uptake of inorganic orthophosphate (Pi) (Mino et al. 1998). Net P removal from wastewater is achieved by wasting sludge with high P content (mainly in the form of polyPs). Therefore, the synthesis and/or degradation of intracellular polyPs, as well as their structural characteristics (e.g., chain length, cation composition, structure), potentially play critical roles in influencing the EBPR performance and stability.

The polyPs inside PAO cells, at levels up to 35% of dry weight in biomass (Wentzel et al. 2008), are biopolymers consisting of three to several hundred Pi units linked by high-energy phosphoanhydride bonds. The impacts of polyP chain length on EBPR performance/stability have not yet been extensively studied. Limited previous studies reported varying polyP lengths (indicated by different distributions of acid-soluble, short-chain (3-20 Pi) and acid-insoluble, long-chain (>20 Pi)) in pilot and full-scale EBPR systems (Lindrea et al. 1998, Müssig-Zufika et al. 1994). Lindrea et al. (1998) suggested that the acid-insoluble longer polyPs seemed to be associated with more stable EBPR performance. Previous studies have also suggested a possible impact of varying polyP-associated metal compositions on EBPR microbial ecology and performance (Li et al. 2019, Lindrea et al. 1998, Lockwood et al. 1990, Schönborn et al. 2001). A positive correlation

between the magnesium content in polyP granules and the long-term stability of EBPR systems has been shown by Li et al. (2019). These observations suggest that the variations in polyP length and composition are possibly associated with and reflect the phylogenetic and/or phenotypic diversity of PAOs in these ecological systems.

Beyond the field of EBPR, polyP has also been persistently identified in all domains of life. This ancient and ubiquitous molecule has been found as a polyanionic scaffold of varying polymer lengths that facilitates the orientation and assembly of major biological macromolecules (e.g., phospholipids, nucleic acids, proteins), and influences their biological roles (Brown and Kornberg 2004). Therefore, in-depth understanding of the metabolism, structure and function of intracellular polyPs are of great importance to many biological and environmental fields, but have been hampered by the lack of tools that enable a finer-resolution differentiation and characterization of polyP fractions. Current methods available for polyP chain length characterization include: (1) Chemical digestion, which can roughly separate extracted polyPs into acid-soluble (3-20 Pi) and acid-insoluble (>20 Pi) fractions (Serafim et al. 2002); (2) High-performance liquid chromatography (HPLC), which can separate extracted polyPs of <35 Pi (Majed et al. 2012b); (3) Thin-layer chromatography (TLC), which can only separate polyPs up to a length of 8 Pi (Majed et al. 2012b); (4) Solution ^{31}P-nuclear magnetic resonance (^{31}P-NMR) spectroscopy, which could estimate average polyP length up to 1000 Pi, potentially without prior extraction (Christ et al. 2020, Hupfer et al. 1995, Smith et al. 2010); and (5) Polyacrylamide gel electrophoresis (PAGE), which can separate extracted polyPs up to a few hundred Pi units, and is an established and widely used method for separating extracted polyPs from pure culture organisms (Ogawa et al. 2000, Smith et al. 2010). So far, only ^{31}P-NMR has been used to study polyP characteristics in full-scale EBPR samples (Staal et al. 2019).

Among all of the aforementioned methods, only ^{31}P-NMR and PAGE could potentially allow for fine-scale separation and determination of polyP length distributions, especially for the long-chain polyPs (up to hundreds of Pi) which are very likely to be present in EBPR systems. More

powerful P analysis tools are in demand to help elucidate the biochemistry, nature, roles and dynamics of intracellular polyPs *in vitro* and/or *in vivo*. Besides, although the progress in (meta)genomics has advanced our knowledge of phylogenetic and metabolic diversities of microbial populations in EBPR systems (Flowers et al. 2013, Gao et al. 2019, Martín et al. 2006), phenotypic profiling of individual cellular characteristics and linking genotypes with phenotypes and functions remain challenging. Therefore, there is a need for a phenotyping method particularly suitable for high-diversity ecosystems and complementary to genomics analysis. Recent advances in single-cell Raman spectroscopy (SCRS) can be applied for *in situ* identification and quantification of the specific functionally relevant populations, such as PAOs and glycogen accumulating organisms (GAOs) in the EBPR process. This can be realized via simultaneous detection of the multiple polymer peaks (i.e., polyP, PHA and glycogen) (Majed et al. 2012a, Majed and Gu 2010, Majed et al. 2009), as well as phenotypic classification of EBPR communities through the single-cell Raman spectral fingerprints (Li et al. 2018, Wang et al. 2020). However, the potential of SCRS for polyP characterization (e.g., chain length) and phenotype-function relationship deduction has not yet been explored for PAOs in EBPR systems.

The objectives of this study were to explore and evaluate SCRS as an intracellular polyP characterization method, and to demonstrate its potential applications for phenotypic profiling of the diverse polyP characteristics of PAOs in EBPR systems. We explored and demonstrated that SCRS can be a label-free and non-destructive method for characterizing intracellular polyPs of PAOs at both population and single-cell levels, and therefore can potentially serve as a sensitive tool to assess the PAO diversity and phenotypic dynamics. Two conventional polyP characterization approaches, namely ^{31}P-NMR and PAGE, were optimized and employed as references for the SCRS-based method validation, and EBPR biomass from both lab-scale reactors and full-scale WWTPs were examined. The results bring new insights into the variations and dynamics of polyP length and composition of PAOs in EBPR systems, where polyP features can

be an indicator of PAO phenotypic diversity and associated EBPR performance. The methods and outcomes can be potentially applied and extended to other environmental and biological systems where biologically originated polyP plays an important role.

2. Materials and methods

2.1 Synthetic P chemicals

Synthetic P chemicals with different chain lengths were purchased for the study (Table S1). Monosodium phosphate ($NaH_2PO_4 \cdot 2H_2O$, MSP), tetrasodium pyrophosphate ($Na_4P_2O_7 \cdot 10H_2O$, TSPP), sodium tripolyphosphate ($Na_5P_3O_{10}$, STPP), sodium hexametaphosphate ($(NaPO_3)_6$, SHMP, an impure compound containing cyclic and linear polyPs) and sodium phosphate glass type 45 (SPG45) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Sodium tetraphosphate ($Na_6P_4O_{13}$, STP) and two ammonium polyPs labeled with alkyne tags ($C_5H_{35}N_7O_{19}P_6$ (pentyne-P6) and $C_5H_{39}N_8O_{22}P_7$ (pentyne-P7)) were synthesized in Dr. Henning J. Jessen's laboratory at the Institute of Organic Chemistry, University of Freiburg, Germany (Singh et al. 2019a, Singh et al. 2019b). Medium-chain and long-chain polyPs, namely p100 and p700 (nominal average chain length of ~100 and ~700 Pi respectively), were purchased from KeraFAST (Boston, MA, USA).

2.2 EBPR activated sludge and PAO-enriched culture

Lab-scale EBPR samples: Lab-scale activated sludge samples for ^{31}P-NMR and PAGE analyses were collected from the aerobic phase of four sequencing batch reactors (SBRs) operated at Northeastern University, USA, with solid retention times (SRTs) of 3, 5, 10, and 20 days respectively. The SBRs, with working volume of 4 L and hydraulic retention time (HRT) of 12 h, had four 6-h cycles per day. Each cycle consisted of: 7 min fill followed by 110 min anaerobic phase, 180 min aerobic phase, 58 min settling, and 5 min decanting. The SBRs were fed with synthetic wastewater as described in Table S2 and previous publications (Li et al. 2018, Schuler and Jenkins 2003). The total chemical oxygen demand (COD) and P concentrations were 200 and 8 mg/L respectively, yielding an influent COD/P ratio of 25 (relevant for full-scale EBPRs in the

USA). Each SBR was operated at 20–22°C for at least 3 times of the SRT to achieve steady-state prior to performance and polyP analysis. To compare the Raman spectra of intracellular polyPs in different PAO genera, a lab-scale SBR was operated to enrich *Accumulibacter*-like PAOs (>50% by biovolume) (Majed and Gu 2020), and enriched culture of *Tetrasphaera* PAOs (>80% by biovolume) (Marques et al. 2018, Marques et al. 2017) from a lab-scale EBPR reactor was kindly provided by Dr. Adrian Oehmen and Dr. Ricardo Marques at Universidade Nova de Lisboa, Portugal. The operational conditions and feed composition of the reactor have been described by Marques et al. (2017).

Full-scale EBPR samples: Full-scale activated sludge samples for PAGE analysis were collected from the aerobic zone of five full-scale EBPR facilities in the USA, including Clark County Water Reclamation District (CL), Las Vegas Water Pollution Control Facility (LV), and three plants of Hampton Roads Sanitation District: Atlantic Treatment Plant (AT), Virginia Initiative Treatment Plant (VIP), and Nansemond Treatment Plant (Nan). Activated sludge samples for SCRS analysis were collected from four different full-scale side-stream EBPR (S2EBPR) configurations in the USA, including: (a) Side-stream return activated sludge (RAS) fermentation (SSR) configuration at the South Cary Water Reclamation Facility (SC); (b) Side-stream RAS fermentation with supplemental carbon addition (SSRC) configuration at the Westside Regional Wastewater Treatment Plant (WR); (c) Side-stream mixed liquor suspended solids (MLSS) fermentation (SSM) configuration at the Cedar Creek Wastewater Treatment Facility (CC); and (d) Unmixed in-line MLSS fermentation (UMIF) configuration at the Kurt R. Segler Water Reclamation Facility (Hen) (Figure S1). The main characteristics, performance and phylogenetic information of the full-scale EBPR plants are summarized in Table S3 and S4.

2.3 P release and uptake batch tests

To investigate the effect of different carbon sources on PAO phenotypic diversity, P release and uptake batch tests were conducted following a previously described protocol (Gu et al. 2008, Wang et al. 2019) to restore polyP reserves in PAO cells for SCRS analysis. Sodium acetate of 80

mg COD/L or real wastewater was then added, and the sludge was continuously purged with nitrogen gas to maintain anaerobic conditions for 1 h. Then, the sludge was bubbled with air to obtain aerobic conditions for 3 h. Similarly, a batch test was performed on the lab-scale *Tetrasphaera*-enriched culture under anaerobic (4 h) and aerobic (4 h) conditions with sodium casein hydrolysate as the sole carbon source. Samples taken from the end of the batch tests were washed, homogenized, and mounted on Raman grade slides, as described previously (Majed et al. 2012a).

2.4 Extraction methods

Prior to polyP analysis by ^{31}P-NMR and PAGE, three polyP extraction methods were applied to compare the effectiveness for sufficient and non-destructive polyP release/extraction, which were as follows:

(1) EDTA extraction (Jing et al. 1992): Fifty milliliters of mixed liquor samples were washed three times in 0.85% NaCl, twice in 0.1 M Na_2EDTA, re-suspended in 2 mL 0.5 M Na_2EDTA, and homogenized with a 26-gauge needle. Centrifugation between washings was set at 6,000 g for 3 min.

(2) EDTA+sonication (Jing et al. 1992): Mixed liquor samples were washed as in method (1) and heated in a 50°C water bath for 10 min to denature enzymes. Cells in the EDTA solution were then disrupted with a sonicator (Digital Sonifier 450, Branson Ultrasonics Co., Ltd., Shanghai, China) at 90 W for 10 min. The lysed sludge was centrifuged (13,000 rpm and 4°C for 10 min), with both the supernatant and the re-suspended and homogenized sludge pellet collected for polyP analysis.

(3) EDTA+NaOH (Cade-Menun et al. 2002): 10 g samples of wet sludge were re-suspended in 100 mL of NaOH/EDTA mixture (1:1 mix of 0.5 M NaOH prepared with autoclaved DI water and 0.1 M autoclaved Na_2EDTA). All samples were extracted overnight (18 h) at room temperature, with occasional stirring (30 min stirring for every 2 h). Samples were then centrifuged at 10,000

rpm for 25 min. The supernatant was frozen at -80°C and defrosted immediately before polyP analysis.

Synthetic polyPs (SHMP and SPG45) were used as controls for evaluating the potential polyP hydrolysis during extraction.

