REVIEW

Proteomics and metabolomics: The molecular make-up of toxic aromatic pollutant bioremediation

Om V. Singh

Department of Pediatrics, The Johns Hopkins School of Medicine, Baltimore, MD, USA

Microbial-mediated attenuation of toxic aromatic pollutants offers great potential for the restoration of contaminated environments in an ecologically acceptable manner. However, incomplete biological information regarding the regulation of growth and metabolism in many microbial communities restricts progress in the site-specific mineralization process. In the postgenomic era, recent advances in MS have allowed enormous progress in proteomics and elucidated many complex biological interactions. These research forefronts are now expanding toward the analysis of low-molecular-weight primary and secondary metabolites analysis, *i.e.*, metabolomics. The advent of 2-DE in conjunction with MS offers a promising approach to address the molecular mechanisms of bioremediation. The two fields of proteomics and metabolomics have thus far worked separately to identify proteins and primary and secondary metabolites during bioremediation. A simultaneous study combining functional proteomics and metabolomics, *i.e.*, proteometabolomics would create a system-wide approach to studying site-specific microorganisms during active mineralization processes. This article deals with advances in environmental proteomics and metabolomics and advocates the simultaneous study of both technologies to implement cell-free bioremediation.

Received: March 21, 2006
Revised: May 16, 2006
Accepted: June 8, 2006

Keywords:
Bioremediation / Environmental pollutants / Mass spectrometry / Metabolomics / Nuclear magnetic resonance

1 Introduction

Environmental pollutants in soil are a major concern worldwide; many toxic, mutagenic, and carcinogenic aromatic pollutants are known to pose serious threats to human health. Contaminated soil can be treated by various physical and chemical methods, such as thermal desorption, soil washing, incineration, and landfilling, but these treatments do not effectively restore natural flora and fauna. Bioremediation – the use of microorganisms to remove toxic pollutants from the environment – is a highly promising technology that is environmentally friendly, safe, and effective [1–3]. Many different site-specific microorganisms are capable of carrying out bioremediation reactions, and they have already been used at sites previously contaminated with polycyclic aromatic hydrocarbons (PAHs), nitroaromatic compounds (NACs), chlorinated organics, *etc.* [3–5]. However, certain pollutants have been unusually recalcitrant, as either microbes are unable to completely mineralize them or their sister metabolites accumulate in the environment [6] and create health hazards. Possible solutions to this dilemma are exploring or engineering new catabolic pathways and studying regulatory control of primary and secondary metabolites to produce effective bioremediation reactions.

In practice, our knowledge of *in situ* bacteria's adaptability to environmental stress and the regulation of catabolic and respiratory genes is far from complete. However, the study of genes, protein expression patterns, and metabolite

Correspondence: Dr. Om V. Singh, Park 316, 600 N. Wolfe Street, Baltimore, MD 21209, USA
E-mail: osingh1@jhmi.edu
Fax: +1-410-955-1030

Abbreviations: AE, alcohol polyethoxylate; **DIGE**, differential in-gel electrophoresis; **HR**, high-resolution; **MAS**, magic angle spinning; **MudPIT**, multidimensional protein identification platform; **PAHs**, polycyclic aromatic hydrocarbons; **PPI**, protein–protein interaction

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

formation continues to make critical contributions. New sophisticated techniques in medical science make it possible to explore global protein expression (proteomics) and low-molecular-weight metabolite expression (metabolomics) in environmental remediation. By applying proteome- and metabolome-based techniques to environmental samples, it is now possible to develop models that can predict microbial activities under various bioremediation strategies.

Proteomics is one of the newest emerging technologies in functional genomics, and it has taken center stage in current scientific thinking. Efforts to sequence the genomes of several bacteria involved in bioremediation (*Colwellia psychrerythraea*, *Dehalococcoides ethanogens*, *Deinococcus radiodurans*, *Geobacter sulfurreducens*, *Pseudomonas putida*, *etc*.) are the first step in the daunting task of exploring the myriad physiological functions of highly complex and adaptable organisms possessing a variety of odd genes. The classical "one gene – one function" and/or "one gene – one enzyme" relationships no longer apply [7], since each gene may encode more than one protein based on differential splicing and translation factors [8]. In addition, PTMs in natural and genetically engineered organisms may generate millions of distinguishable functional entities at the protein level. Therefore, there are many good reasons to study the global proteomic profile during *ex situ* and *in situ* bioremediation.

Metabolomics involves a nontargeted, holistic analysis of the set of metabolites produced by cellular proteins in response to various environmental stimuli. One particular advantage of metabolomics is that, when combined with multivariate data analysis (MVDA) tools like principal component discriminant analysis and partial least squares, it allows us to monitor changes in an organism as it is exposed to environmental pollutants. This dynamic information allows us to follow degradative pathways and track their intermediates and responses during mineralization; it is particularly valuable for studying the complete *in situ* mineralization process, which evolves dynamically and has indirect effects on many heterogeneous cell populations.

The key question is whether we will be able to exploit the complexity of the proteome and metabolome. Rather than summarizing all of the existing literature, this article will focus on current questions and developments in environmental proteomics and metabolomics for bioremediation purposes. Various sources of bioinformatics will also be discussed, emphasizing the interpretation of vast amounts of proteome and metabolome data that can help us to understand the regulatory mechanisms of microbial physiology in bioremediation.

2 Proteomics

The first step in any proteomics-based approach is protein separation. There are numerous ways to achieve this, each with its own pros and cons; the choice of method is usually based on the particular experiment and the available instruments and expertise. Based on the typical workflow in Fig. 1, discussion will focus on the most common methodologies for environmental proteomics.

2.1 2-DE for protein and enzyme expression profiling

The mainstay of protein expression profiling for the last 30 years has been 2-DE, where proteins are separated by the complementary techniques of IEF in the first dimension and SDS-PAGE size fractionation in the second dimension. IPGs and advanced bioinformatics have remarkably improved the reproducibility and comparability of 2-DE gels. Many studies have shown that 2-DE can detect differences between proteins from normal and contaminated environments [9–15].

