Environmental Microbiology (2004) doi:10.1111/j.1462-2920.2004.00687.x

The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms

Paul Wilmes[1]* and Philip L. Bond[1,2]
[1]*School of Environmental Sciences and* [2]*School of Biological Sciences, University of East Anglia, Norwich, Norfolk NR4 7TJ, UK.*

Summary

In the post-genomic era, the focus of numerous researchers has moved to studying the functional products of gene expression. In microbiology, these 'omic' approaches have largely been limited to pure cultures of microorganisms. Consequently, they do not provide information on gene expression in a complex mixture of microorganisms as found in the environment. Our method enabled the successful extraction and purification of the entire proteome from a laboratory-scale activated sludge system optimized for enhanced biological phosphorus removal, its separation by two-dimensional polyacrylamide gel electrophoresis and the mapping of this metaproteome. Highly expressed protein spots were excised and identified using quadrupole time-of-flight mass spectrometry with *de novo* peptide sequencing. The proteins isolated were putatively identified as an outer membrane protein (porin), an acetyl coenzyme A acetyltransferase and a protein component of an ABC-type branched-chain amino acid transport system. These proteins possibly stem from the dominant and uncultured *Rhodocyclus*-type polyphosphate-accumulating organism in the activated sludge. We propose the term 'metaproteomics' for the large-scale characterization of the entire protein complement of environmental microbiota at a given point in time.

Introduction

The diversity and abundance of prokaryotic microorganisms in the biosphere is extensive. They make the earth hospitable by managing its biogeochemistry, driving its nutrient cycles and breaking down both natural and anthropogenic waste products (Rodriguez-Valera, 2004). Unfortunately, the study of *in situ* prokaryotic communities is riddled with obstacles and results in one of the greatest knowledge gaps in the biological and environmental sciences. Partly, this is due to enrichment bias, in which standard culturing techniques often select for 'weed' organisms, leading to a distorted understanding of microbial community structure and function. Recent estimates predict that, in some environments, 99% of prokaryotes are, as yet, unculturable (Schloss and Handelsman, 2003). The advent of molecular techniques, 16S rRNA gene libraries in particular, and their application to a variety of different microbial communities have reinforced the need for developing new tools for the study of prokaryotic diversity and function.

Currently, metagenomics, in which large fragments of DNA are directly cloned from a particular environment, is being hailed as a path forward (Rodriguez-Valera, 2004). Although the metagenomic approach allows relevant clones to be sequenced and screened for the expression of useful gene products, it is limited by DNA extraction bias, the need for specific screening criteria and the requirement for large sequencing capacities.

Recently, whole-genome sequences of an acid mine drainage biofilm (Tyson *et al*., 2004) and of microbial populations from the Sargasso Sea (Venter *et al*., 2004) have been reported. As more genome and environmental sequencing projects are performed, the application of functional environmental 'omic' approaches is becoming feasible. Knowledge of complete metagenomes will allow the construction of DNA microarrays to monitor global community gene expression patterns within mixed communities of microorganisms. Consequently, both transcriptomic and proteomic investigations will provide insights into microbial activities in the environment (Zhou and Thompson, 2002).

Biological wastewater treatment processes represent complex microbial systems that are of particular biotechnological and environmental interest. Effluent from wastewater treatment plants (WWTPs) is a major contributor of phosphorus (P) to receiving surface waters and may lead to eutrophication. Eutrophication events are predicted to get worse in response to global climate change and increases in population density, unless the P load is

Received 19 April, 2004; revised 15 June, 2004; accepted 15 June, 2004. *For correspondence. E-mail P.Wilmes@uea.ac.uk; Tel. (+44) 1603 593633; Fax (+44) 1603 591327.

reduced. Activated sludge wastewater treatment systems with alternating anaerobic and aerobic phases operate globally and are designed for enhanced biological phosphorus removal (EBPR) (Seviour *et al*., 2003). In EBPR, microorganisms are conditioned to accumulate large amounts of P in the form of intracellular polyphosphate (polyP), resulting in the desired P removal (Blackall *et al*., 2002). The microorganisms involved are termed polyphosphate-accumulating organisms (PAOs), reflecting their ability to incorporate polyP in excess of the levels normally required to satisfy the metabolic demand for growth (>15% cell dry weight) (Bond and Rees, 1999). Over the past 30 years, successive studies of pure cultures from activated sludge have failed to match isolates with the expected PAO phenotype (Blackall *et al*., 2002). However, the construction of 16S rRNA gene libraries in conjunction with fluorescence *in situ* hybridization (FISH) reveal that a predominant PAO in laboratory-scale EBPR systems is an uncultured relative of *Rhodocyclus* spp. (Bond *et al*., 1999). Further recent studies have confirmed this (Hesselmann *et al*., 1999; Crocetti *et al*., 2000; McMahon *et al*., 2002). Investigations using combinations of FISH and microautoradiography (^{33}P-labelled polyP; Lee *et al*., 2001) and DAPI (Zilles *et al*., 2001) convincingly demonstrate that *Rhodocyclus*-type organisms accumulate polyphosphate in EBPR full-scale systems. Several empirical biochemical models have been proposed since the accidental discovery of EBPR more than 40 years ago (Mino *et al*., 1998; Seviour *et al*., 2003). In addition, there is a clear interest in gaining enzymological details of the EBPR process (Hesselmann *et al*., 2000). Despite the long-standing interest in the microbiology of EBPR, numerous biochemical and metabolic details remain elusive, and a fresh incentive to elucidate these is required.

Ogunseitan (1993) first reported the use of single-dimensional SDS-PAGE for analysing the protein products of gene expression from different environmental samples, including wastewater. More recently, single-dimensional SDS-PAGE has been used to analyse protein expression in different EBPR WWTPs (Ehlers and Cloete, 1999a,b). These studies failed to reveal proteins specific for P removal. However, the investigations were carried out on activated sludge from full-scale WWTPs, which contain a large microbial diversity, performing an array of complex biochemical transformations. It is likely that these studies were hampered by the complexity of the samples and the limitations of single-dimensional SDS-PAGE. The present focus on high-performance EBPR laboratory-scale reactors, as model systems, is providing the most useful insights into EBPR microbial ecology and function (Crocetti *et al*., 2000; Filipe *et al*., 2001).