2.5 ^{31}P-NMR analysis

^{31}P-NMR spectra were obtained with a Bruker Avance 400 MHz NMR spectrometer (Bruker Instrument, Billerica, MA, USA). Prior to ^{31}P-NMR analysis, each sample was mixed with 10% D_2O for locking. The chemical shifts (relative peak positions) of P compounds were determined relative to an external alkaline standard of 0.05 M KH_2PO_4 in a NaOH/EDTA solution with 10% D_2O. Resonance peaks for different P species were assigned according to literature at the sample pH condition (Burt 2018). The 90° pulse width was determined to be 18 μs on a 0.0485 M triphenyl phosphate standard in $CDCl_3$. NMR spectra were acquired with a 30° radiofrequency pulse, 555.7 ms acquisition time, and 1.444 s delay time. The gain and line broadening were set to 60 and 4 Hz respectively. The number of transients (scans) for each spectrum was typically between 1,000 and 25,000. With this NMR instrument, it was determined that several thousand scans for each sample were necessary for the smaller peaks (e.g., terminal polyP peak) in the spectra to be visible (i.e., to have a signal-to-noise ratio (SNR) of at least 5:1). The average polyP chain length determination, based on the polyP-related peaks (PP1-PP4), is described in the Supplementary Information. The amount of released and visible polyPs were semi-quantified by integrating all of the detected polyP peaks (Shi and Lee 2007).

2.6 PAGE analysis

PolyP analysis by PAGE was performed according to the procedures in previous literature (Clark and Wood 1987). Briefly, a 15% (w/v) polyacrylamide gel (280 mm×16.5 cm×0.75 mm) was prepared with 7 M urea in 1X TBE buffer, which contains a 19:1 ratio (w/w) of acrylamide to bisacrylamide, and pre-electrophoresed at 800 V, 4°C for 1 h. The EDTA+NaOH extracted samples were centrifuged at 6,000 g for 2 min, with the supernatant loaded in the loading dye solution (1X

TBE buffer, 30% sucrose, 0.05% bromophenol blue). The gels were then electrophoresed at 800 V, 4°C until the bromophenol blue migrated approximately 14 cm. The gel was then soaked into 10% acetate, 10% methanol for 15 min, stained with 0.5% toluidine blue O, 25% methanol, 5% glycerol, 5% acetate for 15 min, and de-stained with 25% methanol, 5% glycerol, 5% acetate for 10 min several times. The stained polyP bands were visualized under white light using a Bio-Rad Gel-Doc XR imaging system (Bio-Rad Laboratories, Hercules, CA, USA). The chain lengths of polyPs less than 30 Pi were counted based on the methods introduced previously (Robinson et al. 1984). The chain lengths of polyPs longer than 30 Pi were extrapolated from a semi-log plot of migration distance *versus* polyP chain length (Ogawa et al. 2000). Synthetic polyPs (SHMP and SPG45) were loaded as polyP size markers, which contain heterogeneous populations of 1–30 Pi (Stover et al. 1994) and 30–150 Pi (Smith et al. 2010), respectively.

2.7 SCRS analysis and data processing

Raman spectra of synthetic P chemicals and activated sludge samples were acquired using a LabRam HR Evolution Raman spectrometer (Horiba Jobin Yvon, Kyoto, Japan) configured with 532 nm Nd:YAG laser and 600 gr/mm grating. A water-immersion 100× objective with a numerical aperture (NA) of 0.9 and working distance (WD) of 0.21 mm was used to observe samples and collect Raman signals in the spectral range from 200 to 3200 cm^{-1}. A silicon wafer was used for automatic Raman shift calibration at the wavenumber of 520.7 cm^{-1} every day before sample measurement. For synthetic P chemicals, the acquisition time for each spectrum was set to 5 s and the laser power at the sample was ~6.0 mW. For each sludge sample, Raman spectra for at least 80-250 single cells were analyzed twice with 20 s acquisition time and ~0.6 mW laser power. Sampling sizes for this study were considered sufficient based on our statistical evaluation method by Li et al. (2020). All Raman spectra were processed with cosmic spike removal, background subtraction, and baseline correction using LabSpec 6 software (Horiba Jobin Yvon, Kyoto, Japan). Quality control was conducted by excluding the spectra showing a low SNR. Raman spectra of the cells lacking major characteristic peaks from bacterial components such as phenylalanine (~1002

cm^{-1}) and amide I (~1657 cm^{-1}) were also excluded. Cells with both of the polyP characteristic peaks (i.e., bands at ~690 cm^{-1} for P–O–P stretch and ~1160 to ~1175 cm^{-1} for PO_2^- stretch) were assigned as candidate PAOs (Li et al. 2018, Majed et al. 2012b). The values of polyP peak positions in individual PAO cells were then extracted to construct the relative frequency histogram with a bin width of 4 cm^{-1}. The Gaussian curve was utilized to fit the histogram and describe the polyP peak distribution pattern in the sample. To evaluate the reproducibility of the results, a total of 150 SCRS-identified PAO cells were randomly selected from three biological replicates (50 PAO cells per sample) in the *Tetrasphaera*-enriched culture to evaluate the varitions in the polyP peak position among biological replicates. The results showed that even for a mixed culture, the obtained values of PO_2^- peak position in each replicate are normally distributed (Figure S2), and the mean value of the three replicates is 1162.6 cm^{-1} with a standard deviation <0.7 cm^{-1}.

2.8 Chemical analyses

Samples periodically collected from the influent and effluent of the four lab-scale SBRs, and during the batch tests, were filtered through 0.45 μm Millipore filters (MilliporeSigma, Burlington, MA, USA). Total phosphorus (TP), ammonia (NH_4–N), and COD in the filtrate were then determined in accordance with Standard Methods (APHA 1985). Total suspended solids (TSS) and volatile suspended solids (VSS) were determined weekly. All data were determined at least in duplicate.

3. Results and Discussion

3.1 Intracellular polyP analysis by PAGE

Synthetic polyP chemicals and EDTA+NaOH extracted polyPs from EBPR sludge samples were resolved to estimate chain length range and distributions in both lab-scale and full-scale EBPR systems. With a polyacrylamide concentration of 15%, polyP chains shorter than ~250 Pi could be resolved and separated (Figure 1). The SHMP and SPG45 heterogeneous standards were resolved to be mixtures of <30 Pi polyPs, and ~20 to ~160 Pi polyPs respectively (Figure 1a), which agreed

with their length ranges reported from previous literature (Smith et al. 2010, Stover et al. 1994). With the PAGE method, different polyP length ranges were found in the samples from lab-scale EBPR reactors (Figure 1b) that were operated with four varying SRTs, as well as for five full-scale (Figure 1c) wastewater treatment plants that performed EBPR. Interestingly, the PAGE-estimated polyP chain lengths for lab-scale reactors (mostly all >100 Pi) were found to be longer than those from the full-scale EBPR sludge samples (from ~50 to ~170 Pi) (Figure 1b and 1c). PAGE bands for the lab-scale samples look darker than those for full-scale samples (Figure 1), indicating a higher polyP content. This is probably because PAO abundance is usually higher in the lab-scale EBPR reactors than in full-scale systems.

For the first time, we semi-quantitatively compared the polyP chain length distributions from different lab-scale (with 3, 5, 10, 20 days SRTs) and full-scale EBPR systems, and revealed the variations. For the four lab-scale EBPR reactors, three types of polyP length distribution patterns were observed (Figure 1b). PolyPs with relatively short length (~100 to ~170 Pi) were observed for the SBR with 3-day SRT, which has the lowest P removal efficiency of only 18%. The SRT of 3 days at 20–22°C in this study is near or below the model-predicted washout SRT for PAOs of 2.7–4.3 days at 20°C (Brdjanovic et al. 1998), therefore leading to poor performance. Longer polyPs (~120 to ~250 Pi) were observed for SBRs with longer SRTs (5, 10, and 20 days) that had relatively effective and stable EBPR performance (P removal efficiency of 96%-98%). For a short period (2 days), the performance in SBRs with 10- and 20-day SRTs experienced deterioration due to malfunction of dissolved oxygen supply (P removal efficiency of 21% and 81%, respectively). The polyPs in these samples exhibited a narrower length range of ~170 to ~250 Pi, missing those polyPs of ~120 to ~170 Pi that were observed when the SBRs had good performance. Previous studies of the same lab-scale SBRs indicated that *Dechloromonas* (note that *Dechloromonas* was covered by the Accumulibacter-like PAO mix FISH probes we used) dominated the PAO populations with SRT-dependent phylogenetic and phenotypic diversities (Li et al. 2018). In this study, the variations in polyP length distribution pattern revealed by PAGE analysis may also be associated with the

phylogenetic or phenotypic variations in PAOs and associated polyP metabolisms as observed in the previous studies (Li et al. 2018, Li et al. 2019).

The five full-scale EBPR plants with different configurations (i.e., CL, LV, VIP, Nan, and AT plants), which had varying P removal efficiency, exhibited a distinctively different range and distribution of polyP length (Figure 1c). Interestingly, it was noticed that the polyP bands covering all the regions between 120 and 170 Pi were only detected in the VIP and CL plants, which had 99% and 91% P removal efficiency, respectively. In comparison, the 150–170 Pi polyP bands were missing in the Nan plant with 79% P removal efficiency, and the whole 120–170 Pi polyP bands were missing in the AT plant with only 44% P removal efficiency. These results seem to suggest a possible association of the intracellular polyPs of ~120 to ~170 Pi lengths with the good P removal performance of both lab-scale and full-scale EBPR systems, although future investigation is recommended to confirm this. The varying polyP length ranges among the different EBPR plants, as revealed for the first time here, suggested that the polyP length distribution is likely associated with the specific PAO populations and their phylogenetic and/or phenotypic diversity. Therefore, the polyP length characterization may be among the indicators for describing the PAO metabolic and functional diversities and dynamics, similar to the metal composition in the polyP. Li et al. (2019) discovered that the varying intracellular polyP-metal compositions and metal/P stoichiometry were associated with different PAO phylogeny and phenotypes, as well as with the performance stability in lab-scale EBPR reactors that were operated with varying SRTs.

In comparison with other polyP length characterization methods (e.g., chemical digestion, ^{31}P-NMR), the PAGE method has some advantages for the semi-quantitative determination of polyP chain length, including: (1) PolyP separation on polyacrylamide gels is not heavily impacted by complexation with metal ions or other cellular compounds; (2) The direct visualization of the separated polyPs shows not just an average length estimation, but also the range and distribution of different polyP lengths; (3) For samples with relatively low polyP content (e.g., the SBR of 3

days SRT in this study), the gel resolved polyP chains could still be detected by simply loading more samples (Figure 1). Overall, the PAGE method seemed to be a suitable method for polyP length determination, especially for polymers with relatively long chain lengths (up to hundreds of Pi) presenting within complex biological samples (e.g., EBPR systems).

3.2 Intracellular polyP analysis by ^{31}P-NMR

Theoretically, polyPs can be directly detected by ^{31}P-NMR, yet intracellular polyPs of PAOs, especially long-chain types, maybe invisible in ^{31}P-NMR spectra, due to complex formation with other cell structures and compounds such as sugars, proteins, nucleic acids, or metal ions (i.e., Mg^{2+}, K^{+}, Ca^{2+} and Na^{+}) (Hupfer et al. 2008). In this study, intact activated sludge collected from the aerobic zone of EBPR plants and the aerobic phase of SBRs, which contained presumably high concentrations of polyP, were directly analyzed with ^{31}P-NMR. None of the polyP-related peaks (PP1-PP4) were observed in these samples, even after concentrating the sludge by 50 times. This showed that prior extraction steps would be a prerequisite for the detection of polyP-related signals by ^{31}P-NMR within an acceptable resolution (Pilatus et al. 1989, Staal et al. 2019, Tao et al. 2020).