Based on system biology, any physiological state of an organism leads to a differential expression in genes and proteins that induces a chain of enzymatic reactions to catalyze the mineralization of contaminants in and around the cell. 2-DE can explore the metabolic pathways of such organisms by identifying the key proteins and enzymes [16–20]. The identification and characterization of such proteins and enzymes that can be removed and/or extracted from cell are in demand for environmental remediation. The cellular lysate of *Mycobacterium* sp. was profiled on a 2-DE gel and an 81-kDa protein similar to catalase-peroxidase was shown to be significantly induced under pyrene exposure [9]. Later, two ring-hydroxylating dioxygenases, Pdo1 and Pdo2, were induced during pyrene catabolism by another *Mycobacterium* sp. [10]. A bacterial membrane proteome of *Acinetobacter radioresistens* S13 during aromatic exposure was profiled and a variety of outer membrane proteins, including OmpA-like proteins, trimeric porin, and glycosyltransferases, were up-regulated [11]. Eighty unique proteins were identified by 2-DE/MS from *P. putida* KT 2440 cultured in the presence of six different organic compounds [19]. Based on 2-DE analysis, an enzymatic pathway has been proposed for phthalate metabolism in *Rhodococcus* sp. strain TFB [20]. A modeling of the 3-D protein homology of a psychrophilic bacterium *C. psychrerythraea* 34H reveals changes in proteome composition that may enhance enzymes' effectiveness at low temperature [21]. Other than 2-DE, MS-based proteomics approaches were used to profile a variety of specific enzymes, such as epoxide hydrolases, peroxisomal antioxidant enzymes, and sarcosine oxidase (SOX) associated with marine pollutant exposure [18]. 2-DE-based identification of these proteins and enzymes during bioremediation makes them candidates as catalysts for cell-free remediation that can direct the use of biomolecules, extracted or removed from cell and/or cell-free systems for field bioremediation.

Many highly sensitive fluorescent stains have been successfully employed in differential in-gel electrophoresis (DIGE) to accurately quantify protein expression in 2-DE [22]. In DIGE, protein samples covalently labeled with red and/or green fluorescent dyes are mixed with a blue dye-labeled internal standard protein [23]. A 2-DE separation of these labeled protein samples helps avoid ambiguities in

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

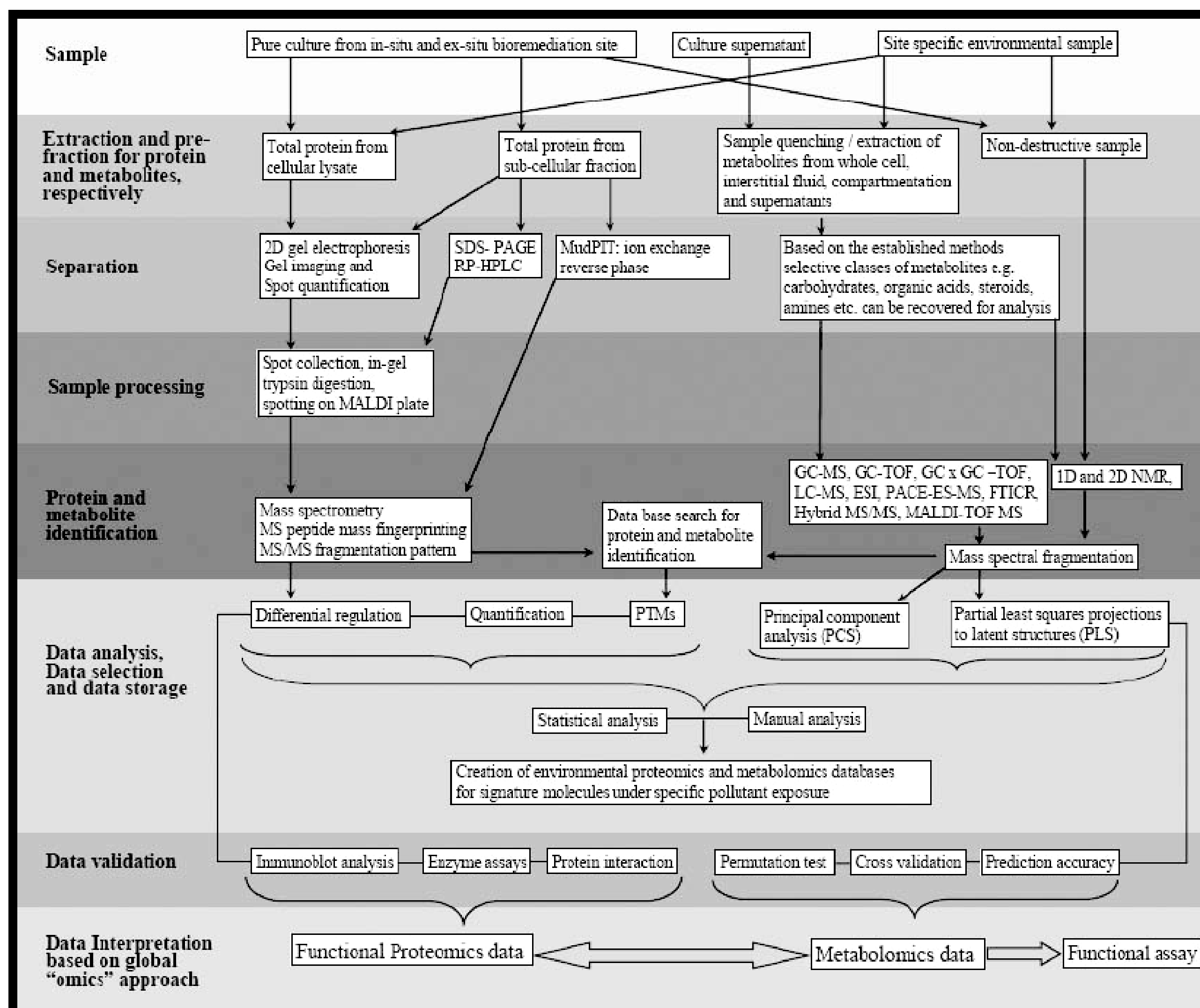

Figure 1. Schematic representation of workflow for proteomics and metabolomics expression profiling experiments. A variety of methods for protein separation and selective metabolite labeling is being established and have been mentioned in the text.

protein spot matching. The sensitivity of DIGE is comparable to that of silver staining, and it is compatible with MS, which can be used to detect the protein of interest after in-gel digestion. However, because of the covalent binding of DIGE with lysine [24], lysine-free proteins would be excluded from the total cellular lysate profiling on 2-DE. In environmental proteomics, DIGE experiments have been performed to profile protein expression in the substrate-dependent regulation of aromatic degradation pathways for toluene and ethylbenzene in a denitrifying bacterium, strain EBN1 [14]. More than 150 protein spots were found in increased abundance in the presence of toluene- or ethylbenzene-adapted cells. Using DIGE, a variety of enzymes has been characterized for phathalate degradation pathway in *Rhodococcus* sp. strain TFB [20]. These studies confirm DIGEs ability to detect proteins of interest in bioremediation and demonstrate the vast heterogeneity of bacterial cell lysate under stress conditions, reinforcing the need for better separation methods.

2.2 Compartmental proteomics in bioremediation

One limitation of 2-DE is the fact that some very basic and hydrophobic membrane proteins are lost due to poor solubility in 2-DE-compatible lysis buffers or precipitation at their p*I* during first-dimension electrophoresis. Membrane proteins, which are particularly prone to precipitation, are of great interest for the biodegradation of aromatic pollutants such as PAHs, NACs, and organophosphates, where many alterations in the organism affect cell–surface proteins and receptors [11, 25, 26]. Membrane proteomics has recently been improved by the development of the multidimensional

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

protein identification platform (MudPIT) where complex protein and mixtures of peptides can be separated using 2-D LC system [27, 28] by omitting 2-DE to avoid the electro-separation of proteins.