So far, the application of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and downstream analyses has been limited to pure cultures of different microorganisms (Washburn and Yates, 2000). However, Rodriguez-Valera (2004) proposes the use of the 'metaproteome' to detect genes most abundantly expressed in environmental samples. The term 'metaproteome' is derived from 'metagenome', which reflects the compound genome of the whole microbiota found in nature (Rondon *et al*., 2000). As microbial populations form highly variable communities, neither the metaproteome nor the metagenome of a microbial community can be regarded as a static entity; both are dynamic functions of the mixture of microorganisms at a given point in time. Hence, metaproteomic investigations, as well as metagenomics, merely portray 'snap-shots' of microbial populations. Proteomics is defined as the 'large-scale characterization of the entire protein complement of a cell line, tissue or organism' (Graves and Haystead, 2002). Consequently, we propose the term 'metaproteomics' for the large-scale characterization of the entire protein complement of environmental microbiota at a given point in time. Our study describes the application of metaproteomics to a mixed community of microorganisms from a laboratory-scale sequencing batch reactor optimized for EBPR and enriched for PAOs. The method developed allowed the extraction and purification of the entire mixed community proteome for 2D-PAGE separation. In addition, we demonstrate the use of mass spectrometry for the identification of proteins from the environmental sample. To our knowledge, this is the first account of a metaproteomic investigation based on 2D-PAGE.

Results and discussion

Sequencing batch reactor performance

Enhanced biological phosphorus removal was achieved in the operation of the laboratory-scale activated sludge reactor and maintained at a constant level for ≈ 3 months. The P removal performance of the sequencing batch reactor (SBR) was monitored by determination of the P levels in the media and in the supernatant at the end of the aerobic phase. Every second week, cycle studies were carried out, for which samples were collected from the reactor every 15 min during one discrete cycle. During the period of operation, 120 mg l^{-1} P was regularly removed during the aerobic phase, and no phosphate was detected in the effluent (Fig. 1). The soluble phosphate P (PO_4-P) concentration in the synthetic wastewater averaged 22.3 mg l^{-1} during the period of stable operation. Mixed liquor suspended solids (MLSS) levels averaged 4.0 g l^{-1}, and total P was 314.5 mg l^{-1}. Consequently, the biomass P content was 7.7%. These parameters were similar to other high-performance EBPR laboratory-scale activated sludge systems (Bond *et al*., 1999; Crocetti *et al*., 2000; McMahon *et al*., 2002). The EBPR sludge in the SBR

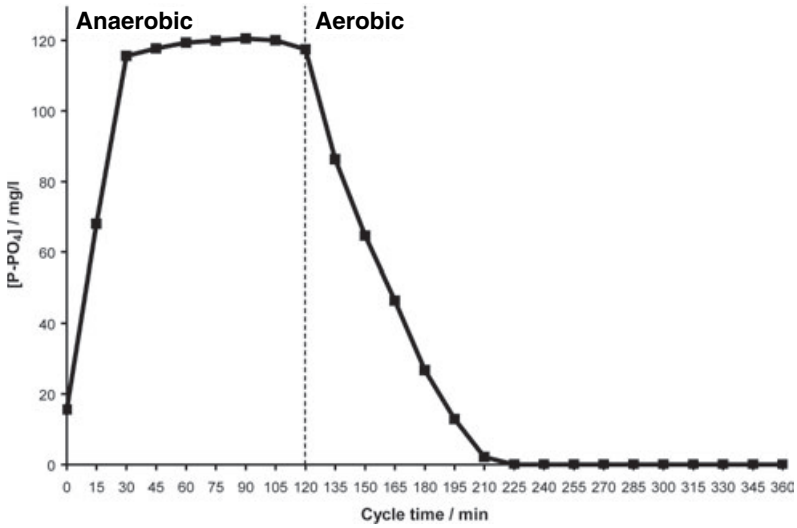

Fig. 1. Extracellular phosphate P concentrations in the SBR during one discrete cycle study.

provided samples for subsequent analyses by FISH and 2D-PAGE.

FISH analysis

Most of the cells that stained positively with DAPI were identified as bacteria with the EUB338 probe (88.3%). From the proportions of cells stained with DAPI and hybridized with the specific probes, it was concluded that the EBPR sludge community was dominated by β-Proteobacteria (probe BET42a detected 73.6% of EUB338-binding cells) and by the *Rhodocyclus*-type PAO organisms (probe PAO651 detected 54.1% of EUB338-binding cells) (Fig. 2). These results were similar to those found previously by other investigators (Hesselmann *et al*., 1999; Crocetti *et al*., 2000; McMahon *et al*., 2002). The laboratory-scale high-performance EBPR systems select for PAOs that are related to *Rhodocyclus* (*R. tenuis* and *R. purpureus*) and *Propionibacter pelophilus* (Crocetti *et al*., 2000). These are assumed to be important to EBPR but are as yet uncultured. There is interest in analysing these high-performance EBPR systems, that are enriched for PAOs, in an effort to obtain insights into the important microbial transformations. Thus, to circumvent problems of studying unculturable microorganisms, culture-independent analyses are required to investigate mixed culture systems such as the metaproteomic approach developed in this study.

2D-PAGE

The extraction and preparation of the entire SBR proteome required trials of different protocols for various steps of the procedure. This involved trialling combinations for (i) various wash steps of different NaCl concentrations to remove excess exopolysaccharides; (ii) lysis buffers [Genomic Solutions' standard buffers, buffers used for single-dimension SDS-PAGE (Ogunseitan, 1993), urea–thiourea–CHAPS (UTCHAPS) buffers]; (iii) cell lysis methods (sonication, freeze–thaw, French press); (iv) protein precipitation methods [acetone, ammonium sulfate, trichloroacetic acid (TCA)]; and (v) resuspension buffers (UTCHAPS buffers, combinations of low- and high-stringency buffers). These trials resulted in a preparation method comprising: (i) a 0.9% (w/v) NaCl wash solution; (ii) a UTCHAPS lysis buffer; (iii) French press lysis; (iv) TCA precipitation [followed by washes in 50 mM Tris and 80% (v/v) acetone]; and (v) resuspension in the low- and high-stringency buffers (see *Experimental procedures*). Protein quantification of the crude protein extracts (before TCA precipitation) enabled the accurate loading of protein to the first-dimension immobilized pH gradient (IPG) strips and was a crucial step for producing quality 2D-PAGE separations. The developed method resulted in well-resolved and reproducible 2D-PAGE gels.