The ideal extraction approach needs to be able to effectively release intracellular polyP from the activated sludge, but with minimized breakdown/hydrolysis of polyP chains. Here, we evaluated and compared the effectiveness of three extraction methods for sufficient and non-destructive polyP release/extraction, using: (1) EDTA only, (2) EDTA+sonication, (3) EDTA+NaOH. Among the three protocols, EDTA+NaOH extraction revealed all the polyP-related peaks (PP1-PP4) in most of the samples (Figure S3a), allowing for estimation of the average polyP chain length using Equation (1) in the Supplementary Information. Contrarily, only the middle P signal (one broad peak for PP2+PP3+PP4) of polyPs could be detected in EDTA only or EDTA+sonication extracted samples (Figure S3b), and thus these were not feasible for ^{31}P-NMR-based high-resolution polyP chain length analysis. EDTA is widely used for polyP extraction as it can release polyP chains from paramagnetic ions (such as Fe^{3+}, Mn^{2+}), which helps to make polyP chains more visible in ^{31}P-NMR, reducing line broadening and improving spectral quality (Cade-

Menun 2005, Staal et al. 2019). However, some studies suggested that EDTA could only enter the periplasm and not the cytoplasmic membrane of bacterial cells (Jing et al. 1992), and thus may only be able to release the polyPs located in the periplasm rather than those inside the cytoplasmic membrane. The semi-quantified amounts of visible polyPs from EDTA+sonication extraction were almost the same as those from EDTA extraction only, and negligible polyP signal was detected in the supernatant from EDTA+sonication extraction (Figure S4). This indicated that the sonication treatment was probably not sufficient enough to break down cell membranes in the studied sludge samples. Alternatively, the EDTA+NaOH protocol has been successfully applied to extract polyP from other environmental samples (e.g., lake sediments, soil) (Cade-Menun 2005), and the addition of NaOH could help lyse the cell membrane, dissolve polyP from the complex, and also inhibit the enzymes for polyP hydrolysis (Cade-Menun et al. 2002, Cade-Menun 2005, Hupfer et al. 1995). As shown in Figure S4, the amounts of polyP extracted by the EDTA+NaOH protocol were more than twice the amounts extracted by the other two methods.

After EDTA+NaOH extraction, the average chain lengths of extracted polyPs from the lab-scale EBPR reactors were estimated to range from 31 to 48 Pi (Table 1). This is significantly lower than the PAGE-resolved length ranges with the same extraction step (~50 to ~250 Pi). The possible reasons for the different results derived from the two methods include: (1) Long-chain polyPs are more likely to remain invisible in ^{31}P-NMR, despite the EDTA+NaOH treatment, as they tend to form more complex and rigid binding with other molecules (Cade-Menun 2005, Hill et al. 1989, Pilatus et al. 1989); (2) ^{31}P-NMR seemed to exhibit more pronounced underestimation of the internal P groups (PP2–PP4) for long-chain polyPs (Kulaev et al. 2005), as well as overestimation of the terminal P group (PP1) due to the overlap with β–P and/or γ–P groups in nucleotides (Christ 2020), leading to potential underestimation of average polyP length; (3) PAGE-estimated polyP chain length may be overestimated as the short-chain polyP (<15 Pi) sometimes cannot be stained by the dye toluidine blue O (Christ et al. 2020). Besides, even with the overnight acquisition, ^{31}P-NMR failed to detect PP1 or PP2 signals in the EDTA+NaOH extracted polyPs from the SBR at

an SRT of 3 days (Table 1), implying its insensitivity towards the assessment of shorter-chain polyPs even with the sample containing its highest abundance.

Previous chemical digestion analysis, with a very limited length resolution, indicated a possible correlation between longer polyP length (roughly >20 Pi) and better EBPR performance stability (Lindrea et al. 1998). In our study, both PAGE and ^{31}P-NMR analyses provided more quantitative polyP length characteristics at a finer scale among different EBPR systems as compared to the previous study. As revealed by ^{31}P-NMR, the lab-scale SBR of 5 days SRT with the highest P removal efficiency (97.5%) was the one with the longest average polyP length (48 Pi) (Table 1), indicating an increasing trend of polyP length along with increased system performance. Although the polyP characteristics (i.e., chain length distribution and average chain length) determined by PAGE and ^{31}P-NMR are not completely consistent for the reasons discussed above, nevertheless, they still revealed new insights into the range and variations in polyP length in various EBPR systems that are likely associated with the PAO phylogenetic and/or phenotypic profiles and, therefore, potentially have linkage with the system performance. Further studies are needed with an extended depth of focus on revealing the correlation between the polyP length and EBPR performance, which is beyond the core focus of this study.

3.3 Intracellular polyP analysis by SCRS

3.3.1 Characterization of synthetic P chemicals

Recognizing the inconsistency and limitations of polyP length characterization with ^{31}P-NMR or PAGE methods that require sample pre-extraction, we further explored the application of SCRS for the direct characterization of polyPs with various lengths. First, we examined synthetic P chemicals with lengths ranging from 2 to 700 Pi, and also polyPs with alkyne tags to evaluate the potential Raman spectrum shifts in response to polyP chain length (Figure 2). The Raman peak for P–O–P stretching vibration of the phosphoanhydride bond occurs at ~740 cm^{-1} in the spectrum of a dimer (TSPP, 2 Pi), and it moves to lower wavelengths of ~685 cm^{-1} in the spectrum of a hexamer

(SHMP, 6 Pi). Meanwhile, the peak representing PO_2^- stretches of the ester chain, shifts from ~1024 cm^{-1} in TSPP to ~1163 cm^{-1} in SHMP, indicating a possible association between polyP chain length and Raman peak position (Majed et al. 2009). However, the changing trends of P–O–P and PO_2^- peaks were not very pronounced for long-chain polyPs of 45 to 700 Pi (constantly at 682-691 cm^{-1} and ~1164 cm^{-1}, respectively), which is probably due to the high heterogeneity of the polyP mixtures. For the organic alkyne-tagged polyPs (pentyne-P6 and pentyne-P7), the PO_2^- peak is shifted from ~1164 cm^{-1} to lower wavenumbers of 1152 and 1137 cm^{-1} respectively, which is probably attributed to the incorporated Raman-active organic tag containing a C≡C triple-bond (peak shown at ~2116 cm^{-1}). The results indicated a potential relationship between the polyP structure and the Raman signature peak shifts, suggesting a possibility of characterizing the polyP length, distribution and bound molecules using SCRS.

3.3.2 Characterization of polyPs in lab-scale EBPR samples via SCRS

To investigate the single-cell Raman spectra of the intracellular polyP granules, we first explored and examined the distribution and shift of the polyP signature peak positions (P–O–P and PO_2^- peaks) among individual PAO cells in two lab-scale EBPR reactors, with one that contained predominantly *Tetrasphaera* PAOs (>80% by biovolume) and the other one consisting mostly of *Accumulibacter*-like PAOs (>50% by biovolume) (Figure 3 and S5). Unlike the wide distribution pattern of P–O–P peak position among *Tetrasphaera* cells (Figure S5a), the PO_2^- peak position showed a narrow normal distribution centered at ~1162 cm^{-1}, with >70% ranging between 1157 cm^{-1} and 1165 cm^{-1} in the samples from the *Tetrasphaera*-dominated EBPR reactor (Figure 3a). In comparison, in the EBPR reactor enriched with *Accumulibacter*-like PAOs, the PO_2^- peaks were mostly observed at ~1170 cm^{-1}, ranging between 1168 cm^{-1} and 1177 cm^{-1} (Figure 3b). The distinguishable and consistently detected shift of PO_2^- peak positions distribution among PAO cells observed in the above two EBPR cultures that were dominated by two different PAO genera indicate that the shift is likely indicative of different microbial populations. In addition, the shifts of polyP signature peak position are most likely associated with polyP structural changes (e.g.,

chain length, linear, branched or cyclic polyP structure, molecular coordination, structure and stability of polyP complexes with metals or macromolecules, as well as the composition of metals or organic ligands). The Raman shift noise due to instrumental and experimental variations is recognized (García-Timermans et al. 2018, Guo et al. 2018), however, it is unlikely the primary contributor to the observed statistical shifts in these different PAOs-enriched cultures, because all the measurements were performed on the same Raman system with pre-verified instrument calibration, spectral acquisition and data processing steps (Li et al. 2018, Majed and Gu 2010). In this study, we performed three biological replicates in a Tetrasphaera-enriched culture. The results showed that the obtained values of PO2- peak position in each replicate are normally distributed, and the mean value of the three replicates is 1162.6 cm-1 with a standard deviation <0.7 cm-1, which is much smaller than the variation among different sludge samples (Figure S2).

Furthermore, besides the selected polyP signature peak, the whole region of the information-rich Raman spectra could also be utilized for phenotypic separation and classification at a finer level. To probe the validity of this concept, we further examined the single-cell Raman spectra from four lab-scale EBPR reactors (operated with four different SRTs with varying performance stability) that contained high proportion of *Accumulibacter*-like PAOs (35–57% by biovolume via FISH analysis) and the 16s rRNA gene analysis indicated that the dominant PAOs was *Dechloromonas* (Li et al. 2018). The 16s rRNA gene analysis indicated that the four reactors with different SRTs exhibited different level and distribution of PAO sub-groups (Li et al. 2018). The unsupervised multivariate statistical analysis, hierarchical clustering analysis (HCA), was therefore performed to classify the collected Raman spectra based on their intra-spectral similarities (see details in Supplementary Information), and four phenotypic groups were identified and defined as operational phenotypic units (OPUs), namely OPU1 to OPU4, with 12 subgroups (sub-OPUs) (Figure S6). Our previous study suggested that the EBPR performance stability was better correlated with Raman OPU-based phenotypic diverisity rather than OUT-based phylogentic

diversity (Li et al. 2018). In this study, we re-analyzed the Raman-identified PAOs that were distributed in 8 out of 12 sub-OPUs, and among them, the PO_2^- peak position (1160-1173 cm^{-1}), P–O–P peak position (680-710 cm^{-1}), peak distance (458-479 cm^{-1}) and levels of intracellular polymers (e.g., polyP, glycogen, PHA) varied largely (Table 2). For the two predominant subgroups with sufficient sample size (OPU4a and 4b, which account for 76% of all PAOs), the PO_2^- peak positions exhibited normal distribution and the mean values separated by a distance of 9 cm^{-1} (Figure 4a), which validates the resolving power of the SCRS-based phenotypic classification at a finer level. Meanwhile, a considerable overlap of P–O–P peak position between OPU4a and 4b was observed (Figure 4b). It is probably due to the amorphous nature of the polyP with the atoms around the P–O–P backbone adopting many different orientations, which leads to a broad P–O–P peak (Figure S6) that is unsuitable for tracing PAO phenotypic change. It was noticed that the OPU4b-PAO subgroup was frequently detected in the samples from all EBPR reactors (Li et al. 2018), implying its potential association with system performance in the acetate-fed EBPR reactors, which requires further validation. Although directly linking a specific PAO phenotypic group with phylogenetic identity via SCRS method only is still challenging and requires building database and further investigation, the preliminary results have supported the notion that the SCRS-based polyP characterization is associated with and therefore reflects the phenotypic heterogeneity of PAOs. The extracted high-resolution fingerprint-like information may serve as a complementary tool for future in-depth studies such as gene-function relationship deduction.