MudPIT directly interfaces with MS and can significantly reduce the analysis time for 2-DE. In brief, MudPIT consists of strong cation exchange (SCX) material back-to-back associates with an RP material inside fused-silica capillaries [28]. The chromatographic separation performed in cycles, each cycle involves an increase in salt concentration to "strike" off peptides of the SCX followed by a gradient of increasing hydrophobicity for progressive elution of peptides from the RP into the ion source [28]. The mass spectrometer's data-dependent acquisition isolates peptides as they elute and subjects them to CID by recording the fragmented ions in a tandem mass spectrum. These spectra are matched to database peptide sequences by the SEQUEST algorithm [28]. The identified peptides from SEQUESTs are assembled and filtered into protein-level information using DTASelect algorithm.

Since MS-compatible chromatographic systems only handle small peptides, the protein samples must be digested with protease (usually trypsin) prior to analysis. Fortunately, most proteins contain at least a few peptides that can be analyzed with MS, and MudPIT makes the detection of basic and hydrophobic proteins more likely. A better representation of low-abundance proteins and quantitative comparisons can be achieved by combining MudPIT with isotope-coded affinity tags (ICAT), mass tags that modify cysteine residues and permit the affinity purification of tagged peptides [28, 29]. This reduces the complexity of the peptide sample that needs to be analyzed; however, no proof-of-principle experiment has yet been done to prove this simplicity for various cellular and compartmental proteomics in bioremediation.

2.3 MS in proteomics

Advances in MS have improved the analysis of peptides for protein identification and helped to create the field of environmental proteomics. MS deduces the composition of molecules by determining the particular peptide mass that results from certain combinations of amino acids. The more accurately one can measure peptide mass, the less ambiguous the composition results will be. To study these combinations of amino acids, proteins are routinely digested with proteases to generate peptides small enough for MS analysis, referred to as PMF. In this process, the controlled fragmentation of a peptide yields a series of overlapping fragment ions that differ by the mass of a particular amino acid, which allows the full or partial amino acid sequence of the peptide to be deduced. This phenomenon is also called MS/MS or tandem MS, as it typically uses one MS analyzer to select ions for fragmentation and a second to measure the fragment ions.

MALDI-TOF-MS is routinely used to identify proteins of interest from 2-D gels, as well as to detect and identify viruses, bacteria, fungal spores, and low-mass compounds in environmental samples [30–32]. The complex mass spectra of environmental samples can be used to create characteristic fingerprinting databases to detect many site-specific microorganisms. In terms of bioremediation, MALDI-TOF-MS can detect specific bacterial signature proteins and biomarkers (primary and secondary metabolites) from site-specific samples for the taxonomic identification of potential microorganisms.

A form of direct sample analysis on a microchip using MALDI-TOF-MS – SELDI-TOF-MS – is another promising analytical technique for site-specific samples. A variety of differentially expressed signature proteins in blue mussels (*Mytilus edulis*) exposed to PAHs and heavy metals were analyzed using SELDI-TOF-MS [33]. Although SELDI analysis has been useful for identifying potential biomarkers in clinical research, some have questioned its reproducibility and specificity [34, 35]. Another emerging technique, Fourier transform ICR (FT-ICR) MS, may allow multistage MS experiments (MS^n) with an enhanced detection limit of ~30 zmol for a ~10 kDa protein [36]. Adding ESI and LC to MS has opened new analytical windows for the detection and identification of potential contaminants in the environment [37]. A direct analysis of a peroxisomal protein pattern associated with marine pollutant exposure was performed using ESI MS/MS to identify various epoxide hydrolases, peroxisomal antioxidant enzymes, and SOX [18]. Hence, 2-DE- and MS-based identification of proteins/enzymes from any site-specific bacterium would eventually pave a way toward cell-free bioremediation.

2.4 Protein aggregation and protein–protein interaction

Large-scale studies of protein aggregation, degradation, and protein-protein interaction (PPI) using microarray platforms have increased in popularity over the last decade. The large-scale road mapping of signaling pathways in bacteria became possible when it was recognized that these pathways are organized as interacting units of protein complexes that form the physical machines for signal processing during chemotactic movement of an organism [38]. These specific protein complexes or peptides representing the proteins of interest can be placed on a glass slide in an array-based technology [39]. The major challenge here is generating the proteins themselves for printing libraries. Unfortunately, this method lacks some PTMs, and the yield of folded full-length proteins may be low [40]. Proteins themselves can be printed by contact and noncontact methods, then immobilized using covalent, noncovalent, oriented, or random attachment procedures [39]. Labeling strategies and detection techniques for protein microarrays vary more than the DNA arrays.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

For bioremediation purposes, proteo-arrays can detect binding of specific inhibitors or ligands of dioxygenase or mono-oxygenase enzymatic activities. Enzymatic activity can also be profiled to characterize the neighboring or corresponding enzymes based on qualitative and quantitative estimation during *in situ* bioremediation. However, no report has been published thus far that can represent the global PPI (interactomics) expression pattern of an organism in bioremediation.

2.5 Protein modification in prokaryotes

In prokaryotes, proteins on average are smaller (single domain) than those of eukaryotes, and their functions are often governed by PTMs. Phosphorylation (addition of a phosphate group) is one of the most frequent covalent PTMs. Any protein deciphering PTMs can reveal the functional state of the signaling pathway. The overall PTM mechanisms can be identified through a combination of MS and 2-DE.

In microorganisms, the addition of a phosphate group (usually to serine, tyrosine, threonine, or histidine) is referred to as PTM, and can be catalyzed by three major systems that affect most of the cellular activities [41]. First is the utilization of nucleoside-triphosphate as a phosphoryl donor, which can lead to the modification of the protein substrate on serine/threonine or tyrosine residues. In *Escherichia coli*, several protein kinase activities can phosphorylate more than 130 different endogenous proteins all together [41]. The second system is a "two-component system" that requires a sensor kinase, which autophosphorylates at a histidine residue at the expense of adenosine-triphosphate (ATP), and thereafter induces metabolic changes within the cell [41]. The carbohydrate phosphotransferase (PTS) system makes use of phosphoenol pyruvate to generate a phosphoryl group that is passed down in a chain of several proteins and finally transferred to a sugar molecule [41]. There is a characteristic mass shift during MS analysis of peptides that indicates the site-specific addition of a phosphate group from a single set of proteins [42]. Thus, MS permits the analysis of global changes in the composition or abundance of proteins and determines key proteins for the response of a microorganism in a given physiological state.