After a complete cycle study and determination of good EBPR performance (Fig. 1), activated sludge samples were taken at the end of the anaerobic (t = 120 min) and aerobic (t = 330 min) phases. In addition, samples were taken at time points of highest activities of phosphate transformations, as determined in the cycle study (t = 20 min and t = 160 min). During these phases, the microbial transformations included the anaerobic degradation of polyP and release as phosphate, followed by phosphate uptake and storage as polyP in the aerobic phase (Fig. 1).

A series of 2D-PAGE separations of a single activated sludge sample (t = 120 min) was performed to determine the optimal pH gradient range for first-dimension isoelectric focusing (IEF) (Fig. 3). Low-stringency fraction (LSF) proteins resolved well on pH 3–10 immobilized pH gradient (IPG) strips (Fig. 3A). However, as most proteins

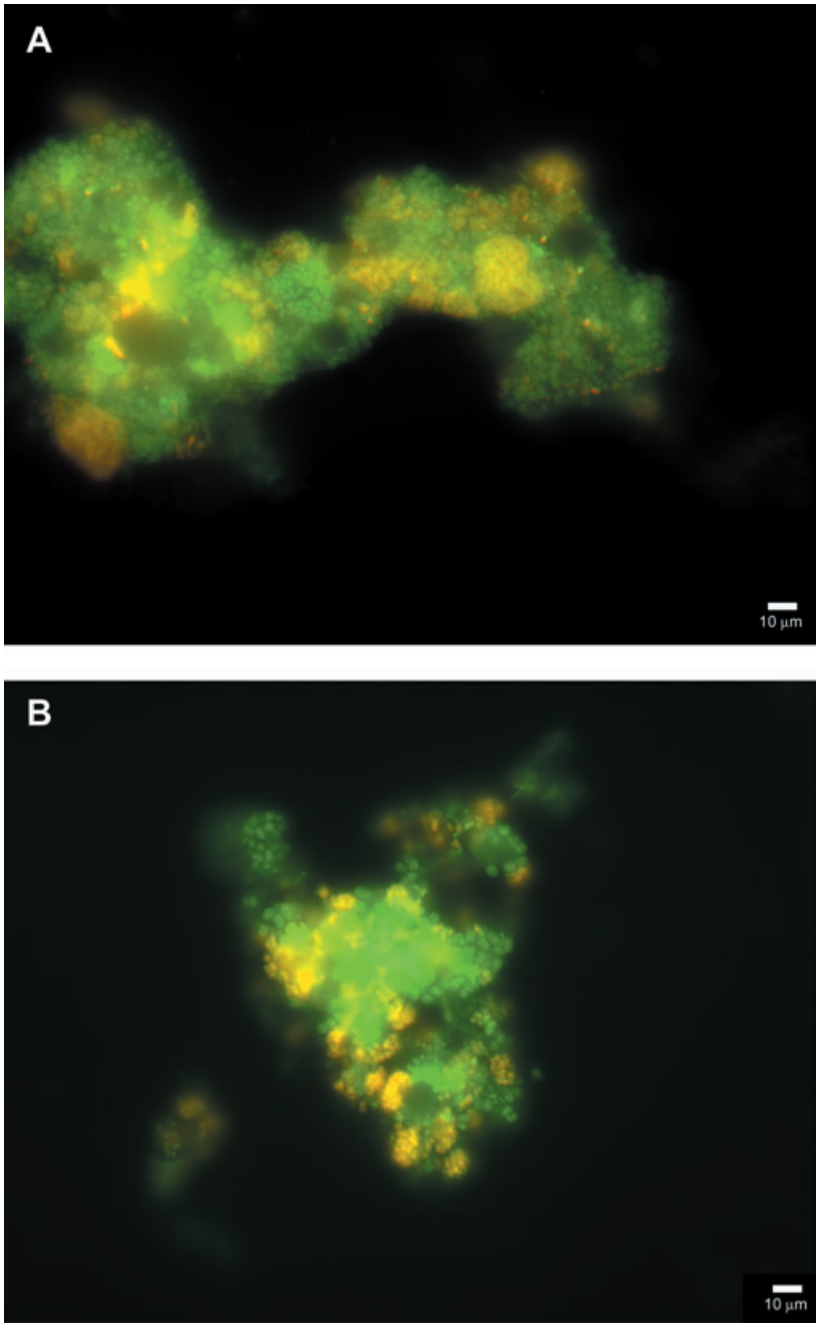

Fig. 2. Representative FISH micrographs of the SBR activated sludge. Cells detected with probe EUB338 only are green (A and B). Cells detected with both EUB338 and BET42a probes (A) and cells detected with both EUB338 and PAO651 probes (B) are yellow-orange. Images taken under the different excitation wavelengths for CY3 and FITC were combined using Adobe PHOTOSHOP.

appeared in the pH 4–7 region, IPG strips of that pH range were used. Improved protein separation was obtained for both LSF and high-stringency fraction (HSF) preparations (Fig. 3B and C). The use of low- and high-stringency buffers allowed the prefractionation of the complex proteome into two fractions based on protein solubility. Similar techniques are commonly used to prefractionate eukaryotic proteomes before 2D-PAGE analysis (Peck *et al.*, 2001). By comparing the resulting gel images (Fig. 3B and C compared with 3D and E), it is apparent that the fractionation resulted in two distinct protein fractions of the activated sludge proteome.

Sequencing batch reactor (SBR) sample proteins (t = 120 min) were resolved further for LSF and HSF preparations using IPG strips in the range pH 4–5 (Fig. 3D and E). Pairwise comparison of the gel images produced (Fig. 3A with 3B, 3B with 3D, and 3C with 3E) demonstrated that corresponding protein separation patterns were obtained using the different pH range IPG strips. This indicated that the method had produced reproducible