3.3.3 Characterization of polyPs in full-scale EBPR samples via SCRS

In order to explore the potential application of SCRS for polyP characterization in complex environmental samples, we further examined the single-cell Raman spectra of intracellular polyPs in individual PAO cells from four full-scale S2EBPR systems (i.e., SC, WR, CC, and Hen plants) (Figure S1). S2EBPR refers to a suite of modified EBPR configurations (e.g., SSR, SSRC, SSM, UMIF) that involve a side-stream anaerobic reactor/zone where a portion or all of return activated

sludge (RAS) or mixed liquor (MLSS) is diverted to promote sludge fermentation (Barnard et al. 2017, Gu et al. 2019). The extended anaerobic retention time and conditions in the side-stream reactor were shown to provide more competitive advantage to many PAOs (e.g., *Tetrasphaera*, *Accumulibacter*), leading to higher phylogenetic/phenotypic diversity and improved process stability than the conventional EBPR systems (Onnis-Hayden et al. 2020, Wang et al. 2019). Unlike the unimodal distribution exhibited in the lab-scale PAO-enriched cultures (Figure 3 and 4a), multiform and multimodal distributions of the PO_2^- peak position in a wider range (1148–1179 cm^{-1}) were observed in the full-scale samples from the four different S2EBPR plants (Figure 5), reflecting the configuration-specific phenotypic diversity and heterogeneity of PAO cells. According to the Gaussian curves of the binned histogram data, the distribution patterns of PO_2^- peak position could be roughly clustered into three phenotypic groups, centered at ~1157 cm^{-1}, ~1164 cm^{-1} (*Tetrasphaera*-like phenotype) and ~1172 cm^{-1} (*Accumulibacter*-like phenotype), respectively. It was noticed that in the CC plant (SSM configuration) that has a relatively poorer P removal efficiency of 67% compared to other facilities, the relative abundance of PAO phenotype with PO_2^- peak centered at ~1157 cm^{-1} was much lower than the other plants (Figure 5c), implying possible association of this unknown PAO phenotype with P removal performance in full-scale EBPRs. In comparison, the WR plant (SSRC configuration) with the highest P removal efficiency of 97% showed a broad PO_2^- peak distribution curve covering all the PAO phenotypes (Figure 5b), indicating a higher PAO phenotypic diversity that is possibly associated with EBPR performance. The relatively low phenotypic diversity in the CC plant might be because the SSM configuration is more similar to a conventional EBPR system with the side-stream fermenter fed with MLSS from the mainstream anaerobic zone, compared to the "classical" S2EBPR configurations (e.g., SSR and SSRC) with the dedicated side-stream RAS fermentation reactors that potentially providing a specialized ecological niche for the co-occurrence of different PAO phenotypes (Gu et al. 2019). These observations, which could not be unveiled by the previously employed 16S rRNA gene amplicon sequencing and fluorescence in situ hybridization (FISH) (Onnis-Hayden et al. 2020),

suggested that the SCRS-based polyP characterization could potentially be a unique phenotyping method that is complementary to phylogenetic analyses for the study of microbial communities in full-scale EBPR processes. Further investigation, with the assistance of advanced SCRS-enabled technique (e.g., Raman–FISH (Fernando et al. 2019, Petriglieri et al. 2021), Raman activated cell sorting (RACS) (Lee et al. 2019)), is needed to solidly establish the correlation between SCRS-resolved polyP features and its specific PAO population, which is beyond the scope of this current proof-of-concept study that aims at exploring the SCRS-enabled polyP characterization and comparison of different methods for polyP characterization in EBPR systems.

3.3.4 PolyP characterization revealed carbon-dependent PAO phenotypic change

For intracellular polyPs in PAOs, the shift of PO_2^- peak position seems to be closely related to specific molecular structure (e.g., linear, branched, or cyclic), chain length, the macromolecules and cellular compounds complexed with polyPs (e.g., proteins, nucleic acids, metals), and/or physiological/metabolic state in specific PAO cells (De Gelder et al. 2008, Majed and Gu 2010). To further confirm that the polyP structural variation and the consequent shift of Raman peak position are associated with and can reflect the PAO phylogenetic difference or phenotypic changes, we compared the distribution of PO_2^- peak position among SCRS-identified PAO cells in the same full-scale EBPR samples that were fed with either acetate or real wastewater during the P release and uptake batch tests (Figure 6). Recently, Qiu et al. (2019) characterized different EBPR-associated carbon usage profiles using activated sludge from full-scale EBPR systems, and revealed a connection between carbon utilization and PAO community composition. Similarly, the results of SCRS analysis in our study showed varied polyP peak distribution patterns when adding different carbon sources, which might be driven by PAOs with different carbon metabolisms. For the samples from the batch test fed with acetate, ~61% of PAOs exhibited the PO_2^- peak at a range between 1165 cm^{-1} and 1177 cm^{-1}, while a small group (~12%) displayed peaks at ~1154 cm^{-1} (Figure 6a). The relative abundance of *Accumulibacter* and *Tetrasphaera* PAOs in the studied

S2EBPR samples, according to the FISH results (Onnis-Hayden et al. 2020), were 4.6-7.6% and 15.3-20.2%, respectively. As suggested by Nguyen et al. (2011), *Accumulibacter* taxa have a preference for acetate, while most *Tetrasphaera* cannot utilize acetate. It implies that the major PAO phenotypic group with PO_2^- peak centered at ~1172 cm^{-1} is likely attributable to the *Accumulibacter*-like PAOs that are active in the presence of acetate for EBPR metabolism, which is also in agreement with our previous observation (Majed et al. 2009). When the carbon source in the batch test was switched to real wastewater, a varying bimodal distribution of PO_2^- peak position was detected (with a declining frequency of the group centered at ~1172 cm^{-1} and increasing frequency of the group centered at ~1160 cm^{-1} (Figure 6b and S6)). This shift in the polyP structure was likely associated with the greater involvement of other non-*Accumulibacter* PAOs, including but not limited to, *Tetrasphaera*-like PAOs with versatile capabilities of taking up and fermenting complex carbon substrates (including sugars, amino acids, etc.) (Fernando et al. 2019, Liu et al. 2019). Therefore, the shift of polyP peak distribution induced by specific carbon sources could be potentially utilized as a novel indicator reflecting PAO phenotypic changes and likely differentiating metabolically active PAO sub-populations at the single-cell level. In addition, the polyP peak distribution curves in Figure 6a and 6b overlapped to a certain extent and covered most of the spectral range from ~1150 to ~1178 cm^{-1}, indicating high PAO metabolic versatility and plasticity in the S2EBPR systems that could partially contribute to their statistically higher performance stability and functional redundancy compared to the conventional EBPR plants (Gu et al. 2019, Onnis-Hayden et al. 2020). Again, further study is needed to investigate the association of specific polyP profiles with certain PAO populations, and to evaluate if the polyP feature can serve as another complementary phenotypic indicator for phenotype-function-performance correlation establishment.

3.4 Comparison of different methods for intracellular polyP characterization

While there have been a few investigations of the intracellular polyPs in EBPR systems using different analytical tools (Alam et al. 2021, Li et al. 2019, Lindrea et al. 1998, Staal et al. 2019), an extensive comparative study of polyP characteristics (e.g., length, composition, structure) with various biological and environmental samples had not previously been performed. Therefore, not much is known about the high-resolution polyP features in EBPR samples and their relationship to specific PAO populations and associated EBPR metabolism. In this study, we, for the first time, examined and explored the application of SCRS for finer-resolution characterization of polyP in EBPR systems. We further compared SCRS with two other techniques (i.e., ^{31}P-NMR and PAGE) for the characterization of intracellular polyPs in EBPR samples (Table 3). Both ^{31}P-NMR and PAGE methods require extraction procedures, however, inefficient extraction and post-extraction procedures (i.e., lyophilization and redissolution) may result in unknown polyP fragmentation and hydrolysis (Staal et al. 2019), thus introducing uncertainties to polyP chain length results and hindering polyP quantification. In addition, neither of the above two methods can detect insoluble polyP-enriched fractions in cells (Mandala et al. 2020), thus limiting their applications for *in vivo* studies of cellular metabolism.

In comparison, the SCRS method, based on vibrations of molecules, offers a label-free and non-destructive way to obtain the *in situ* biochemical fingerprints of intracellular polymer (e.g., polyP) at a single-cell level with minimal intrusion of the cell native state (Table 3). For both lab-scale and full-scale EBPR systems (Section 3.3), the new insights of the distinctive distributions of polyP peak (e.g., PO_2^-) position (indicates intracellular polyP structure and/or assembled compounds) from SCRS analysis implies the phenotypic diversity of intracellular polyP characteristics. These phenotypic distributions are likely associated with both genotypes and system-dependent factors (e.g., carbon source, configuration), which can therefore be potential phenotypic indicators. The SCRS-based polyP characterization revealed distinct polyP length/composition for different PAO-enriched EBPR systems, potentially enabling high-

resolution phenotyping such as differentiating phenotypic subgroups or *Accumulibacter versus Tetrasphaera*.

In comparison with PAGE and ^{31}P-NMR, there are also certain limitations associated with SCRS-based polyP characterization, including: (1) As discussed in Section 3.3, the average chain length for long-chain polyP mixtures in PAO cells cannot be quantified by SCRS without a pre-established standard Raman spectral database and/or advanced data analysis tools such as machine learning; (2) The inherently weak Raman scattering may sometimes be interfered or even overlapped by the other cellular constituents and/or fluorescence background in sample matrices, leading to low spectral quality (Moudříková et al. 2016); (3) Although the reproducibility of SCRS-based measurement has been verified previously (He et al. 2017, Majed et al. 2012a) and in this study (Figure S2), long acquisition time is often required to guarantee sufficient sampling depth (i.e., number of cells to be scanned), especially for a given complex sample without prior knowledge of the diversity (He et al. 2017, Li et al. 2020). It is expected that more advanced SCRS-based techniques and bioinformatics (Wang et al. 2020) will be applied in the field of EBPR to circumvent and/or overcome the aforementioned limitations, enabling the comprehensive, enhanced, and high-throughput phenotyping of complex EBPR communities at a finer resolution.

Combined phenotypic profiling and phylogenetic composition analysis was shown to provide better insights and prediction of EBPR systems performance as indicated by the previous study (Li et al. 2018), highlighting the need and importance of phenotyping methods. The SCRS-based intracellular polyP characterization, as well as the simultaneous detection of other functionally relevant intracellular polymers (e.g., PHA, glycogen) (Majed and Gu 2010) and Raman spectral signature-based phenotypic profiling and clustering (Li et al. 2018, Wang et al. 2020), could together provide the diverse and comprehensive metabolic profiles in EBPR systems at not only population level but also down to the single-cell level. Furthermore, SCRS-based phenotypic profiles will provide more dynamic functions and traits of EBPR-relevant populations at various

metabolic stages (Fernando et al. 2019, Majed et al. 2012a, Majed and Gu 2020), which are more sensitive and responsive to changing conditions than genome-based phylogenetic analyses. In the future, it is possible to develop a Raman-based remote sensor (Gupta et al. 2020, Nagy et al. 2018) as a real-time on-line monitoring tool for capturing time-resolved polyP profiles and metabolic states of functionally relevant populations, which will potentially enhance the understanding and control of the EBPR systems. Beyond that, since polyP is a universal biopolymer in various natural and engineered ecosystems, the SCRS-based tools, as demonstrated in this case study in EBPR system, could also be applied to elucidate phenotypic or metabolic dynamics of PAOs in many other fields, such as environmental, agricultural and geoscience studies (Akbari et al. 2021).

4. Conclusions

In this study, we explored and demonstrated that SCRS can be employed for characterizing intracellular polyPs of PAOs in complex environmental samples. We further employed and compared three different methods, namely ^{31}P-NMR, PAGE, and SCRS, for polyP characterization in both lab-scale and full-scale EBPR systems, with the aim to reveal the diversity and dynamics of polyP length and structure that are indicative of phenotypic profiles of PAOs. The following conclusions are obtained:

- SCRS-enabled polyP characterization revealed distinct polyP characteristics between *Accumulibacter*- and *Tetrasphaera*-dominated EBPR reactors, as well as among different phenotypic subgroups. SCRS-based polyP characterization for both known and unknown PAOs therefore can potentially serve as a sensitive tool to assess PAO phenotypic diversity and dynamics at a finer-resolution level.
- For the first time, varied polyP features, captured and reflected by the shift and distribution of polyP signature peak positions, were revealed in the full-scale EBPR plants with varying operational conditions (i.e., configurations, carbon sources) and performance, suggesting

the possible association of polyP length characteristics with PAO phenotypic diversity, and their potential associations with EBPR performance and stability.

- In comparison to PAGE and ^{31}P-NMR that require sample extraction and pretreatment and therefore suffer from limitations such as loss or alteration of polyPs due to hydrolysis, the non-invasive SCRS approach enables "*in situ*" intracellular polyP characterization at both population and single-cell levels through clustering analysis. However, SCRS does not provide quantitative information on the average and distribution of polyP chain length as PAGE and ^{31}P-NMR can.

In the future, the SCRS-based method could be potentially extended and advanced, by employing and combining with other techniques, such as Raman–FISH, or RACS followed by downstream sequencing, for simultaneous detection and phenotypic/genotypic characterization of unculturable PAOs. Considering the diverse and important roles of polyPs in cells across taxonomic domains, SCRS-enabled technologies could be applied in a broad range of engineered or natural ecosystems for delineating the phenotypic profiles and potential roles of PAOs and deducting the genotype-phenotype-function relationships.