2.6 Challenges in environmental proteomics

Proteomics is now successfully established as a valuable tool to characterize the functional molecules of various signaling pathways in biomedical research, but only a few laboratories are currently applying it to environmental concerns. There are no doubts that this technology is expensive and requires highly specialized facilities and needs skilled staff to perform the analyses. Clearly, proteomics technology needs to be developed for environmental cleanup at a more reasonable cost. Nevertheless, progress is still being made to implement proteomics studies in environmental biotechnology laboratories. Even though its applications are mainly beyond the scope of current science, this technology is already offering tangible benefits by helping us to understand the principles of protein expression and identifying new target candidate proteins that can be exploited by other technologies. Also, a bigger challenge has arisen to integrate proteomics with other "-omics" technologies, particularly metabolomics, where low-molecular-weight primary and secondary metabolites are key role players in bioremediation.

3 Metabolomics

While proteomics studies the global expression of proteins, metabolomics characterizes and quantifies their end products: the metabolites, produced by an organism under a given set of conditions [43]. Unlike past studies based on predefined metabolites, metabolomics examines all the metabolites present in a biological system; thus, there is no bias associated with the choice of metabolites to be studied. However, metabolites in a site-specific organism are part of an *in vivo* metabolite flux that regulates entire metabolic pathways. Additionally, metabolism-based wide fluxes (fluxomes) allow us to pinpoint scenarios of physiological regulation in an organism. The key issue in metabolomics is how to exploit the hidden information that exists in different metabolite compositions.

A microbial cell liberates hundreds of primary and secondary metabolites during its life span in response to environmental or cellular changes. These metabolites were previously identified through a process of metabolite fingerprinting [44]. The major difference in the current high-throughput approach is that it concentrates on a relative comparison between samples, rather than on absolute determination of specific known compounds or compound groups. Metabolomics takes the next step beyond metabolite fingerprinting; instead of simply making an inventory of the metabolites in a cell, it aims to quantify every single metabolite in its functional role.

3.1 Principles and measurement strategies for the metabolome

Metabolites are low-molecular-weight organic compounds ($<$1000 Da) that are involved in general metabolic reactions or required for the maintenance, growth, and normal functioning of an organism. Several microorganisms produce primary and secondary metabolites that have been considered as protectors for certain environmental stresses during *in vitro* and *in vivo* decontamination [45, 46]. These are complicated mixtures of compounds, each with its own characteristics. The quantification of all metabolites in a cellular system is an ultimate principle for any metabolomic experiment. This can be achieved by technologies combining automation and miniaturization that have been developed to isolate and characterize metabolites, including technologies for sampling, extraction of specific molecular classes, storage

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

temperature, sample preparation, and analysis [47]. A typical workflow for metabolite experimentation is represented in Fig. 1.

Techniques for analyzing extracted metabolites must be sensitive and robust, with the capacity to screen large numbers of samples. Considerable advances have been made in the development of analytical technologies to measure and interpret complex metabolite profiles. However, because of the wide dynamic and chemical range of low-molecular-weight metabolites in biological mixtures, it has not yet become possible to analyze a global metabolome within a single analytical platform [48]. The techniques most frequently employed for metabolomic studies are NMR spectroscopy and GC-MS and LC-MS.

3.2 NMR spectroscopy and metabolite profiling during bioremediation

It has been over three decades since Eakin *et al.* [49] investigated the catabolism of [1-^{13}C]-glucose by a living cell suspension of *Candida utilis* yeast using NMR. Although the detection sensitivity was poor, they were able to detect the specifically labeled substrate in a nondestructive and noninvasive way. Since that time, significant advances have been made in NMR equipment and methods, but the technology still relies on principles implemented in that pioneering study: using enriched precursors for improved sensitivity and selectivity, recording a series of successive spectra to determine kinetic constants, and examining cell extracts to identify metabolites.

Most of the NMR studies in bioremediation of organic pollutants have been performed in aquatic environments, although a few have used soil or soil models. Bioremediation in soil is detected by solid-state NMR or high-resolution (HR) magic angle spinning (MAS) NMR instead of liquid-state NMR. Microbial degradation of xenobiotics using liquid-state NMR has been limited to heteronuclei NMR studies [50, 51]. Although ^{31}P and ^{39}F are 100% naturally abundant, very few reports concerning them have been published because pollutants bearing these types of atoms are not common. Girbal *et al.* [52] studied the reductive cleavage of demeton-*S*-methyl, an organophosphorus pesticide, by *Corynebacterium glutamicum*, while Rietjens and coworkers published a series of results on the biodegradation of fluorinated aromatic compounds and were reviewed in [53, 54]. The disappearance of 3-fluorobenzoate and appearance of fluoride ions and various fluorinated compounds were observed using ^{19}F NMR from *Sphingomonas* sp. HB-1 in culture medium [55]. Specifically labeled [^{13}C] xenobiotics studies are more limited. The degradation of [1-^{13}C]-dimethyl sulfoniopropionate and [1-^{13}C]-acrylate by a marine isolate has been studied [56].

The kinetics of bioremediation can be studied directly on incubation media using ^{1}H 1D NMR whereas the identification of new metabolites can be performed using 2-D ^{1}H-^{13}C and ^{1}H-^{15}N NMR. Due to the low sensitivity of the ^{15}N nucleus, metabolites must be extracted in higher concentrations in the latter case. The biodegradation pathways of morpholine, thiomorpholine, and piperidine by *Mycobacterium aurum* MO1 and *Mycobacterium* sp. RP1 have been elucidated by *in situ* ^{1}H 1D NMR [50, 51, 57]. A pathway for the degradation of benzo[a]anthracene by *Mycobacterium vanbaalenii* PYR-1 in liquid culture has been proposed using ^{1}H NMR and HPLC analysis [58]. Various degradation pathways have also been proposed based on identified metabolites from mass and ^{1}H NMR spectral analysis when studies were performed for phenanthrene and pyrene degradation at different pH levels by *Mycobacterium* sp. strain PPY-1 [59]. In the future, ^{1}H NMR will likely be used to study biodegradation directly in the soil. Indeed, ^{1}H high resolution-magic angle spinning (HR-MAS) NMR has been successfully applied to study the interaction of pesticides with hydrated matrices of whole soil [60]. HR-MAS NMR has several distinct advantages over existing NMR approaches: (a) it can be performed with a high natural abundance, (b) hydrated matrices representing real environments can be used for detection, (c) "mobile" pollutants can be detected with ease, and (d) bioavailability can be defined in depth.

Another advanced form of NMR, crosspolarization-magic angle spinning (CP-MAS) NMR, has been used to study covalent interactions between organic pollutants and organic matter in the soil. In CP-MAS NMR, ^{15}N, ^{13}C or ^{31}P-enriched xenobiotics were used to detect ^{15}*N*-trinitrotoluene (TNT) [61, 62], ^{13}*C*-benzothiazoles [63], ^{13}*C*-PAHs [64], and ^{31}*P*-organothiophosphate [65]. This type of NMR only provides partial information about the molecular structure, since single-positioned labeled organic pollutants do not yield a straightforward NMR spectrum, which causes misinterpretations of NMR results. Knicker [62] subsequently proposed a double crosspolarization-magic angle spinning (DCP-MAS) NMR technique that combines ^{15}*N*-TNT and ^{13}C-labeled plant material. This technique proved extremely powerful, but was limited by its low sensitivity and longer detection time.