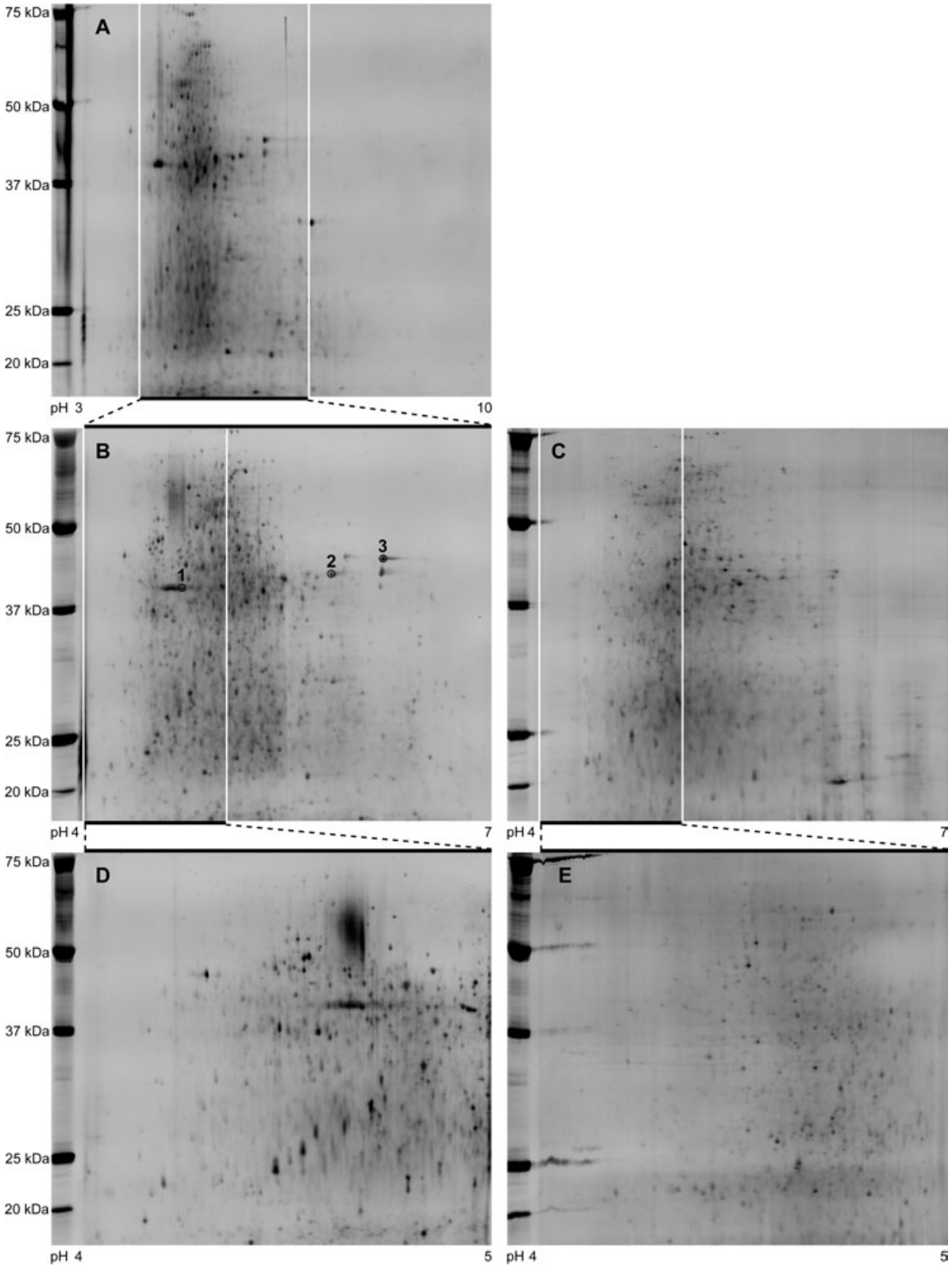

Fig. 3. 2D-PAGE gels of an identical activated sludge protein sample ($t = 120$ min) analysed for different pH regions: (A) LSF, pH 3–10; (B) LSF, pH 4–7; (C) HSF, pH 4–7; (D) LSF, pH 4–5; (E) HSF, pH 4–5. Highlighted areas (white bars) refer to narrower pH regions, which correspond to the 2D-PAGE gels depicted in the row below. Highlighted spots in (B) (1, 2 and 3) correspond to the protein spots excised and analysed using Q-ToF MS.

proteome maps of all fractions. The improved resolution obtained in the narrower pH ranges allows better protein visualization in areas of potential interest (proteins with known isoelectric point and known molecular mass).

For visual detection of the large amount of proteins present in the SBR samples, it was necessary to load maximum amounts of protein onto the IPG strips. This caused problems when attempting to load enough protein on the narrow-range IPG strips. However, that limitation could be overcome by preparative liquid-phase isoelectric focusing to remove the superfluous proteome fractions.

Comparative gels of LSFs from the different sampling time points ($t = 20$ min, 120 min, 160 min and 330 min) in the pH 4–7 region were generated (Fig. 4). These exhibited excellent reproducibility of protein separation. A total of 728 corresponding spots was identified on all four gels using the PROTEOMWEAVER software. A correlation matrix based on the spot intensities was constructed using the Pearson's product moment correlation coefficient (r). The mean r-value of the correlation matrix was 0.827 with a standard deviation of 0.014, demonstrating strong correlations between all four gels and, hence, implying low gel-to-gel variability. Regions of potential differential protein expression were analysed, and a single area of interest is highlighted on all four gels (Fig. 4A–D). Spot no. 1 was 27% less intense in the $t = 330$ min gel than in the $t = 120$ min gel, indicating potential downregulation of that particular gene during the aerobic phase (Fig. 4C and D). Furthermore, the one spot on gels A, B and C (Fig. 4) was resolved into two spots on the $t = 330$ min gel (Fig. 4D),