Supporting information avaible online. Raman spectra are available upon request.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Water Environment Research Foundation, United States (Grant No. U1R13) and National Natural Science Foundation of China (No. 52070156). HJJ gratefully acknowledges financial support from the HFSP (Human Frontier Science Program RGP0025/2016). This work was performed in part at the Harvard University Center for Nanoscale

Systems (CNS), a member of the National Nanotechnology Coordinated Infrastructure Network (NNCI), which is supported by the National Science Foundation, United States. The authors thank Doug Drury, LeAnna Risso, Bruce N. Dacko, Kevin M. Parker, Gina Serge, Heather Phillips, Howard Analla, Josh Cummings, Angela Lambrecht, and all the staff at the Clark County Water Reclamation District, Las Vegas WPCF, Atlantic Treatment Plant, Virginia Initiative Treatment Plant, Nansemond Treatment Plant, Cedar Creek WWTF, Kurt R. Segler WRF, South Cary WRF, and Westside Regional WWTP for their assistance with sample and data collection.

References

Akbari, A., Wang, Z., He, P., Wang, D., Lee, J., Han, I., Li, G. and Gu, A.Z. (2021) Unrevealed roles of polyphosphate-accumulating microorganisms. Microbial Biotechnology 14(1), 82-87.

Alam, M.M., Srinivasan, V., Mueller, A.V. and Gu, A.Z. (2021) Status and advances in technologies for phosphorus species detection and characterization in natural environment- A comprehensive review. Talanta 233, 122458.

APHA (1985) Standard methods for the examination of water and wastewater, Apha.

Barnard, J.L., Dunlap, P. and Steichen, M. (2017) Rethinking the Mechanisms of Biological Phosphorus Removal. Water Environment Research 89(11), 2043-2054.

Brdjanovic, D., Van Loosdrecht, M.C., Hooijmans, C.M., Alaerts, G.J. and Heijnen, J.J. (1998) Minimal aerobic sludge retention time in biological phosphorus removal systems. Biotechnology and bioengineering 60(3), 326-332.

Brown, M.R. and Kornberg, A. (2004) Inorganic polyphosphate in the origin and survival of species. Proceedings of the National Academy of Sciences 101(46), 16085-16087.

Burt, C. (2018) Phosphorus NMR in Biology, CRC press.

Cade-Menun, B., Liu, C., Nunlist, R. and McColl, J. (2002) Soil and litter phosphorus-31 nuclear magnetic resonance spectroscopy. Journal of Environmental Quality 31(2), 457-465.

Cade-Menun, B.J. (2005) Characterizing phosphorus in environmental and agricultural samples by 31P nuclear magnetic resonance spectroscopy. Talanta 66(2), 359-371.

Christ, J.J. (2020) Novel polyphosphate analytics for the development of biotechnological polyphosphate production, Apprimus Verlag.

Christ, J.J., Willbold, S. and Blank, L.M. (2020) Methods for the analysis of polyphosphate in the life sciences, ACS Publications.

Clark, J.E. and Wood, H.G. (1987) Preparation of standards and determination of sizes of long-chain polyphosphates by gel electrophoresis. Analytical biochemistry 161(2), 280-290.

Cordell, D., Drangert, J.-O. and White, S. (2009) The story of phosphorus: global food security and food for thought. Global environmental change 19(2), 292-305.

De Gelder, J., Willemse-Erix, D., Scholtes, M.J., Sanchez, J.I., Maquelin, K., Vandenabeele, P., De Boever, P., Puppels, G.J., Moens, L. and De Vos, P. (2008) Monitoring poly (3-hydroxybutyrate) production in Cupriavidus necator DSM 428 (H16) with Raman spectroscopy. Analytical chemistry 80(6), 2155-2160.

Fernando, E.Y., McIlroy, S.J., Nierychlo, M., Herbst, F.-A., Petriglieri, F., Schmid, M.C., Wagner, M., Nielsen, J.L. and Nielsen, P.H. (2019) Resolving the individual contribution of key microbial populations to enhanced biological phosphorus removal with Raman–FISH. The ISME Journal, 1.

Flowers, J.J., He, S., Malfatti, S., del Rio, T.G., Tringe, S.G., Hugenholtz, P. and McMahon, K.D. (2013) Comparative genomics of two 'Candidatus Accumulibacter' clades performing biological phosphorus removal. The ISME Journal 7(12), 2301-2314.

Gao, H., Mao, Y., Zhao, X., Liu, W.-T., Zhang, T. and Wells, G. (2019) Genome-centric metagenomics resolves microbial diversity and prevalent truncated denitrification pathways in a denitrifying PAO-enriched bioprocess. Water research 155, 275-287.

García-Timermans, C., Rubbens, P., Kerckhof, F.-M., Buysschaert, B., Khalenkow, D., Waegeman, W., Skirtach, A.G. and Boon, N. (2018) Label-free Raman characterization of bacteria calls for standardized procedures. Journal of Microbiological Methods 151, 69-75.

Gu, A.Z., Rahman, S.M., Eckelman, M.J. and Onnis-Hayden, A. (2016) Sustainability evaluation of nutrient removal technologies using comprehensive life cycle assessment, Water Environmental Research Foundation.

Gu, A.Z., Saunders, A., Neethling, J., Stensel, H. and Blackall, L. (2008) Functionally relevant microorganisms to enhanced biological phosphorus removal performance at full-scale wastewater

treatment plants in the United States. Water Environment Research 80(8), 688-698.

Gu, A.Z., Tooker, N.B., Onnis-Hayden, A., Wang, D., Srinivasan, V., Li, G., Takács, I. and Vargas, E. (2019) Optimization and design of a side-stream EBPR process as a sustainable approach for achieving stable and efficient phosphorus removal, Water Research Foundation.

Guo, S., Heinke, R., Stöckel, S., Rösch, P., Popp, J. and Bocklitz, T. (2018) Model transfer for Raman-spectroscopy-based bacterial classification. Journal of Raman Spectroscopy 49(4), 627-637.

Gupta, S., Huang, C.H., Singh, G.P., Park, B.S., Chua, N.-H. and Ram, R.J. (2020) Portable Raman leaf-clip sensor for rapid detection of plant stress. Scientific Reports 10(1), 20206.

He, Y., Zhang, P., Huang, S., Wang, T., Ji, Y. and Xu, J. (2017) Label-free, simultaneous quantification of starch, protein and triacylglycerol in single microalgal cells. Biotechnology for Biofuels 10(1), 275.

Hill, W.E., Benefield, L.D. and Jing, S.R. (1989) 31P-NMR spectroscopy characterization of polyphosphates in activated sludge exhibiting enhanced phosphorus removal. Water research 23(9), 1177-1181.

Hupfer, M., Glöss, S., Schmieder, P. and Grossart, H.P. (2008) Methods for detection and quantification of polyphosphate and polyphosphate accumulating microorganisms in aquatic sediments. International Review of Hydrobiology 93(1), 1-30.

Hupfer, M., Gtichter, R. and Ruegger, R.R. (1995) Polyphosphate in lake sediments: 31P NMR spectroscopy as a tool for its identification. Limnology and Oceanography 40(3), 610-617.

Jing, S., Benefield, L. and Hill, W. (1992) Observations relating to enhanced phosphorus removal in biological systems. Water research 26(2), 213-223.

Kulaev, I.S., Vagabov, V. and Kulakovskaya, T. (2005) The biochemistry of inorganic polyphosphates, John Wiley & Sons.

Lee, K.S., Palatinszky, M., Pereira, F.C., Nguyen, J., Fernandez, V.I., Mueller, A.J., Menolascina, F., Daims, H., Berry, D. and Wagner, M. (2019) An automated Raman-based platform for the sorting of live cells by functional properties. Nature microbiology 4(6), 1035-1048.

Li, G., Wu, C., Wang, D., Srinivasan, V., Kaeli, D.R., Dy, J.G. and Gu, A.Z. (2020) Machine Leaning-based Determination of Sampling Depth for Complex Environmental Systems: Case Study with Single-Cell Raman Spectroscopy Data in EBPR Systems. bioRxiv, 2020.2012.2018.423496.

Li, Y., Cope, H.A., Rahman, S.M., Li, G., Nielsen, P.H., Elfick, A. and Gu, A.Z. (2018) Toward Better

Understanding of EBPR Systems via Linking Raman-Based Phenotypic Profiling with Phylogenetic Diversity. Environmental science & technology 52(15), 8596-8606.

Li, Y., Rahman, S.M., Li, G., Fowle, W., Nielsen, P.H. and Gu, A.Z. (2019) The Composition and Implications of Polyphosphate-Metal in Enhanced Biological Phosphorus Removal Systems. Environmental science & technology 53(3), 1536-1544.

Lindrea, K., Lockwood, G. and Majone, M. (1998) The distribution and movement of polyphosphate and associated cations in sludges from NDEBPR plants in different configurations at pilot scale. Water Science and Technology 37(4-5), 555-562.

Liu, R., Hao, X., Chen, Q. and Li, J. (2019) Research advances of Tetrasphaera in enhanced biological phosphorus removal: A review. Water research, 115003.

Lockwood, G., Lindrea, K. and Seviour, R. (1990) The distribution of phosphorus and ions in the biological phosphorus removal process, pp. 195-204.

Majed, N., Chernenko, T., Diem, M. and Gu, A.Z. (2012a) Identification of functionally relevant populations in enhanced biological phosphorus removal processes based on intracellular polymers profiles and insights into the metabolic diversity and heterogeneity. Environmental science & technology 46(9), 5010-5017.

Majed, N. and Gu, A.Z. (2010) Application of Raman microscopy for simultaneous and quantitative evaluation of multiple intracellular polymers dynamics functionally relevant to enhanced biological phosphorus removal processes. Environmental science & technology 44(22), 8601-8608.

Majed, N. and Gu, A.Z. (2020) Phenotypic dynamics in polyphosphate and glycogen accumulating organisms in response to varying influent C/P ratios in EBPR systems. Science of The Total Environment 743, 140603.

Majed, N., Li, Y. and Gu, A.Z. (2012b) Advances in techniques for phosphorus analysis in biological sources. Current opinion in biotechnology 23(6), 852-859.

Majed, N., Matthäus, C., Diem, M. and Gu, A.Z. (2009) Evaluation of intracellular polyphosphate dynamics in enhanced biological phosphorus removal process using Raman microscopy. Environmental science & technology 43(14), 5436-5442.

Mandala, V.S., Loh, D.M., Shepard, S.M., Geeson, M.B., Sergeyev, I.V., Nocera, D.G., Cummins, C.C. and

Hong, M. (2020) Bacterial Phosphate Granules Contain Cyclic Polyphosphates: Evidence from 31P Solid-State NMR. Journal of the American Chemical Society 142(43), 18407-18421.

Marques, R., Ribera-Guardia, A., Santos, J., Carvalho, G., Reis, M.A., Pijuan, M. and Oehmen, A. (2018) Denitrifying capabilities of Tetrasphaera and their contribution towards nitrous oxide production in enhanced biological phosphorus removal processes. Water research 137, 262-272.

Marques, R., Santos, J., Nguyen, H., Carvalho, G., Noronha, J.P., Nielsen, P.H., Reis, M.A.M. and Oehmen, A. (2017) Metabolism and ecological niche of Tetrasphaera and Ca. Accumulibacter in enhanced biological phosphorus removal. Water research 122, 159-171.

Martín, H.G., Ivanova, N., Kunin, V., Warnecke, F., Barry, K.W., McHardy, A.C., Yeates, C., He, S., Salamov, A.A., Szeto, E., Dalin, E., Putnam, N.H., Shapiro, H.J., Pangilinan, J.L., Rigoutsos, I., Kyrpides, N.C., Blackall, L.L., McMahon, K.D. and Hugenholtz, P. (2006) Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. Nature Biotechnology 24(10), 1263-1269.

Mino, T.v., Van Loosdrecht, M. and Heijnen, J. (1998) Microbiology and biochemistry of the enhanced biological phosphate removal process. Water research 32(11), 3193-3207.

Moudříková, Š., Mojzeš, P., Zachleder, V., Pfaff, C., Behrendt, D. and Nedbal, L. (2016) Raman and fluorescence microscopy sensing energy-transducing and energy-storing structures in microalgae. Algal Research 16, 224-232.