NMR-based analysis has significant advantages, such as little or no sample preparation, nondestructive sample analysis, comprehensive profiling of low-molecular-weight metabolites, and inherent quantification, which can permit mass balance studies to be performed on chemical reactions. However, it is less sensitive than MS and requires relatively large sample volumes. The sensitivity issue has been addressed by developing magnets with increased field strength and improving the design of NMR detectors [66], but long acquisition times are still necessary to detect low-abundance metabolites with NMR.

3.3 MS in environmental metabolomics

The goal of metabolomics is to analyze all metabolites using the extreme sensitivity of MS. The most commonly used technique for holistic analysis of biological samples is quadrupole (Q) GC-MS in electron impact (EI) ionization mode.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

This ionization mode allows the detection of any metabolite eluting from the analytical GC column with a comparable response factor in full-scan acquisition mode, rather than selective detection using chemical positive or negative ionization. A procedure for cocamidopropylbetaine (CAPB) and nonionic alcohol polyethoxylate (AE) surfactant biodegradation was characterized in a hydroponic plant growth system using ESI-Q-IT-MS (ESI-MS) [67]; a predominant homologs, 12 C alkyl CAPB and 9 EO AE, were monitored to represent the total amount of the respective surfactants. This method represents a dynamic linear range of 0.5–250 ng for CAPB and 8–560 ng for AE homolog mixture.

Among other more sensitive MS instruments that have recently become commercially available, GC–TOF-MS instruments are 5–20 times more sensitive than quadrupole detectors, and even greater sensitivity can be achieved using comprehensive GC × GC-TOF-MS rather than GC-TOF-MS. Because of its high sensitivity and speed of application, TOF-MS has now become the method of choice for detecting high-molecular-weight hydrocarbons in various petroleum products [68].

A wide range of compounds can be detected using an ion-trap (Q)-LC-MS system in full-scan acquisition mode. LC-MS provides broad quantitative and qualitative identification of compounds, while fractionation of the LC eluent has often occurred with the MS/MS approach. In an effort to eliminate the polar compounds alkylphenolethoxylates (APEOs) from wastewater in a membrane-assisted and conventional biological pilot plant bioreactor, the effluent was detected by substance-specific analysis applying an atmospheric pressure chemical ionization (APCI) method in combination with MS and MS/MS detection, either in the flow injection (FIA) or liquid chromatographic separation (LC) mode [69]. Degradation products for fluorinated surfactants in advanced oxidation processes were then detected using flow injection-MS, LC-MS, and LC multistage MS [70].

Since metabolites often appear in limited volumes within a complex supernatant of microbial culture, the recent developments in FT-ICR MS have also aroused interest within the field of metabolomics. Since FT-MS has limitations in distinguishing the isomers than the other MS techniques, its capacity to generate spectral data relating to the elemental composition of metabolites suggests that this analytical technique may play a major role in future metabolomic strategies [71].

MALDI-MS has become a workhorse in proteomics, but it can also be used in metabolomics to analyze low-molecular-weight compounds in ionic form. MALDI has recently been applied to low-molecular-weight compounds produced by eukaryotic cells (Islets of Langerhans) and prokaryotic cells (*E. coli* strain DH5-α) [72]; it detected over 100 metabolites from *E. coli*.

The high sensitivity and wide dynamic range of MS make it excellent for analyzing large populations of metabolites. However, its weaknesses include extensive sample preparation, loss of low-abundance metabolites during extraction, variable ionization efficiencies of metabolites, and the inability to provide absolute quantification. The ionic suppression and effects of different matrix systems can also make it difficult to detect certain metabolites [72].

4 Bioinformatics

Protein and metabolite expression profiling is increasingly being performed in the search for molecules that are potentially useful for cell-free bioremediation, which may lead to mass production of specific enzymes and metabolites. To analyze proteome and metabolome data, which are produced in vast amounts, one must understand the analytical procedures used to obtain the data and the statistical principles underlying high-dimensional data. It has been said that "Data are an extremely valuable asset, but like a cash crop, unless harvested, it is wasted". An extensive bioinformatics infrastructure is required to generate, store, analyze, and interpret these results because of the amounts of data involved.

In microbial biotechnology, bioinformatics is used for the computational analysis of laboratory data, identification of protein coding segments, development of proteomics and metabolomics databases, and inference of phenotypes from genotypes [73–75]. So far, four major streamlined criteria have been considered for proteome and metabolome bioinformatics: automated construction of proteome- and metabolome-based regulatory pathways, study of protein–protein and protein–metabolite interaction to understand the regulatory pathways, modeling the 2-D and 3-D structures of proteins, and the construction of 3-D models of protein docking. These studies have been rendered through three major bioinformatics approaches: (i) the use of mathematical modeling techniques such as data mining, statistical analysis, genetic algorithms, and graph matching to identify common patterns, features, and high-level functions; (ii) the use of computation search alignment techniques to compare new proteomes and metabolomes against the set of known proteins and metabolites and annotate the structure and function of new genes up to metabolite level; and (iii) an integrated approach that combines critical mathematical modeling and search techniques.

4.1 Proteome bioinformatics

Information at the proteome level is critical for understanding cellular phenotypes (such as chemotactic movement) under normal and pollutant exposure conditions. The proteomes of pollutant-exposed cells can be altered in many ways [8–14, 19], and some of them are unpredictable from earlier reported genomic/transcriptomic analyses. The identification and development of biomarkers during bioremediation is of great interest to environmentalists. To aid researchers in identifying target proteins directly on 2-DE gels, proteomics databases are

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

now available on the World Wide Web [75], and updated proteomics resources can be found in a recent database issue of *Nucleic Acids Research* (http://nar.oupjournals.org/content/vol33/suppl_1/index.dtl).

Proteomics generation and data analysis are becoming increasingly popular. Typical 2-D-PAGE databases store 2-D gel images obtained from specific organisms and allow users to select protein spots for more information to match and identify proteins of interest. Most of these databases also provide "clickable" map functionality (http://ca.expasy.org/links.html): based on the protein information, gel data, and spot coordination, the 2-D library can dynamically display circles in 2-D gel spot size on the reference gels [75]. Goesmann *et al.* [76] described a newly developed "BRIDGE"-based proteome browser for the Web that currently provides access to three applications: GenDB, EMMA, and ProDB. "DynoProt 2D" is another advanced proteomic database for dynamic online access to proteomes and 2-DE gels based on spot identification and annotation [77]. Major proteomics databases and resources for 2-D gel spot identification are listed in Table 1. In addition to experimental proteome data, complete theoretical proteomes can also be retrieved from databases such as the Proteome Analysis Database (PAD) or generated by tools that calculate theoretical proteome maps [78]. When they are integrated into 2-D gel databases, complete theoretical proteomes are a powerful tool for adding new experimental data.