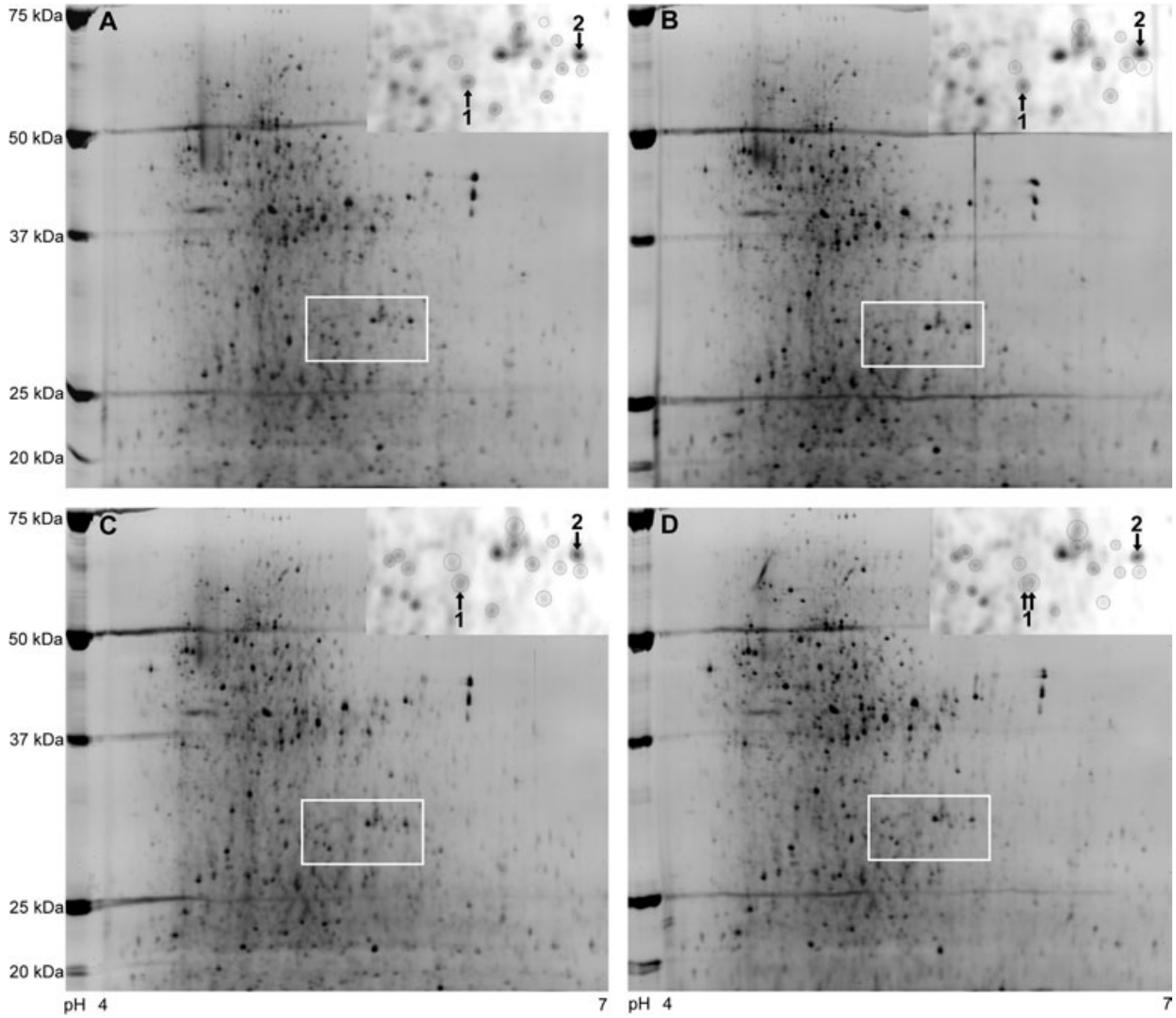

Fig. 4. 2D-PAGE gels of LSF SBR protein samples taken at different time points: (A) $t = 20$ min; (B) $t = 120$ min; (C) $t = 160$ min; (D) $t = 330$ min. Highlighted areas refer to the areas of interest analysed using PROTEOMWEAVER (inserts in top righthand corner of each gel image). Gel images were normalized using PROTEOMWEAVER.

which may indicate possible protein modification between the aerobic and anaerobic phases. Spot no. 2 was 41% less intense in the $t = 330$ min gel compared with the $t = 120$ min gel (Fig. 4D and B). As can be seen in Fig. 4, the decreases in intensity were gradual for both spots throughout the aerobic phase. Further in-depth analysis of these gels is required for the detection of differential protein expression between the anaerobic and aerobic phases of the activated sludge to discover functional metabolic constituents of EBPR.

Matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-ToF MS)

Nine strong protein spots were excised from an EZBlue-stained gel, digested and analysed by MALDI-ToF MS. All nine spots revealed good mass spectra with clear and distinct peaks, indicating that pure proteins had been analysed. Database searches of the resulting peptide mass fingerprints revealed bacterial proteins. However, the taxonomic affiliation of the different proteins was varied, and statistically relevant hits were not obtained.

Although good mass spectra data were obtained from the activated sludge proteins, poor database matches probably resulted because the genomes of the organisms involved in EBPR are as yet unsequenced. Thus, this study demonstrates that database matches of peptide mass fingerprints generated from environmental samples using MALDI-ToF MS do not produce useful information.

Quadrupole (Q)-ToF MS

Q-ToF MS of three excised spots (Fig. 3B) was performed to assess the possibility of protein identification through *de novo* peptide sequencing. As found with MALDI-ToF MS, the Q-ToF MS spectra exhibited clear and distinct peaks. *De novo* peptide sequencing was carried out on three peptide fragments for each protein spot (Table 1).

The peptide sequence of the fragment (a) for protein spot no. 1 (Table 1) resulted in a high database match (*E*-value of 0.082). The fragments (b) and (c) had very low database matches (high *E*-values). However, fragments (a) and (b) revealed identical protein matches (COG3203: outer membrane protein, porin) from two β-Proteobacteria, *Burkholderia fungorum* and *Ralstonia metallidurans* respectively. These are close relatives of *Rhodocyclus* spp. In addition, fragments (a) and (b) exhibited a sequence overlap. A BLASTP search of the combined sequence (QQFVGLQSDKLGTVALGR) revealed an identical protein match to fragment (a) with an even lower *E*-value (0.022). Hence, the excised protein is an outer membrane protein (porin) (COG3203) possibly from the *Rhodocyclus*-type PAO.

The closest match for protein spot no. 2 (Table 1) was obtained for fragment (b) (*E*-value of 2×10^{-4}). This was the best statistical match of all the fragments obtained from the Q-ToF MS analyses. The next best match for this protein was fragment (a) (*E*-value of 0.48), and fragment (c) resulted in a low match (*E*-value of 9.1). Both protein matches for fragments (a) and (b) were identical, acetyl coenzyme A acetyltransferase, although they differed in their taxonomic affiliations [*Staphylococcus aureus* for fragment (a) and *Pseudomonas aeruginosa* for fragment (b)]. *Pseudomonas aeruginosa* (highest match) is a γ-Proteobacterium, therefore related to the β-Proteobacteria. Acetyl coenzyme A acetyltranferase catalyses the condensation of two molecules of acetyl-CoA to form acetoacetyl-CoA, a precursor in the synthesis of polyhydroxyalkanoates (PHAs). The conversion of acetyl-CoA to PHAs is a predominant feature of representative metabolic models for EBPR (Mino *et al.*, 1998). The protein isolated from the high-performance EBPR system, acetyl coenzyme A acetyltranferase, may be actively involved in the anaerobic production of PHAs.