Müssig-Zufika, M., Kornmüller, A., Merkelbach, B. and Jekel, M. (1994) Isolation and analysis of intact polyphosphate chains from activated sludges associated with biological phosphate removal. Water research 28(8), 1725-1733.

Nagy, B., Farkas, A., Borbás, E., Vass, P., Nagy, Z.K. and Marosi, G. (2018) Raman Spectroscopy for Process Analytical Technologies of Pharmaceutical Secondary Manufacturing. AAPS PharmSciTech 20(1), 1.

Neethling, J., Bakke, B., Benisch, M., Gu, A., Stephens, H., Stensel, H.D. and Moore, R. (2006) Factors influencing the reliability of enhanced biological phosphorus removal, Water Environment Research Foundation.

Nguyen, H.T.T., Le, V.Q., Hansen, A.A., Nielsen, J.L. and Nielsen, P.H. (2011) High diversity and abundance of putative polyphosphate-accumulating Tetrasphaera-related bacteria in activated sludge systems. FEMS microbiology ecology 76(2), 256-267.

Oehmen, A., Lemos, P.C., Carvalho, G., Yuan, Z., Keller, J., Blackall, L.L. and Reis, M.A. (2007) Advances in enhanced biological phosphorus removal: from micro to macro scale. Water research 41(11), 2271-2300.

Ogawa, N., DeRisi, J. and Brown, P.O. (2000) New components of a system for phosphate accumulation and polyphosphate metabolism in Saccharomyces cerevisiae revealed by genomic expression analysis. Molecular biology of the cell 11(12), 4309-4321.

Onnis-Hayden, A., Srinivasan, V., Tooker, N.B., Li, G., Wang, D., Barnard, J.L., Bott, C., Dombrowski, P., Schauer, P., Menniti, A., Shaw, A., Stinson, B., Stevens, G., Dunlap, P., Takács, I., McQuarrie, J., Phillips, H., Lambrecht, A., Analla, H., Russell, A. and Gu, A.Z. (2020) Survey of full-scale sidestream enhanced biological phosphorus removal (S2EBPR) systems and comparison with conventional EBPRs in North America: Process stability, kinetics, and microbial populations. Water Environment Research 92(3), 403-417.

Petriglieri, F., Singleton, C., Peces, M., Petersen, J.F., Nierychlo, M. and Nielsen, P.H. (2021) "Candidatus Dechloromonas phosphoritropha" and "Ca. D. phosphorivorans", novel polyphosphate accumulating organisms abundant in wastewater treatment systems. The ISME Journal.

Pilatus, U., Mayer, A. and Hildebrandt, A. (1989) Nuclear polyphosphate as a possible source of energy during the sporulation of Physarum polycephalum. Archives of biochemistry and biophysics 275(1), 215-223.

Qiu, G., Zuniga-Montanez, R., Law, Y., Thi, S.S., Nguyen, T.Q.N., Eganathan, K., Liu, X., Nielsen, P.H., Williams, R.B. and Wuertz, S. (2019) Polyphosphate-accumulating organisms in full-scale tropical wastewater treatment plants use diverse carbon sources. Water research 149, 496-510.

Robinson, N., Goss, N. and Wood, H. (1984) Polyphosphate kinase from Propionibacterium shermanii: formation of an enzymatically active insoluble complex with basic proteins and characterization of synthesized polyphosphate. Biochemistry international 8(6), 757-769.

Schönborn, C., Bauer, H.-D. and Röske, I. (2001) Stability of enhanced biological phosphorus removal and composition of polyphosphate granules. Water research 35(13), 3190-3196.

Schuler, A.J. and Jenkins, D. (2003) Enhanced biological phosphorus removal from wastewater by biomass with different phosphorus contents, part I: experimental results and comparison with metabolic models.

Water Environment Research 75(6), 485-498.

Selman, M. and Greenhalgh, S. (2009) Eutrophication: policies, actions, and strategies to address nutrient pollution. World Resources Institute (3).

Serafim, L.s.S., Lemos, P.C., Levantesi, C., Tandoi, V., Santos, H. and Reis, M.A. (2002) Methods for detection and visualization of intracellular polymers stored by polyphosphate-accumulating microorganisms. Journal of Microbiological Methods 51(1), 1-18.

Shi, H.-P. and Lee, C.-M. (2007) Use of^ 3^ 1P NMR to Investigate Polyphosphate Accumulation in Brachymonas sp. P12 with Acetate as the Sole Carbon Source under Repeated Anoxic Batch Cultivations. JOURNAL OF ENVIRONMENTAL ENGINEERING AND MANAGEMENT 17(5), 319.

Singh, J., Ripp, A., Haas, T.M., Qiu, D., Keller, M., Wender, P.A., Siegel, J.S., Baldridge, K.K. and Jessen, H.J. (2019a) Synthesis of modified nucleoside oligophosphates simplified: Fast, pure, and protecting group free. Journal of the American Chemical Society 141(38), 15013-15017.

Singh, J., Steck, N., De, D., Hofer, A., Ripp, A., Captain, I., Keller, M., Wender, P.A., Bhandari, R. and Jessen, H.J. (2019b) A phosphoramidite analogue of cyclotriphosphate enables iterative polyphosphorylations. Angewandte Chemie International Edition 58(12), 3928-3933.

Smith, S.A., Choi, S.H., Davis-Harrison, R., Huyck, J., Boettcher, J., Rienstra, C.M. and Morrissey, J.H. (2010) Polyphosphate exerts differential effects on blood clotting, depending on polymer size. Blood 116(20), 4353-4359.

Staal, L.B., Petersen, A.B., Jorgensen, C.A., Nielsen, U.G., Nielsen, P.H. and Reitzel, K. (2019) Extraction and quantification of polyphosphates in activated sludge from waste water treatment plants by (31)P NMR spectroscopy. Water Res 157, 346-355.

Stover, F.S., Bulmahn, J.A. and Gard, J.K. (1994) Polyphosphate separations and chain length characterization using minibore ion chromatography with conductivity detection. Journal of Chromatography A 688(1-2), 89-95.

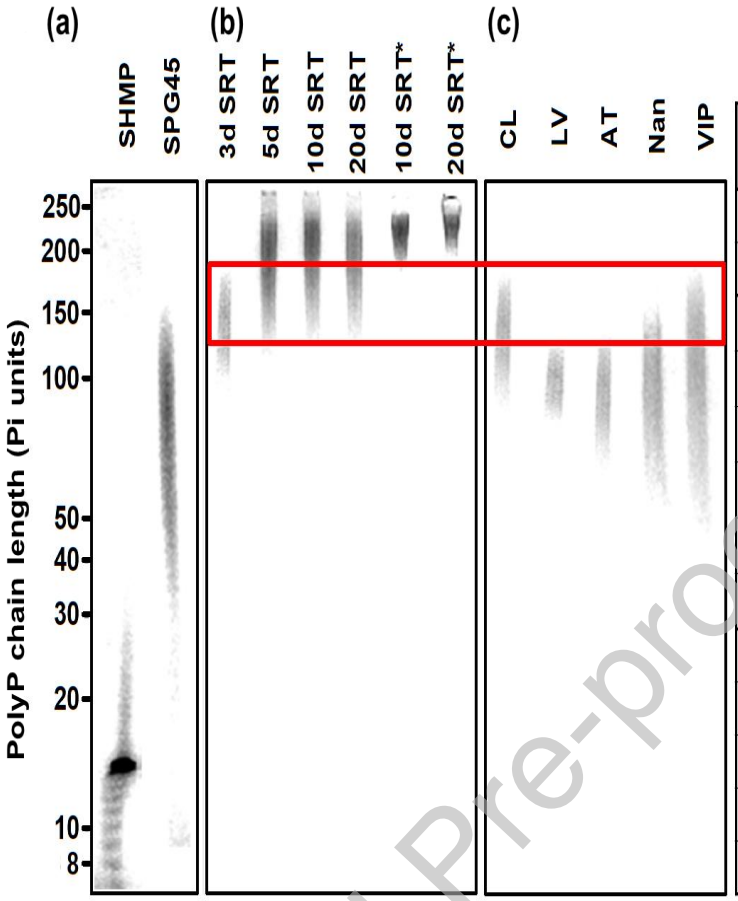
Tao, G.-J., Long, X.-Y., Tang, R., Wang, J.-Y., Fang, Z.-D., Xie, C.-X., Wang, T. and Peng, X.-H. (2020) Comparison and optimization of extraction protocol for intracellular phosphorus and its polyphosphate in enhanced biological phosphorus removal (EBPR) sludge. Science of The Total Environment 699, 134389.

Wang, D., He, P., Wang, Z., Li, G., Majed, N. and Gu, A.Z. (2020) Advances in single cell Raman spectroscopy technologies for biological and environmental applications. Current opinion in biotechnology 64, 218-229.

Wang, D., Tooker, N.B., Srinivasan, V., Li, G., Fernandez, L.A., Schauer, P., Menniti, A., Maher, C., Bott, C.B., Dombrowski, P., Barnard, J.L., Onnis-Hayden, A. and Gu, A.Z. (2019) Side-stream enhanced biological phosphorus removal (S2EBPR) process improves system performance - A full-scale comparative study. Water research, 115109.

Wentzel, M.C., Comeau, Y., Ekama, G.A., van Loosdrecht, M.C. and Brdjanovic, D. (2008) Enhanced biological phosphorus removal. Biological Wastewater Treatment-Principles, Modelling and Design, ed. Mogens Henze, 155-220.

Yuan, Z., Pratt, S. and Batstone, D.J. (2012) Phosphorus recovery from wastewater through microbial processes. Current opinion in biotechnology 23(6), 878-883.



Sample	Chain length range (Pi)	P removal efficiency (%)	Ranking
SHMP	5-30	-	-
SPG45	30-150	-	-
3 days SRT	100-170 (~135)	18	6
5 days SRT	120-250 (~185)	98	1
10 days SRT	120-250 (~185)	96	3
20 days SRT	120-250 (~185)	98	2
10 days SRT*	170-220 (~195)	21	5
20 days SRT*	180-250 (~215)	81	4
CL	80-170 (~125)	99	1
LV	70-120 (~95)	95	2
AT	60-120 (~90)	44	5
Nan	50-150 (~100)	79	4
VIP	50-170 (~110)	91	3

Figure 1. PolyP length characterization resolved on 15% polyacrylamide gels: (a) 20 µg of heterogeneous standards of synthetic sodium hexametaphosphate (SHMP) and sodium phosphate glass type 45 (SPG45); (b) Extracted polyP samples from lab-scale EBPR system of SRT of 3, 5, 10, and 20 days. The asterisk labelled SBRs of 10 and 20 days SRT were experiencing deterioration due to low dissolved oxygen supply; (c) Extracted polyP samples from full-scale EBPR facilities including: Clark County Water Reclamation District (CL), Las Vegas Water Pollution Control Facility (LV), and Atlantic Plant (AT), Nansemond Plant (Nan), and Virginia Initiative Treatment Plant (VIP) of Hampton Roads Sanitation District (HRSD). Note that SPG45 refer to the nominal mean polyP length given by the supplier, while the polyP size ladders on the left were determined either by counting (up to 30 Pi) or were generated by extrapolation from a linear standard curve of migration distance and log scale of chain lengths. The red box indicates the polyP bands within 120–170 Pi range. Values in brackets represent the polyP length at the center of the band. The EBPR performance for lab-scale and full-scale systems were ranked respectively, based on long-term stability of P removal performance (measured by the median P removal efficiency and cumulative frequency of effluent ortho-P <1 mg P/L), with lower numbers indicating better P removal performance.

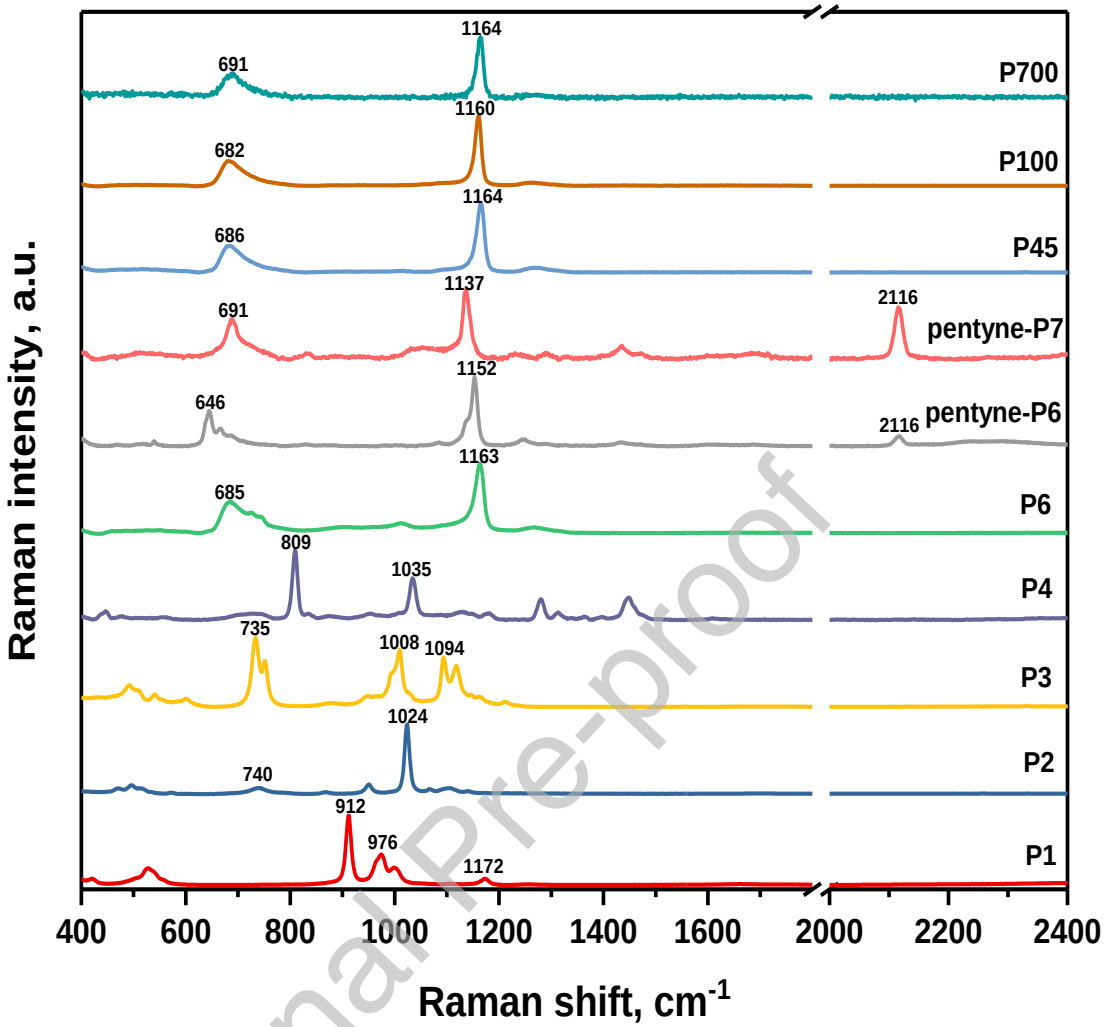


Figure 2. Raman spectra of phosphate and polyphosphate with different chain lengths. P1: monosodium phosphate (MSP); P2: tetrasodium pyrophosphate (TSPP); P3: sodium tripolyphosphate (STPP); P4: sodium tetraphosphate (STP); P6: sodium hexametaphosphate (SHMP); pentyne-P6: ammonium polyP with 6 Pi units, labeled with alkyne (pentyne) tag; pentyne-P7: ammonium polyP with 7 Pi units, labeled with alkyne (pentyne) tag; P45: sodium phosphate glass type 45; P100: polyP with nominal mean length of 100 Pi; P700: polyP with nominal mean length of 700 Pi. Distinct Raman peak of pentyne tag was observed at 2116 cm^{-1}.

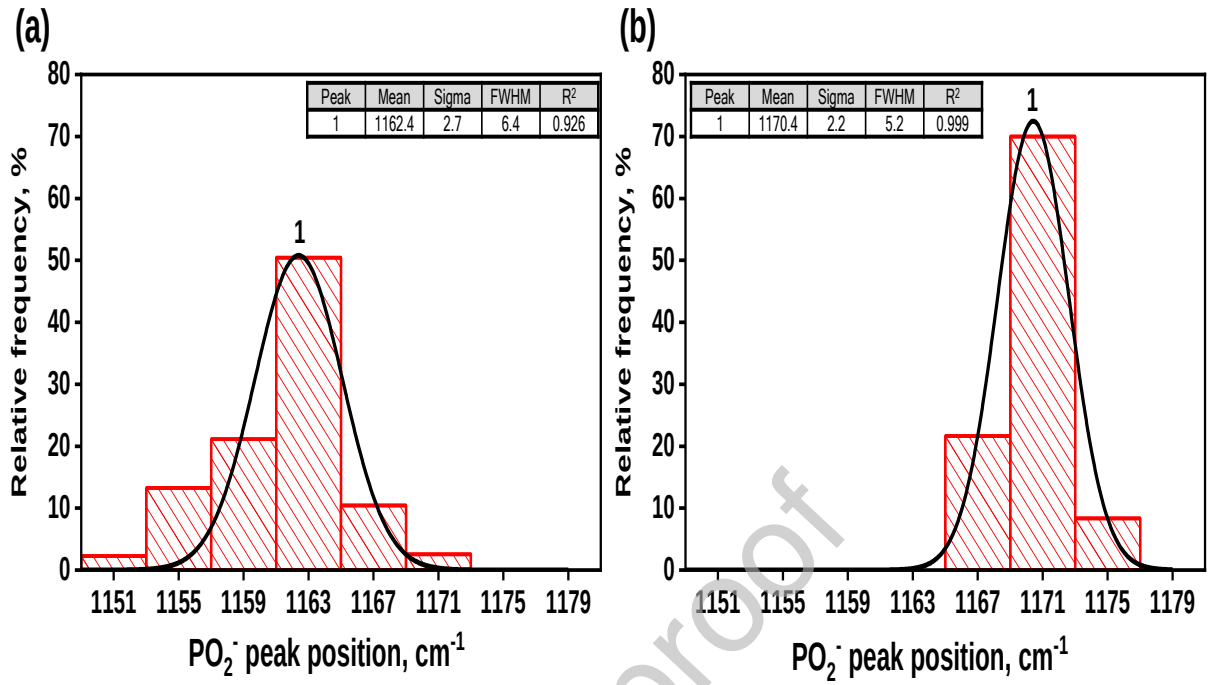


Figure 3. Relative frequency histogram and distribution curve of PO_2^- peak position of polyPs in SCRS-identified PAOs from (a) a *Tetrasphaera*-enriched culture fed with sodium casein hydrolysate (n=355) and (b) an *Accumulibacter*-like culture fed with acetate (n=60). FWHM: full width at half maximum.

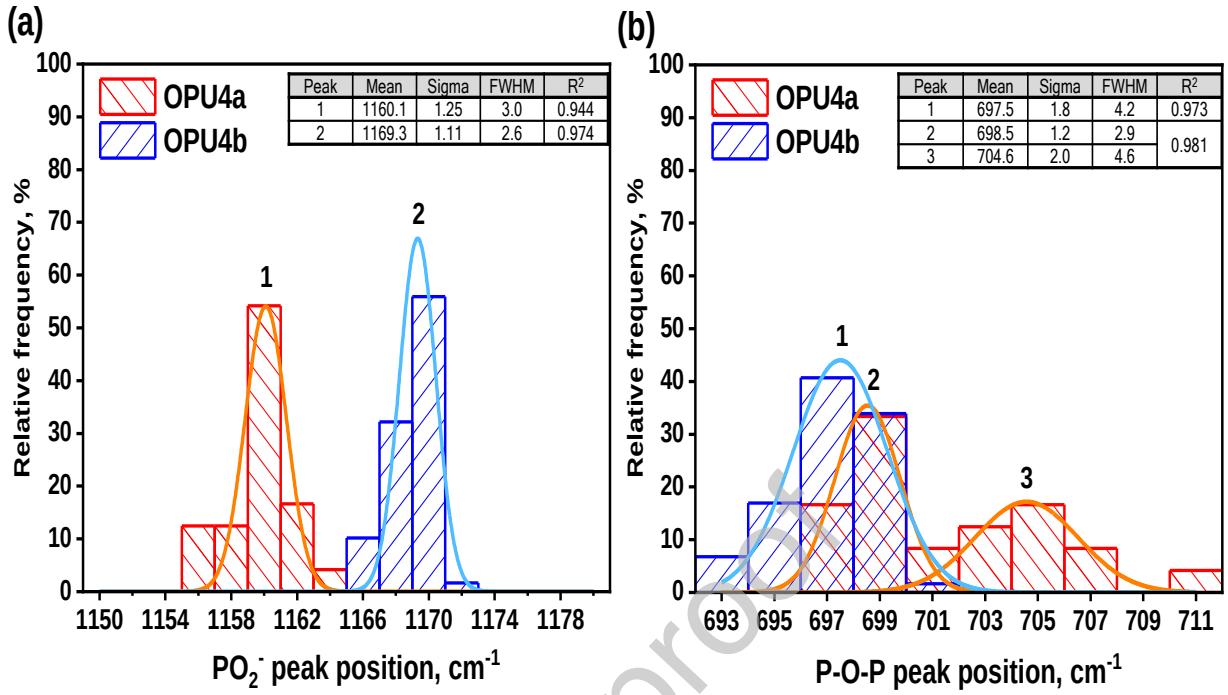


Figure 4. Relative frequency histogram and distribution curve of (a) PO_2^- and (b) P–O–P peak position of polyPs in two predominant sub-OPUs (i.e., OPU4a (n=24) and OPU4b (n=59)) from four lab-scale acetate-fed EBPR reactors. Original Raman data were extracted from previous publication (Li et al. 2018) and subjected to histogram analysis. FWHM: full width at half maximum.