4.2 Metabolome bioinformatics

Metabolomics is a comprehensive extension of traditional metabolite analysis that includes metabolite fingerprinting, metabolite profiling, and metabolome analysis. Metabolome analysis is a powerful tool for verifying *in silico* pathway construction based on available genome sequences. In such cases, stoichiometric model pathways are constructed *in silico* based on information from well-studied microorganisms and established databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG), enzyme and pathway databases (MetaCyc) and enzyme databases (BRENDA) [73]. A metabolomics database can be any of the following: a database storing detailed metabolite profiles, a single-species-based database, a database of complex metabolite profiles from many different species in different physiological states, a listing of known metabolites for each biological species, a database compilation intended to establish biochemical facts, or a database integrating genome and metabolome data to model metabolic fluxes. The majority of the metabolomics databases listed in Table 2 are based on the construction of biochemical pathways that can be compared to annotated genes and proteins from related microorganisms to identify the known metabolic pathways. The University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) is one of the prime sources for microbial catabolism products of environmental pollutant intermediates [79].

The enormously diverse chemical structures of metabolites are being identified and quantified using various technology platforms. In particular, GC-MS-based metabolite profiling of microcosms and culture media is rapidly becoming one of the cornerstones of functional proteomics, moving toward functional metabolomics. Thus, there is an urgent need for a publicly accessible metabolome database. Based on GC-MS data, Kopka *et al.* [80] have established the Golm Metabolome Database (GMD), which currently focuses on analytical technologies, information to support unequivocal metabolite identification, stored metabolite profiles, and database tracking used to confirm the validity of constructed *in silico* pathways.

Table 1. Major proteomics database resources

Database website	Resources
http://www.expasy.ch/ch2d/	Swiss 2DPAGE
http://www-lecb.ncifcrf.gov/2dwgDB	2DWG Image Meta-database
http://bioinformatics.icmb.utexas.edu/OPD	Open Proteomics Database
http://www.systemsbiology.org	Systems Biology Institute
http://www.sbeams.org	SBEAMS
http://www.expasy.ch/ch2d/2d-index.html	Index
http://mipsgsf.de	MIPS
http://www.bind.ca	BIND
http://dip.doe-mbi.ucla.edu	For interacting proteins (DIP)
http://www.ebi.ac.uk/intact	PPI database (IntAct)
http://www.mysql.com	Management system
http://proteome.ibi.unicamp.br/tools/pimw/index.htm	Protein p*I*/Mw prediction
http://spock.jouy.infra.fr/RL000801.html	MOLOKO
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein	NCBI protein database
http://www.wzw.tum.de/proteomik/lactis/	DynaProt 2D
http://www.cebitec.uni-bielefeld.de/groups/brf/software/brifep	BRIGEP-BRIDGE-based browser
http://compbio.mcs.anl.gov/sentra/	SENTRA (Signal transduction proteins)
http://www.boutell.com/gd	GD library for 2D gels

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Major metabolomic databases and pathways resources

Website	Resources
http://csbdb.mpimp-golm.mpg.de GC-MS-based metabolite profiling	Golm metabolom database
http://MetaCyc.org/	MetaCyc
http://BioCyc.org/ Cellular metabolic networks	BioCyc
http://umbbd.ahc.umn.edu/ Bioremediation pathways for a variety of environmental pollutants	UM-BBD
http://www.genome.ad.jp/kegg/ GeneBank-based pathways flow chart	KEGG
http://ecocyc.org/ *E. coli* pathways	EcoCyc
http://www.chem.qmw.ac.uk/iubmb/enzyme/ For enzyme Nomenclature (EC numbers)	Enzyme Nomenclature
http://www.amaze.ulb.ac.be/ To determine the analytic pathways	AMAZE
http://ch.expasy.org/enzyme/ Enzymatic nomenclature	ENZYME
http://www.brenda.uni-koeln.de/ Enzymatic information	BRENDA
http://www.ncgr.org/pathdb/ Pathways search tools	PathDB
http://www.biocarta.com/genes/allPathways.asp Designing the Pathways maps	BioCarta
http://ch.expasy.org/tools/pathways/ Pathways chart	Roche Biochemical Pathways
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pccompound Organic compound search with Enzyme and protein sequence link	PubChem
http://wit.mcs.anl.gov/WIT2 Metabolic reconstruction	WIT
http://compbio.mcs.anl.gov/puma2/cgi-bin/met_recon.cgi Evolutionary analysis of metabolism	PUMA2
http://www.genome.ad.jp/ligand/ A composite database comprising three sections: compound, reaction and enzyme	LIGAND
http://www.genome.jp/kegg/pathway.htmlenvironmental KEGG pathway database	KEGG
http://gabi.rzpd.de/projects/MapMan/ Large datasets from Arabidopsis Affymetrix arrays, metabolic pathways	MapMan

4.3 Challenges in developing proteome and metabolome databases

The generation and analysis of proteome and metabolome data are becoming more common, and emerging high-throughput techniques will continue to increase the complexity of applied system biology approaches. Despite the mushrooming growth of proteomics and metabolomics, databases still contain many gaps and incomplete evidence for annotations. Annotations that do not state whether the information is based on direct experimental evidence or on a theoretical model can be misleading. Therefore, there is a strong need to develop more detailed meta-data describing the experiments in detail and the source of obtained values. It is crucial to identify specific proteins or related metabolites that appear to change in order to make comparisons between normal and pollutant-exposed conditions. Future databases with a focus on integrative biology must allow the exploration and retrieval of data from a universal point of view under specific physiological conditions.

The miniaturized and automated procedures that are now being applied to proteomics might be useful in future metabolomics to provide a basis for generating highly reproducible results. These technologies produce ever-increasing amounts of data to be stored; therefore, standard ways of representing data and an agreed minimum level of annotation must be established. Steps must also be taken to promote data exchange between labs. The exchange and re-analysis of proteome and metabolome data would be a time-consuming task, but would confirm the applicability of output results. A few efforts have been made in the past to propose recommendations for standardization and reporting

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

metabolic analysis data [81, 82]. One of the challenges that both technologies share is to distinguish between the biological variability and analytical reproducibility of the molecule of interest.

5 Combined proteomics and metabolomics for bioremediation

The analysis of low-molecular-weight metabolites related to proteins through functional genes is essential in bioremediation. mRNAs, which are highly unstable transmitters on the path from genes toward functional proteins and metabolites, are assumed to depict the final products of gene expression. Proteomics, on the other hand, offers detailed information regarding quantitative correlation with gene expression, protein modifications by PTMs, and differential expression of compartmental proteins into membrane, cytosol, and periplasm. A comprehensive whole-proteome analysis, including the metabolome, would be especially helpful to understand the functions of bioremediation-relevant microorganisms that have not yet been studied in detail. The global protein expression study supports the view that proteo-arrays can record differential expression in known proteins more accurately and completely [39, 40] than the 2-DE.