For protein spot no. 3 (Table 1), the sequences of peptide fragments (a) and (c) showed high database matches (low *E*-values). Fragment (b) had a high *E*-value and did not match any known bacterial peptide sequence. Fragments (a) and (c) matched highly conserved sequences from two β-Proteobacteria, *Dechloromonas aromatica* RCB and *Ralstonia solanacearum*, the closest relatives of

Table 1. Results of the Q-ToF MS and BLASTP searches of the fragment peptide sequences.

Spot no.	Fragment	Peptide sequence	Most significant alignment	*E*[a]
1	(a)	FVGLQSDKLGTVALGR	COG3203: outer membrane protein, porin (*Burkholderia fungorum*)	0.082
	(b)	QQFVGLQSDK	COG3203: outer membrane protein, porin (*Ralstonia metallidurans*)	72
	(c)	SQSAALGWTTAMSQR	Hypothetical protein Rxyl206301 (*Rubrobacter xylanophilus* DSM 9941)	39
2	(a)	QLPETTPAFTVNR	Acetyl coenzyme A acetyltransferase (*Staphylococcus aureus*)	0.48
	(b)	SVTAGNASGVNDGQAVVLTR	Probable acyl-CoA thiolase (*Pseudomonas aeruginosa*)	2e-04
	(c)	AAFAEELAPFEVNR	Prolyl-tRNA synthetase (*Streptococcus agalactiae* NEM316)	9.1
3	(a)	SVDPVKYLPEVGK	Hypothetical protein Daro127401 (*Dechloromonas aromatica* RCB)	2
	(b)	AQFTTGDGGCSPELLK	Conserved hypothetical protein (*Trypanosoma brucei*)	6.5
	(c)	LAGEASEGVLCTQAGLPADQLP	Putative amino acid transport signal peptide protein (*Ralstonia solanacearum*)	0.12

a. *E*-value describes the random background noise that exists for matches between sequences, see http://www.ncbi.nlm.nih.gov/BLAST/

Rhodocylus spp. with available genome sequences. A BLASTP search of the complete amino acid sequence from the two conserved proteins revealed identical matches (COG0683, LivK, ABC-type branched-chain amino acid transport systems, periplasmic component; amino acid transport and metabolism). Consequently, the excised protein is possibly a cellular amino acid transporter in the *Rhodocyclus*-type PAO.

All the excised protein spots were strongly expressed on the 2D-PAGE gels, suggesting putative 'housekeeping' roles or roles pertinent to EBPR biochemistry. Consequently, these few protein sequencing results are consistent with general microbial activity and perhaps the EBPR-specific activity of the activated sludge sample. In addition, proteins well resolved on these 2D-PAGE gels were cytoplasmic or periplasmic proteins in Gram-negative bacteria (β-Proteobacteria). This was consistent with the microbial community composition of this EBPR sludge as determined by FISH.

Conclusion

The method described using 2D-PAGE and Q-ToF MS with *de novo* peptide sequencing enables the isolation and positive identification of proteins from complex mixtures of microorganisms. The diversity and complexity of the activated sludge proteome reinforces the idea that a mixed community of microorganisms may be regarded as a kind of 'metaorganism' (Rodriguez-Valera, 2004). The metaorganism's proteome is analogous to that of a complex eukaryotic organism. This is reflected in the need for fractionation of the proteome into low- and high-stringency fractions in order to resolve the entire protein complement.

Metaproteomics will allow the mining of mixed community proteomes to identify novel proteins from previously uncultured organisms that are of particular environmental interest. Our study focuses on EBPR, for which pure cultures and comprehensive metabolic details are elusive. Further application of this metaproteomic approach will provide insights into EBPR biochemistry. It follows that the isolation and identification of key metabolic enzymes from the activated sludge mixed community has enormous biotechnological potential for future wastewater treatment processes.

This study demonstrates the feasibility of carrying out proteomics on a mixed community and will lead to further studies on a variety of different environmental samples. The availability of complete metagenome sequences will diminish the need for Q-ToF MS as peptide mass fingerprints generated with MALDI-ToF MS should result in statistically valid database matches.

Experimental procedures

Operation of the SBR

Sequencing batch reactor (SBR) set-up. A laboratory-scale SBR was operated for activated sludge EBPR. It was inoculated with activated sludge from the Whitlingham Wastewater Treatment Plant, Norwich, UK. The SBR had a working volume of 2 l and was operated on a 6 h cycle consisting of a 2 h anaerobic phase, 3.5 h aerobic phase and 0.5 h settling/decanting phase. The aerobic phase was maintained at pH 7 ± 0.2, and the SBR was kept at a constant temperature of 20°C. One litre of treated water was decanted at the end of each cycle followed by the introduction of 1 l of fresh synthetic wastewater, resulting in a hydraulic retention time of 12 h. Biomass was wasted during the aeration period to maintain a sludge retention time of 7 days. The synthetic wastewater comprised 950 ml of bulk medium and 50 ml of concentrated medium for each cycle. The bulk medium contained (per l) 90 mg of $MgSO_4.7H_2O$, 160 mg of $MgCl_2.6H_2O$, 42 mg of $CaCl.2H_2O$, 64 mg of K_2HPO_4, 60 mg of KH_2PO_4 and 0.3 ml of nutrient solution. The nutrient solution was made up according to the method of Bond *et al.* (1999). The concentrated medium contained (per l) 10 mg of allyl-thiourea, 14 g of $NaAc.3H_2O$, 2.44 g of peptone, 0.4 g of yeast extract and 1 g of NH_4Cl. All the solutions were made up with MilliQ (Millipore) water, adjusted to pH 7 and autoclaved.

Chemical analyses. Orthophosphate analyses were carried out on filtered (Sartorius Minisart 0.2 μm filter) sludge samples using the appropriate Merck Spectroquant kit and analysed on a Philips PU8730 fixed bandwidth spectrophotometer. Mixed liquor suspended solids (MLSS) and total P of the mixed liquor were determined by standard methods (Clesceri *et al.*, 1998), and the P was then quantified as orthophosphate as described above.

FISH

Sampling, cell fixation and hybridization. The sampling and cell fixation was carried out as described previously (Bond *et al.*, 1999). Dispersion of the sludge flocs was not necessary. Oligonucleotide probes (Table 2) were obtained from MWG Biotech. Probes BET42a and PAO651 were labelled with the sulfoindocyanide dye CY3, whereas EUB338 was labelled with fluorescein isothiocyanate (FITC). uBET was

Table 2. Oligonucleotide probe sequences and target sites for FISH.