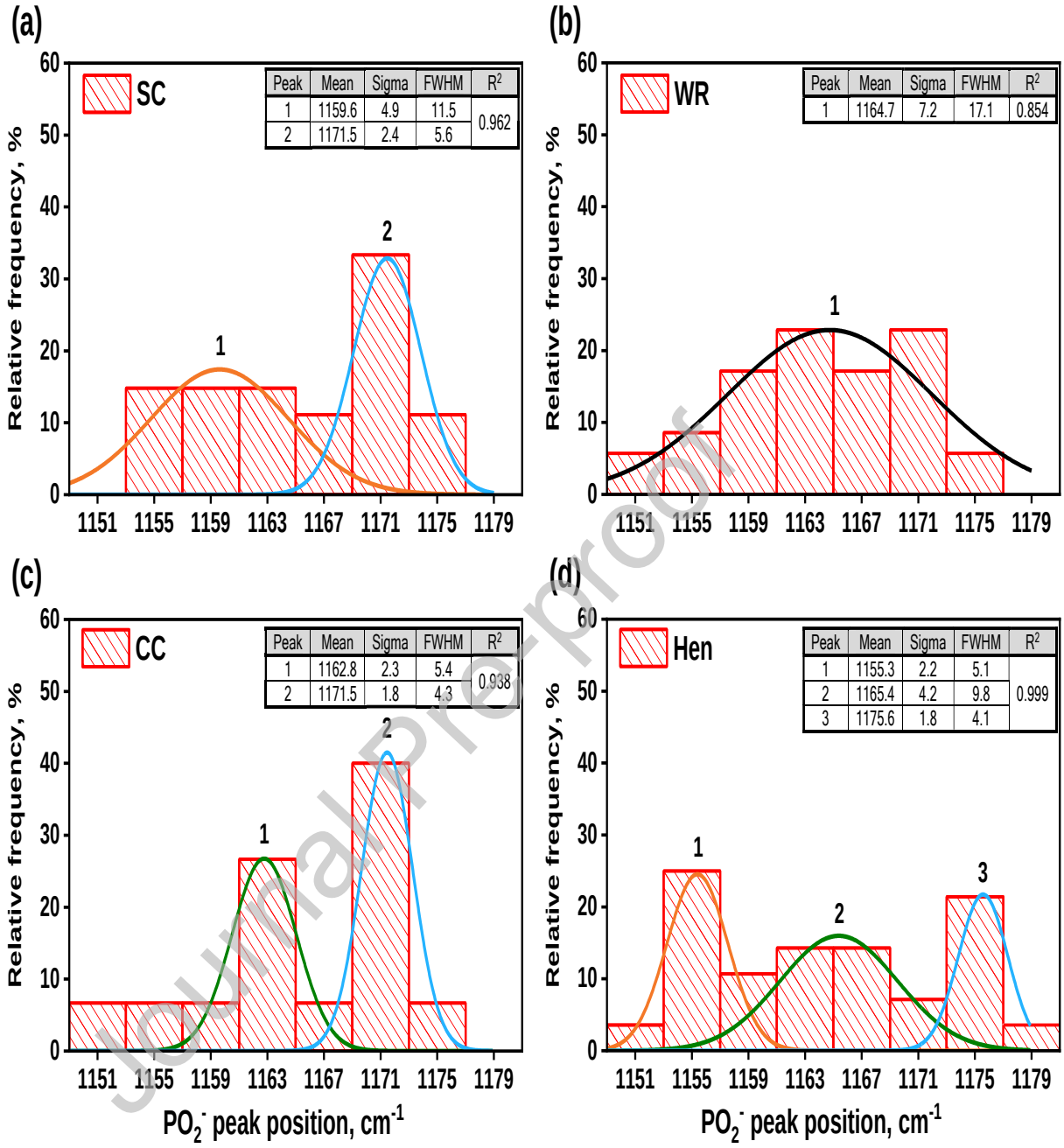


Figure 5. Relative frequency histogram and distribution curve of PO_2^- peak position of polyPs in SCRS-identified PAOs from different full-scale side-stream EBPR (S2EBPR) sludge samples. (a) Side-stream RAS fermentation (SSR) configuration at the South Cary Water Reclamation Facility (SC) (n=37), (b) Side-stream RAS fermentation with supplemental carbon addition (SSRC) configuration at the Westside Regional Wastewater Treatment Plant (WR) (n=35), (c) Side-stream mixed liquor suspended solids (MLSS) fermentation (SSM) configuration at the Cedar Creek Wastewater Treatment Facility (CC) (n=30), and (d) Unmixed in-line MLSS fermentation (UMIF) configuration at the Kurt R. Segler Water Reclamation Facility (Hen) (n=28). Note that the SCRS sampling size is much larger (>80) than the n values (Raman identified PAO cells) because the relative low PAO abundance. FWHM: full width at half maximum.

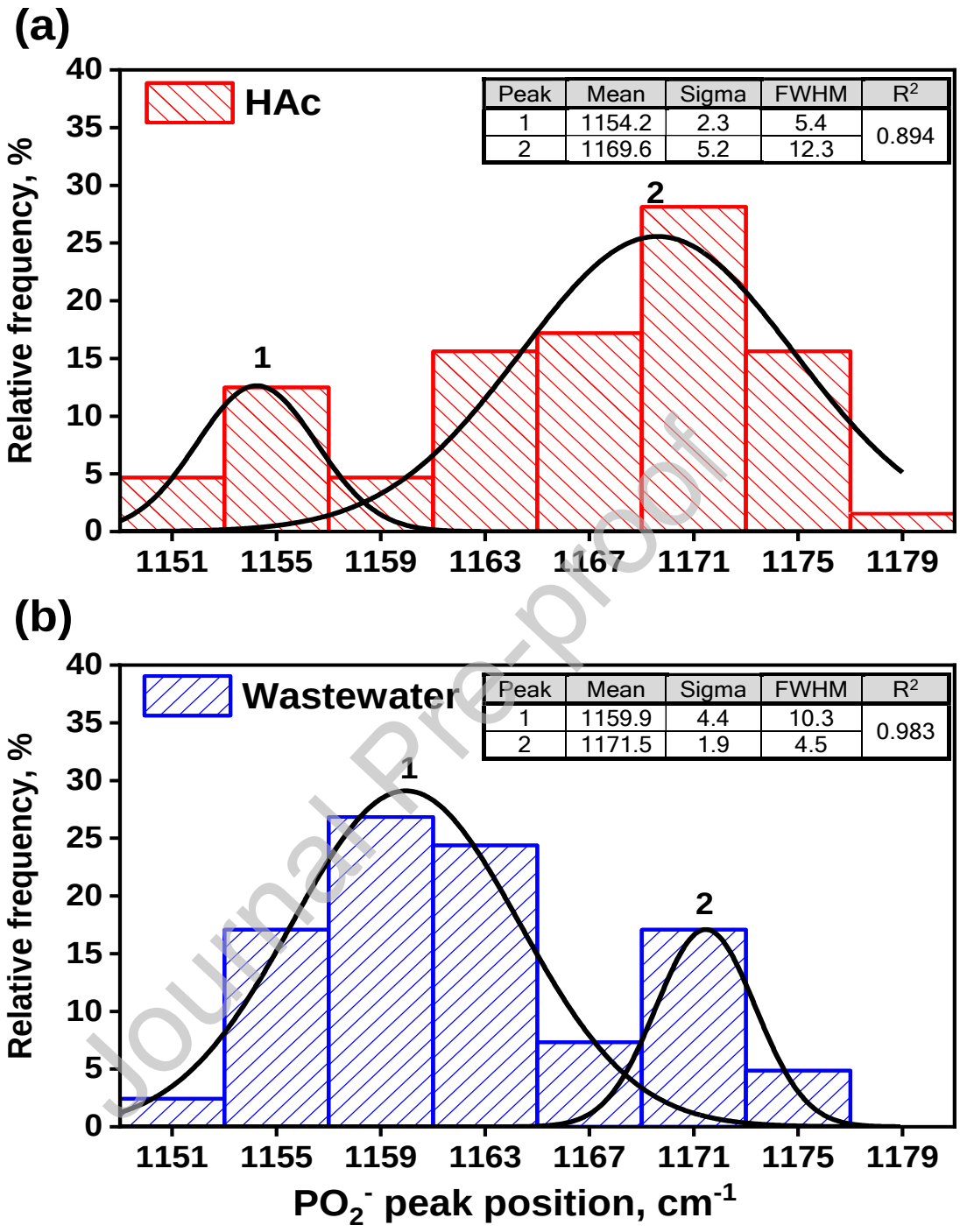


Figure 6. PolyP phenotypic changes in response to different carbon sources: relative frequency histogram of PO_2^- peak position of polyPs in SCRS-identified PAOs from the full-scale S2EBPR sludge samples from the same WWTP, which was fed with either (a) acetate (HAc) (n=64) or (b) real wastewater (n=41) during the P uptake and release tests. Note that the SCRS sampling size is much larger (>80) than the n values (Raman identified PAO cells) because the relative low PAO abundance.FWHM: full width at half maximum.

Table 1. Summarized polyP chain length estimations from ^{31}P-NMR and PAGE analysis.

Sample	P removal efficiency[a] (%)	Chain length (Pi units)	
		Mean of the average length from ^{31}P-NMR[b]	Chain length range from PAGE[c]
Sodium hexametaphosphate	–	8.7±1	5–30
Sodium phosphate glass type 45	–	54.4±2	30–150
3 days SRT SBR	18	ND	100–170
5 days SRT SBR	98	48.1±3	120–250
10 days SRT SBR	96	–	120–250
20 days SRT SBR	98	–	120–250
10 days SRT SBR*	21	31.0±7	170–220
20 days SRT SBR*	81	40.1±8	180–250

ND: not determined due to the absence of polyP-related peaks (PP1, PP2, PP3) in the ^{31}P-NMR spectrum of the sample from the SBR with 3-day SRT, which is probably due to the low PAO abundance and/or polyP storage in this most deteriorated EBPR system.

[a]: P removal efficiency at the sampling time.

[b]: mean of the average polyP lengths from duplicate ^{31}P-NMR measurements of each sample. The average polyP length from each ^{31}P-NMR measurement was calculated with Equation (1) in Supplementary Information.

[c]: with the polyacrylamide concentration (15%) used in this study, only polyP chains shorter than approximately 250 Pi could be resolved and separated.

*The asterisk labeled SBRs were experiencing deterioration due to diffuser malfunction and resulting low dissolved oxygen supply during the sampling time.

Table 2. Raman peak positions and intensities in SCRS-identified PAOs in different sub-OPUs from four lab-scale acetate-fed EBPR reactors that were operated with different SRTs. Original Raman data are extracted from Li et al. (2018) and available upon request. Values are expressed in the form of mean (range).

Sub-OPU	Abundance[a] (%)	PolyP peak position (cm^{-1})			Peak intensity (CCD counts)			
		P–O–P	PO_2^-	Peak distance	P–O–P	PO_2^-	Glycogen[b]	PHA[c]
OPU1	2	698	1166	468	2205 (1200–3496)	2888 (1614–4630)	BD	BD
OPU2d	0.6	695	1169	474	1496	2108	1031	89
OPU2e	4	693 (690–695)	1169 (1166–1172)	476	883 (561–1250)	1189 (1073–1759)	187 (0-520)	BD
OPU2g	2	693 (690–698)	1172 (1171–1173)	479	1714 (1056–2529)	2502 (1500–3824)	206 (0-411)	BD
OPU3a	2	696 (695–698)	1162 (1161–1172)	466	1681 (1538–1900)	2142 (1926–2405)	449 (308-700)	BD
OPU3b	5	699 (698–710)	1162 (1160–1163)	463	2002 (1456–2500)	2488 (2034–3105)	636 (545-781)	BD
OPU4a	14	702 (695–708)	1160 (1158–1161)	458	4359 (2300–9091)	5436 (3054–11278)	BD	BD
OPU4b	36	693 (680–694)	1169 (1167–1174)	477	4253 (2600–9200)	5841 (4150–12400)	BD	BD

BD: Below detection limit;
[a]: determined by SCRS method;
[b]: peak at ~480 cm^{-1};
[c]: peak at ~1730 cm^{-1}.

Table 3. Summary of ^{31}P-NMR, PAGE and SCRS methods for intracellular polyP characterization.

		^{31}P-NMR	PAGE	SCRS
Analysis	Sample size required	Large	Small	Minimal
	Extraction required	Yes	Yes	No
	Label/dye required	No	Yes	No
	Time required per sample[a]	~20 h extraction, ~0.5–14 h measurement[b]	~20 h extraction, ~12 h measurement	~0.5 h sample preparation, ~4 h measurement
Outcomes	P species identification	Yes	No	Yes/No[c]
	Chain length measurement	Yes, average[d]	Yes, range/distribution[e]	No
	PolyP quantification	Yes (semi- or absolute[f])	No	Yes (semi- or absolute[g])
	Measurement level	Bulk	Bulk	From population to single-cell
Advantages		Suitable for polyP quantification; potential to detect all P species present	Reliable for semi-quantitative determination of average polyP chain length and distribution in EBPR systems	Label-free; non-destructive; *in situ* measurement; potential to indicate structural/biochemical variations of polyP; simultaneous observation of multiple biopolymers; allows for phenotypic profiling and clustering; polyP can be localized in the cell
Disadvantages		Requires extraction for complicated EBPR sample; risk of polyP hydrolysis during extraction; relatively low sensitivity; time-consuming; potential to underestimate polyP length in EBPR samples; peak overlap or absence due to sample heterogeneity; cannot detect immobilized polyPs	Requires extraction for complicated EBPR sample; risk of polyP hydrolysis during extraction; relatively time-consuming	Low sensitivity to chain length changes; fluorescence interference; peak overlap or absence due to sample heterogeneity

[a]: consumed in this study;
[b]: depends on the numbers of transients for different samples (1000–25000 scans);
[c]: relies on availability and reliability of Raman spectroscopic database, sometimes not applicable due to sample heterogeneity;

[d]: may underestimate the average polyP length due to metal or cellular compounds rigidly bonded with long-chain polyPs (Hill et al., 1989), and/or underestimation of the internal P groups (PP2–PP4) for long-chain polyPs (Kulaev et al., 2005);
[e]: with the polyacrylamide concentration (15%) used in this study, only polyP chains shorter than approximately 250 Pi could be resolved and separated;
[f]: for absolute quantification, total P content in the sample should be determined by other methods (e.g., chemical digestion and spectrophotometry, inductively coupled plasma (ICP));
[g]: for absolute quantification, it should be coupled with FISH, see Fernando et al., (2019).