Metabolomics has shown that true metabolic fluxes precisely correlate with proteins and mRNA abundance [83]. Since we know that the differential expression levels of various metabolites show physiological changes in organisms, it is surprising that global metabolite profiling is not yet possible; thus far, proteomics data have been deemed necessary to complement the metabolomics approach. However, this means that proteomics and metabolomics will retain their central positions among postgenomic techniques. Detailed information about the bioremediation-related physiological pathways of an organism and its proteins and metabolites can only be discovered *via* a joint proteomics and metabolomics. Given the current state of the technology, it seems feasible to integrate these two technologies into one unified term "proteometabolomics".

6 Conclusion

Environmental contamination can be viewed as an ecological malaise for which bioremediation can be prescribed as a "perfect medicine" [3]. However, total field bioremediation is a difficult task, whether using genetically engineered microorganisms or intrinsic microorganisms. The solution to the problems with bioremediation lies in discovering to what extent the microbes' physiological machinery contributes to the degradation process and which biomolecules and their mechanisms are responsible for different factors within the degradation system, such as PPI, metabolite and enzymatic chemical transformation, *etc*., in order to pursue cell-free bioremediation. Continued scientific advancement will ultimately allow for comprehensive integrated approaches using gene, protein, and metabolite expression to study the functional physiology of an organism. Proteomics will certainly play an essential role in bacterial site remediation, although it will benefit from integration with metabolomics. Combining metabolomics with existing transcriptomics methods will also enable us to address gene function comprehensively; this will be useful in situations where protein expression is a poor monitor of regulation, and it may yield a significantly greater understanding of functional proteomics. This combined approach will allow us to understand better microbial physiology, which may eventually pave the way toward cell-free bioremediation.

OVS is the sole contributor of this original review article. The review is based on published current research in the area of environmental proteomics and metabolomics.

Technical support rendered by Ms. Rashmi Singh for preparing this manuscript is gratefully acknowledged. The author thanks the anonymous reviewers for excellent comments and suggestions.

7 References

[1] Labana, S., Pandey, G., Paul, D., Sharma, N. K. *et al.*, *Environ. Sci. Technol.* 2005, *39*, 3330–3337.

[2] Labana, S., Singh, O. V., Basu, A., Pandey, G., Jain, R. K., *Appl. Microbiol. Biotechnol.* 2005, *68*, 417–424.

[3] Samanta, S., Singh, O. V., Jain, R. K., *Trends Biotechnol.* 2002, *20*, 243–248.

[4] Carvalho, M. F., Ferreira, J. R., Pacheco, C. C., De Marco, P., Castro, P. M., *Environ. Microbiol.* 2005, *7*, 294–298.

[5] Zocca, C., Di Gregorio, S., Visentini, F., Vallini, G., *FEMS Microbiol. Lett.* 2004, *238*, 375–382.

[6] Esteve-Nunez, A., Caballero, A., Ramos, J. L., *Microbiol. Mol. Biol. Rev.* 2001, *65*, 335–352.

[7] Lehmann, A. R., *Genes Dev.* 2001, *15*, 15–23.

[8] Chin, L., Pomerantz, J., DePinho, R. A., *Trends Biochem. Sci.* 1998, *23*, 291–296.

[9] Wang, R. F., Wennerstrom, D., Cao, W. W., Khan, A. A., Cerniglia, C. E., *Appl. Environ. Microbiol.* 2000, *66*, 4300–4304.

[10] Krivobok, S., Kuony, S., Meyer, C., Louwagie, M. *et al.*, *J. Bacteriol.* 2003, *185*, 3828–3841.

[11] Pessione, E, Giuffrida, MG, Prunotto, L, Barello, C. *et al.*, *Proteomics* 2003, *3*, 1070–1076.

[12] Kim, S. J., Jones, R. C., Cha, C. J., Kweon, O. *et al.*, *Proteomics* 2004, *4*, 3899–3908.

[13] Santos, P. M., Benndorf, D., Sa-Correia, I., *Proteomics* 2004, *4*, 2640–2652.

[14] Kuhner, S., Wohlbrand, L., Fritz, I., Wruck, I. *et al.*, *J. Bacteriol. 2005, 187, 1493–1503.*

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[15] Singh, O. V., Nagaraj, N. S., *Brief. Funct. Genomic Proteomic* 2006, *4*, 355–362.

[16] Vasseur, C., Labadie, J., Hebraud, M., *Electrophoresis* 1999, *20*, 2204–2213.

[17] Wilkins, J. C., Homer, K. A., Beighton, D., *Appl. Environ. Microbiol.* 2001, *67*, 3396–3405.

[18] Mi, J., Orbea, A., Syme, N., Ahmed, M. *et al.*, *Proteomics* 2005, *5*, 3954–3965.

[19] Kim, Y. H., Cho, K., Yun, S. H., Kim, J. Y. *et al.*, *Proteomics* 2006, *6*, 1301–1318.

[20] Tomas-Gallardo, L., Canosa, I., Santero, E., Camafeita, E. *et al.*, *Proteomics* 2006, *6*, S119–S132.

[21] Methe, B. A., Nelson, K. E., Deming, J. W., Momen, B. *et al.*, *Proc. Natl. Acad. Sci. USA* 2005, *102*, 10913–10918.

[22] Marouga, R., David, S., Hawkins, E., *Anal. Bioanal. Chem.* 2005, *382*, 669–678.

[23] Alban, A., David, S. O., Bjorkesten, L., Andersson, C. *et al.*, *Proteomics* 2003, *3*, 36–44.

[24] Unlu, M., Morgan, M. E., Minden, J. S., *Electrophoresis* 1997, *18*, 2071–2077.

[25] Wang, Y., Rawlings, M., Gibson, D. T., Labbe, D. *et al.*, *Mol. Gen. Genet.* 1995, *246*, 570–579.

[26] Richins, R. D., Kaneva, I., Mulchandani, A., Chen, W., *Nat. Biotechnol.* 1997, *15*, 984–987.

[27] Paoletti, A. C., Zybailov, B., Washburn, M. P., *Expert Rev. Proteomics* 2004, *1*, 275–282.

[28] Somiari, R. I., Somiari, S., Russell, S., Shriver, C. D., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2005, *815*, 215–225.

[29] Jue-Liang, H., Shu-Hui, C., *Curr. Proteomics* 2005, *2*, 287–302.

[30] Lay, J. O., Jr., *Mass Spectrom. Rev.* 2001, *20*, 172–194.

[31] Ruelle, V., Moualij, B. E., Zorzi, W., Ledent, P., Pauw, E. D., *Rapid Commun. Mass Spectrom.* 2004, *18*, 2013–2019.

[32] Hu, L., Xu, S., Pan, C., Yuan, C. *et al.*, *Environ. Sci. Technol.* 2005, *39*, 8442–8447.

[33] Knigge, T., Monsinjon, T., Andersen, O. K., *Proteomics* 2004, *4*, 2722–2727.

[34] Jacobs, I. J., Menon, U., *Mol. Cell. Proteomics* 2004, *3*, 355–366.

[35] Diamandis, E. P., *Mol. Cell. Proteomics* 2004, *3*, 367–378.