Probe	Sequence (5′-3′)	rRNA target site	Specificity	% Formamide	Reference
EUB338	GCTGCCTCCCGTAGGAGT	16S, 338–355	Bacteria	35	Amman *et al.* (1996)
BET42a	GCCTTCCCACTTCGTTT	23S, 1027–1043	β-Proteobacteria	35	Manz *et al.* (1992)
uBET	GCCTTCCCACATCGTTT	23S, 1027–1043	Competitor for BET42a	35	Falkentoft *et al.* (2002)
PAO651	CCCTCTGCCAAACTCCAG	16S, 651–688	*Rhodocyclus*-type PAO	35	Crocetti *et al.* (2000)

unlabelled. The *in situ* hybridization was carried out according to the method of Bond *et al.* (1999). After hybridization, the slides were rinsed with MilliQ water, air dried and mounted with Vectashield mounting medium for fluorescence (Vector Laboratories) containing 4′,6′-diamidino-2-phenylindole (DAPI). The cells were observed under a Zeiss Axioplan 2 epifluorescent microscope.

Image processing. Images of each field of view were taken at the different excitation wavelengths using a Zeiss Axio Cam Mrm camera and processed using the Zeiss Axioplan software package. For each field of view, three images corresponding to the DAPI, CY3 and FITC signals were captured and imported into Adobe PHOTOSHOP. The proportions of fluorescent pixels of the CY3-labelled probes (BET42a and PAO651) in relation to DAPI and EUB338 were calculated. In addition, the proportion of EUB338 in relation to DAPI was calculated. Nine fields of view were analysed for DAPI, FITC and CY3 excitation wavelengths at two different magnifications (×630 and ×1000) for each of the two sets of probe combinations.

2D-PAGE

Protein extraction and purification. All the solutions were made up with analytical grade water (Fisher). Cells from the activated sludge were harvested by centrifugation at 4500 r.p.m. for 20 min. The supernatant was discarded, and the cells were washed in 50 ml of a 0.9% (w/v) NaCl solution to remove most of the exopolysaccharides. The cells were centrifuged as above, and the resulting pellet was resuspended in 50 mM Tris-HCl, pH 7. After centrifugation, 10 ml of urea–thiourea–CHAPS (UTCHAPS) sample buffer was added to the cell pellet. The sample buffer consisted of 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 10 mM Tris–1 mM EDTA, 50 mM dithiothreitol (DTT), 25 μM Pefabloc SC and 2 mM Pefabloc protector (Roche). Cells were resuspended by pipette and pulse-vortexing, then placed on ice for 2 h with mixing every 15 min. Cells were lysed by three passes through a French press at 1000 p.s.i. The suspension was then centrifuged at 13000 r.p.m. for 30 min to remove cell debris, and the supernatant was retained as the crude protein extract. Protein was quantified using a 2-D Quant system (Amersham Biosciences) and precipitated in 10% (w/v) TCA, incubated on ice for 30 min and spun at 13 000 r.p.m. for 15 min. The protein pellet was then washed twice in 50 mM Tris-HCl, pH 8, twice in 50 mM Tris and twice in 80% (v/v) ice-cold acetone. The dried protein pellet was resuspended in 400 μl of low-stringency buffer, comprising 9 M urea, 1% (v/v) Triton X-100 (Sigma), 1% (v/v) IPG buffer with the appropriate pH range (Amersham Biosciences), 0.5% (w/v) DTT and a trace of bromophenol blue, by vortex mixing for 2 h. After resuspension in the low-stringency buffer, the sample was centrifuged at 13 000 r.p.m. for 45 min, and the resulting supernatant constituted the low-stringency fraction (LSF). The remaining protein pellet was resuspended in 400 μl of high-stringency buffer, consisting of 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 1% (v/v) IPG buffer with the appropriate pH range (Amersham Biosciences), 2% (w/v) DTT and a trace of bromophenol blue and placed on a vortex shaker for 2 h. The sample was then centrifuged at 13 000 r.p.m. for 45 min, and the supernatant was conserved forming the high-stringency fraction (HSF).

2D-PAGE separation. 18 cm IPG strips (Immobiline DryStrips with the appropriate pH range; Amersham Biosciences) were individually rehydrated in the presence of either LSF or HSF in an Immobiline DryStrip reswelling tray (Amersham Biosciences) for 16 h. For first-dimension separation, the strips were placed in a pHaser IEF unit (Genomics Solutions), covered with Plusone DryStrip cover fluid (Amersham Biosciences) and focused for 100 000 volt–hours with a maximum voltage of 3500 V. The strips were equilibrated according to the method of Hesketh *et al.* (2002) and applied to 14% (v/v) Duracryl (Genomic Solutions) gels. Precision Plus Protein Standard Plugs (Bio-Rad) were applied to the gels according to the manufacturer's instructions. Second-dimension separation was carried out at 500 V. Gels for spot excision were stained with EZBlue (Sigma) according to the manufacturer's instructions. Gels for image analysis were stained with silver nitrate (Blum *et al.*, 1987), scanned using a ProXPRESS proteomic imaging system (Perkin-Elmer Life Sciences) and analysed using PROTEOMWEAVER, version 1.3 (Definiens).

Protein identification using mass spectrometry

MALDI-ToF MS. Protein spots of interest were excised from EZBlue-stained gels. The samples were digested and analysed using MALDI-ToF MS (Bruker Reflex III) (Hesketh *et al.*, 2002). Searches of the peptide mass fingerprint data against databases were performed with the MASCOT search engine at http://www.matrixscience.com.

Q-ToF. Digested protein samples were analysed further on a Micromass Q-ToF-2 mass spectrometer. *De novo* peptide sequencing was carried out on three fragments for each protein sample. Peptide sequences were searched using BLASTP at http://www.ncbi.nlm.nih.gov/BLAST using the default settings of the 'Search for short, nearly exact matches' function.

Acknowledgements

The authors would like to acknowledge the help of Craig Baker-Austin and Dr Mark Dopson with the initial 2D-PAGE separations. Helpful comments and advice were received from Prof. Andrew Johnston at University of East Anglia (UEA), from Dr Francis Mullholland at the Institute of Food Research (IFR), Norwich, UK, from Dr Scott Peck at the Sainsbury Laboratory [John Innes Centre (JIC), Norwich, UK] and from Dr Andrew Hesketh at the JIC. MALDI-ToF MS and Q-ToF MS were carried out by Dr Andrew Bottrill and Dr Mike Naldrett at the IFR and JIC Joint Proteomics Facility. Paul Wilmes was funded by a School of Environmental Sciences Studentship in conjunction with a British Council Chevening Scholarship.

References

Amman, R., Snaidr, J., Wagner, M., Ludwig, W., and Schleifer, K.H. (1996) Combination of 16S rRNA-targeted oligo-

nucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56:** 1919–1925.

Blackall, L.L., Crocetti, G.R., Saunders, A.M., and Bond, P.L. (2002) A review and update of the microbiology of enhanced biological phosphorus removal in wastewater treatment plants. *Antonie van Leeuwenhoek* **81:** 681–691.

Blum, H., Beier, H., and Gross, H.J. (1987) Improved silver staining of plant-proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8:** 93–99.

Bond, P.L., and Rees, G.N. (1999) Microbiological aspects of phosphorus removal in activated sludge systems. In *The Microbiology of Activated Sludge*. Seviour, R.J., and Blackall, L.L. (eds). London, UK: Kluwer Academic Publishers, pp. 227–256.

Bond, P.L., Erhart, R., Wagner, M., Keller, J., and Blackall, L.L. (1999) Identification of some of the major groups of bacteria in efficient and nonefficient biological phosphorus removal activated sludge systems. *Appl Environ Microbiol* **61:** 4077–4084.

Clesceri, L.S., Greenberg, A.E., and Eaton, A.D. (1998) *Standard Methods for the Examination of Water and Wastewater*. Washington, DC, USA: American Public Health Association, American Water Works Association and Water Environment Federation.

Crocetti, G.R., Hugenholtz, P., Bond, P.L., Schuler, A., Keller, J., Jenkins, D., and Blackall, L.L. (2000) Identification of polyphosphate-accumulating organisms and design of 16S rRNA-directed probes for their detection and quantification. *Appl Environ Microbiol* **66:** 1175–1182.

Ehlers, M.M., and Cloete, T.E. (1999a) Protein profiles of phosphorus- and nitrate-removing activated sludge systems. *Water SA (South Africa)* **25:** 351–356.

Ehlers, M.M., and Cloete, T.E. (1999b) Comparing the protein profiles of 21 different activated sludge systems after SDS-PAGE. *Water Res* **33:** 1181–1186.

Falkentoft, C.M., Müller, E., Arnz, P., Harremoës, P., Mosbaek, H., Wilderer, P.A., and Wuertz, S. (2002) Population changes in a biofilm reactor for phosphorus removal as evidenced by the use of FISH. *Water Res* **36:** 491–500.

Filipe, C.D.M., Daigger, G.T., and Grady, C.P.L., Jr (2001) Stoichiometry and kinetics of acetate uptake under anaerobic conditions by an enriched culture of phosphorus-accumulating organisms at different pHs. *Biotechnol Bioeng* **76:** 32–43.

Graves, P., and Haystead, T.A.J. (2002) Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* **66:** 39–63.

Hesketh, A.R., Chandra, G., Shaw, A.D., Rowland, J.J., Kell, D.B., Bibb, M.J., and Chater, K.F. (2002) Primary and secondary metabolism, and post-translational protein modifications, as portrayed by proteomic analysis of *Streptomyces coelicolor*. *Mol Microbiol* **46:** 917–932.

Hesselmann, R.P.X., Werlen, C., Hahn, D., van der Meer, J.R., and Zehnder, A.J.B. (1999) Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. *Syst Appl Microbiol* **22:** 454–465.

Hesselmann, R.P.X., Von Rummell, R., Resnick, S.M., Hany, R., and Zehnder, A.J.B. (2000) Anaerobic metabolism of bacteria performing enhanced biological phosphate removal. *Water Res* **34:** 3487–3494.

Lee, N., Jansen, J., Aspegren, H., Dircks, K., Henze, M., Nielsen, P.H., *et al.* (2001) Population dynamics and *in situ* physiology of phosphorus-accumulating bacteria in wastewater treatment plants with enhanced biological phosphorus removal operated with and without nitrogen removal. In *Microorganisms in Activated Sludge and Biofilm Processes*. Tandoi, V., Passino, R., and Blundo, C.M. (eds). Rome, Italy: IWA, pp. 114–122.

McMahon, K.D., Dojka, M.A., Pace, N.R., Jenkins, D., and Keasling, J.D. (2002) Polyphosphate kinase from activated sludge performing enhanced biological phosphorus removal. *Appl Environ Microbiol* **68:** 4971–4978.

Manz, W., Amman, R., Ludwig, W., Wagner, M., and Schleifer, K.H. (1992) Phylogenetic oligonucleotide probes for the major subclasses of Proteobacteria: problems and solutions. *Syst Appl Microbiol* **15:** 593–600.

Mino, T., van Loosdrecht, M.C.M., and Heijnen, J.J. (1998) Microbiology and biochemistry of the enhanced biological phosphate removal process. *Water Res* **32:** 3193–3207.

Ogunseitan, O.A. (1993) Direct extraction of proteins from environmental samples. *J Microbiol Methods* **17:** 273–281.

Peck, S.C., Nühse, T., Iglesias, A., Hess, D., Meins, F., and Boller, T. (2001) Directed proteomics identifies plant-specific protein rapidly phosphorylated in response to bacterial and fungal elicitors in a range of plant species. *Plant Cell* **13:** 1467–1475.

Rodriguez-Valera, F. (2004) Environmental genomics, the big picture? *FEMS Microbiol Lett* **231:** 153–158.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., *et al.* (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* **66:** 2541–2547.

Schloss, P.D., and Handelsman, J. (2003) Biotechnological prospects from metagenomics. *Curr Opin Biotechnol* **14:** 303–310.

Seviour, R.J., Mino, T., and Onuki, M. (2003) The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiol Rev* **27:** 99–127.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428:** 37–43.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.

Washburn, M.P., and Yates, J.R. (2000) Analysis of the microbial proteome. *Curr Opin Microbiol* **3:** 292–297.

Zhou, J., and Thompson, D.K. (2002) Challenges in applying microarrays to environmental studies. *Curr Opin Biotechnol* **13:** 204–207.

Zilles, J.L., Hung, C.-H., and Noguera, D.R. (2001) Presence of *Rhodocyclus* in a full-scale wastewater treatment plant and their participation in enhanced biological phosphorus removal. In *Microorganisms in Activated Sludge*. Tandoi, V., Passino, R., and Blundo, C.M. (eds). Rome, Italy: IWA, pp. 75–81.