[36] Laskin, J., Futrell, J. H., *Mass. Spectrom. Rev.* 2005, *24*, 135–167.

[37] Bossi, R., Seiden, P., Andersen, S. M., Jacobsen, C. S., Streibig, J. C., *J. Agric. Food Chem.* 1999, *47*, 4462–4468.

[38] Larsen, M. H., Blackburn, N., Larsen, J. L., Olsen, J. E., *Microbiology* 2004, *150*, 1283–1290.

[39] Sauer, S., Lange, B. M., Gobom, J., Nyarsik, L. *et al.*, *Nat. Rev. Genet.* 2005, *6*, 465–476.

[40] Predki, P. F., *Curr. Opin. Chem. Biol.* 2004, *8*, 8–13.

[41] Cozzone, A. J., *Biochimie* 1998, *80*, 43–48.

[42] Mann, M., Jensen, O. N., *Nat. Biotechnol.* 2003, *21*, 255–261.

[43] Whitfield, P. D., German, A. J., Noble, P. J., *Br. J. Nutr.* 2004, *92*, 549–555.

[44] Shockcor, J. P., Holmes, E., *Curr. Top Med. Chem.* 2002, *2*, 35–51.

[45] Park, S. J., Lee, S. Y., Cho, J., Kim, T. Y. *et al.*, *Appl. Microbiol. Biotechnol.* 2005, *68*, 567–579.

[46] van der Werf, M. J., Jellema, R. H., Hankemeier, T., *J. Ind. Microbiol. Biotechnol.* 2005, *32*, 234–252.

[47] Buchholz, A., Hurlebaus, J., Wandrey, C., Takors, R., *Biomol. Eng.* 2002, *19*, 5–15.

[48] Glassbrook, N., Ryals, J., *Curr. Opin. Plant Biol.* 2001, *4*, 186–190.

[49] Eakin, R. T., Morgan, L. O., Gregg, C. T., Matwyoff, N. A., *FEBS Lett.* 1972, *28*, 259–264.

[50] Delort, A.-M., Combourieu, B., *J. Ind. Microbiol. Biotechnol.* 2001, *26*, 2–8.

[51] Combourieu, B., Haroune, N., Besse, P., Sancelme M., Delort, A.-M., in: Mohan, R. M. (Ed.), *Res. Adv. Microbiol. 3*, Global Research Network, 2003, pp. 1–22.

[52] Girbal, L., Hilaire, D., Leduc, S., Delery, L. *et al.*, *Appl. Environ. Microbiol.* 2000, *66*, 1202–1204.

[53] Grivet, J. P., Delort, A. M., Portais, J. C., *Biochimie* 2003, *85*, 823–840.

[54] Solyanikova, I. P., Moiseeva, O. V., Boeren, S., Boersma, M. G. *et al., Appl. Environ. Microbiol.* 2003, *69*, 5636–5642.

[55] Boersma, F. G., McRoberts, W. C., Cobb, S. L., Murphy, C. D., *FEMS Microbiol. Lett.* 2004, *237*, 355–361.

[56] Ansede, J. H., Pellechia, P. J., Yoch, D. C., *Appl. Environ. Microbiol.* 2001, *67*, 3134–3139.

[57] Combourieu, B., Besse, P., Sancelme, M., Godin, J. P. *et al.*, *Appl. Environ. Microbiol.* 2000, *66*, 3187–3193.

[58] Moody, J. D., Freeman, J. P., Cerniglia, C. E., *Biodegradation* 2005, *16*, 513–26.

[59] Kim, Y. H., Freeman, J. P., Moody, J. D., Engesser, K. H., Cerniglia, C. E., *Appl. Microbiol. Biotechnol.* 2005, *67*, 275–285.

[60] Simpson, A. J., Kingery, W. L., Shaw, D. R., Spraul, M. *et al.*, *Environ. Sci. Technol.* 2001, *35*, 3321–3325.

[61] Knicker, H., Achtnich, C., Lenke, H., *J. Environ. Qual.* 2001, *30*, 403–410.

[62] Knicker, H., *Sci. Total Environ.* 2003, *308*, 211–220.

[63] Witte, E. G., Philipp, H., Vereecken, H., *Org. Chem.* 2002, *33*, 1727–1735.

[64] Kacker, T., Haupt, E. T., Garms, C., Francke, W., Steinhart, H., *Chemosphere* 2002, *48*, 117–131.

[65] Seger, M. R., Maciel, G. E., *Environ. Sci. Technol.* 2006, *40*, 791–796.

[66] Griffin, J. L., *Philos. Trans. R Soc. Lond. B Biol. Sci.* 2004, *359*, 857–871.

[67] Levine, L. H., Garland, J. L., Johnson, J. V., *Anal. Chem.* 2002, *74*, 2064–2071.

[68] Dutta, T. K., Harayama, S., *Anal. Chem.* 2001, *73*, 864–869.

[69] Schroder, H. F., *Water Sci. Technol.* 2002, *46*, 57–64.

[70] Schroder, H. F., Meesters, R. J., *J. Chromatogr. A* 2005, *1082*, 110–119.

[71] Aharoni, A., Ric de Vos, C. H., Verhoeven, H. A., Maliepaard, C. A. *et al.*, *OMICS* 2002, *6*, 217–234.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[72] Edwards, J. L., Kennedy, R. T., *Anal. Chem.* 2005, *77*, 2201–2209.

[73] Arita, M., *Brief. Funct. Genomic. Proteomic.* 2004, *3*, 84–93.

[74] Bansal, A. K., *Microb. Cell Fact.* 2005, *4*, 19.

[75] Kremer, A., Schneider, R., Terstappen, G. C., *Biosci. Reports* 2005, *25*, 95–106.

[76] Goesmann, A., Linke, B., Bartels, D., Dondrup, M. *et al.*, *Nucleic Acid Res.* 2005, *33*, W710–W716.

[77] Drews, O., Gorg, A., *Nucleic Acid Res.* 2005, *33*, D583–D587.

[78] Pruess, M., Fleischmann, W., Kanapin, A., Karavidopoulou, Y. *et al.*, *Nucleic Acid Res.* 2003, *31*, 414–417.

[79] Ellis, L. B. M., Hou, B. K., Kang, W., Wackett, L. P., *Nucleic Acid Res.* 2003, *31*, 262–265.

[80] Kopka, J., Schauer, N., Krueger, S., Birkemeyer, C. *et al.*, *Bioinformatics* 2005, *21*, 1635–1638.

[81] Lindon, J. C., Nicholson, J. K., Holmes, E., Keun, H. C. *et al.*, *Nat. Biotechnol.* 2005, *23*, 833–838.

[82] Kell, D. B., *Biochem. Soc. Trans.* 2005, *33*, 520–524.

[83] Kromer, J. O., Sorgenfrei, O., Klopprogge, K., Heinzle, E., Wittmann, C., *J. Bacteriol.* 2004, *186*, 1769–1784.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim