5 Soil Proteomics: Extraction and Analysis of Proteins from Soils

Oladele A. Ogunseitan

5.1 Introduction

The diversity of organisms in soil habitats is influenced by complex interactions among various physical, chemical, and ecological factors. Microorganisms belonging to all three major phylogenetic domains, namely Archaea, Bacteria, and Eukarya, inhabit soils according to the availability and configuration of niches. These microbes engage in symbiotic, neutral, or pathogenic relationships with large multicellular organisms such as plants. These interactions result in the biogeochemical cycling of several elements that form permanent or transient features of soils. However, these cycles also link soil processes to atmospheric and aquatic environments. Therefore, methods that are developed to investigate structural and functional diversity of soil microorganisms must be sensitive to the fluidic nature of the processes attributed to the organisms and to the possibility of trans-compartmental interactions. However, it has proven difficult to estimate biological diversity in soils because most of the available methods are dependent on cultivation of organisms under laboratory conditions, a notoriously inadequate strategy for accounting for the majority of microbial species. It is for this reason that methodological approaches which focus on molecular signatures have become so powerful since they attempt to resolve questions of structure and function without the need to cultivate organisms on the basis of habitat characteristics (Torsvik et al. 1996, 1998; Nannipieri et al. 2003).

Nucleic acid-based approaches for investigating soil microbial communities rely on the coordinated functions of protein molecules that serve as enzymes for recognising, replicating, and amplifying specific nucleic acid sequences. In addition to this fundamental linkage between protein function and the assessment of genetic diversity, protein molecules represent the final result of genetic expression, and, through their functions as physiological catalysts, structural components, signal transducers, and

Oladele A. Ogunseitan: Laboratory for Molecular Ecology, Department of Environmental Health, Science, and Policy, University of California, Irvine, California 92697–7070, USA, E-mail: oaogunse@uci.edu

Soil Biology, Volume 8
Nucleic Acids and Proteins in Soil
P. Nannipieri, K. Smalla (Eds.)

mediators of intercellular communication, they control key reactions in ecological processes performed by microorganisms in soils. The analysis of microbial proteins has traditionally focused on comparative quantitative assessments and functional characteristics of a few phylogenetically conserved molecules. For example, cytochromes, protein elongation factors, and ATPases are among the most widely studied protein molecules for which their genetic coding sequences have been used in systematics (Liu and Stahl 2002). Whereas for ecological functions, most research employing protein-based approaches has focused on enzymes of significance to soil fertility (Nannipieri et al. 2003). However, these previous studies can hardly be called "molecular" because inferences were made without adequate resolution of molecular-scale observations that could facilitate reliable conclusions on system diversity and predictive modeling of responses to variable environmental factors. This chapter focuses on the extraction and analysis of protein molecules as an approach for elucidating structural and functional diversity of soil microbial communities. A survey of methods is presented first, followed by selected case studies on the uses of protein-based methods to solve critical ecological questions. The first case study concerns the production and activities of glomalin by soil microbes. The second case study concerns the extraction of soil proteins to indicate the biological availability of toxic chemicals.

5.2 Rationale and Context of Soil Proteomics

Biogeochemists interested in the nitrogen cycle are credited with the invention of quantitative methods for direct extraction of amino acids and proteins from sediments and soils (Greenfield et al. 1970; Cheng 1975). There is a rather long and detailed historical record of numerous attempts to use semi-qualitative methods for extracting proteins, particularly enzymes, from soil samples. These early techniques were subsequently developed with varying degrees of success to investigate specific enzymes important in agriculture and environmental contamination (Briggs and Segal 1963; Weetall et al. 1965; Burns et al. 1972; Bremner and Zantua 1975; Mayaudon et al. 1975; McLaren et al. 1975; Thornton and McLaren 1975; Pettit et al. 1976; Skujins 1976; Burns 1978; Mayaudon 1986; Nybroe et al. 1990; Tabatabai and Fu 1991; Nannipieri et al. 1996). More recent developments have focused on increasing the throughput of enzyme assays in soil extracts to facilitate rapid assessment of soil conditions and to minimise assay variability due to differences in soil characteristics (e.g. Stemmer 2004).

Soils are extremely complex media in terms of biochemical characteristics. As such, all methods of soil analysis are, in practice, limited to or optimised for a certain fraction of the inherent diversity of organisms and molecules. The total protein fraction of soils represents contributions from prokaryotic and eukaryotic microorganisms as well as from large multicellular eukaryotes and anthropogenic wastes. The portion of the protein fraction that is not immobilised by soil colloids such as clay and humic matter is located in one or more of three major compartments depending on the type and physicochemical conditions of the soil. Intracellular proteins occur in microbial cells as well as animal and plant tissues present in soils. Extracellular protein molecules may exist in the aqueous interphase between particles in moist soils, and this protein fraction results from lysed cells or secreted enzymes and other bioactive proteins. The third soil protein fraction is located on the surface of cells, where the molecules function as exoenzymes active in the transportation of nutrients and growth factors into the cells, or in the active excretion of toxic metabolites out of the cell. Methods that have been developed to assess soil proteins, particularly enzymes, optimise to different extents the recovery of these three fractions. In addition to focusing on specific soil protein fractions, certain methods have also been developed for exhaustive recovery of protein molecules without necessarily preserving activity or secondary structure; on the other hand, other methods strive to preserve protein functions such as catalysis, but, in doing so, the gentle approach of these methods does not achieve complete recovery of proteins present in the soil environment.

Strategies for bulk soil protein assessments are based in part on the rationale that a complete proteome map of a microbial community will facilitate the discovery of unique polypeptides whose production is mediated by specific environmental cues. The molecular resolution of soil proteins for elucidating species diversity and metabolic productivity has proven difficult because of the large number of different proteins synthesised by different species, even in axenic microbial cultures (Blom et al. 1992; Ogunseitan et al. 2001). Despite this limitation, the refinement of immunological techniques coupled with the increasing number of available enzyme assays have been used to supplement the need for detailed molecular resolution in cases where attention is focused on particular polypeptides (Feinstein and Lindahl 1973; Mayaudon and Sarkar 1974; Bohlool and Schmidt 1980; Selander et al. 1986; Wright 1992; Ogunseitan 1993; Paul 1993; Jehr and Hussain 1994). For example, the discovery of glomalin, the complex proteinaceous substance which has been implicated in various structural and functional features of soils, was facilitated by exhaustive soil protein extraction and immunological detection, but the molecular details of this protein remain elusive (Wright and Upadhyaya 1996; Wright et al. 1996; Rillig et al. 2003).

Although the frequency of discoveries in proteomics is increasing, there are major gaps in the state of knowledge linking the presence of specific proteins in complex microbial communities to ecological functions within different fields of investigation. For example, the use of protein sequence information to construct phylogenetic trees has greatly contributed to microbial systematics and the exploration of both quantitative and qualitative aspects of microbial diversity, whereas information on the biochemical ecology of proteins has typically been inferred from axenic cultures lacking the context of multi-species community that enables protein molecules to implement various levels of interactions among organisms (Kersters and De Ley 1980; Jackman 1985; Hantula et al. 1990, 1991; Ogunseitan et al. 2002; Ogunseitan 2005). Therefore, it is questionable whether phenomena discovered under axenic conditions are meaningful in natural ecosystems, but the emergence of a repertoire of methods and techniques that support the analysis of protein synthesis, diversity, and function in natural heterogeneous microbial communities is enhancing research towards realistic environmental proteomics (Nannipieri et al. 1974; Busto et al. 1995; Ogunseitan 1997, 1998).

Elucidating ecosystem-level activities by assessing molecular diversity in soil environments demands an awareness of the complexity of protein structures when compared to other signature biomolecules such as nucleic acids and fatty acids. Protein complexity increases the richness of information extractable from an environmental sample, but the available techniques for analysing the information are not only few, but also complicated by many layers of uncertainty. The fundamental level of information contained in protein molecules is represented by the primary structure, or the sequence of ~20 possible amino acids making up the polypeptide chain. Comparative protein profiles based on physical size and electric charge properties can be routinely constructed through one- or two-dimensional polyacrylamide gel electrophoreses (PAGE) which are readily available for studies focused on temporal and spatial variations in genetic induction of protein synthesis. Amino acid sequence analysis can also address questions of evolutionary diversity among specific groups of organisms occupying similar ecological niches. However, protein function also depends on its secondary (folded protein), tertiary (globular protein), or quaternary (several interacting folded polypeptides) forms. These morphological conformations contribute significantly to the ability of proteins to perform crucial functions such as enzymatic catalysis and organelle construction. The abundance and diversity of microbial proteins suggest fairly straightforward extraction methods, but attention must be paid to extraction conditions that preserve protein integrity and function, while reducing interference from co-extracted substances such as nucleic acids and humic materials. It is particularly important that functional enzyme

studies employ extraction, resolution, and detection techniques that optimise the stability of protein conformation while minimising interference by potential inhibitors such as metal ions and detergents which act as denaturing agents.

The rationale for direct extraction of proteins from soils and other environmental samples includes monitoring the activities of particular enzymes without necessarily identifying specific organism sources (Ogunseitan 1998; Ogunseitan et al. 2000); monitoring the fate of genetically engineered organisms that produce commercially desirable protein molecules such as the *Bacillus thuringiensis* crystal toxin (Saxena et al. 2002; Lee et al. 2003); detection of specific antigens associated with non-culturable bacteria or pathogenic virus particles in soil (Vettori et al. 2000); and characterisation of the response of microorganisms to stressful physical-chemical environmental conditions (Ogunseitan 1996, 1999). In the following sections, techniques useful for extracting proteins from complex microbial communities are described with emphasis on the preservation of enzymatic activities. New techniques, including protein chip arrays, are described briefly because they offer high resolving power, although their application has thus far principally been limited to the study of pure cultures and simple assemblages of microbial populations.

5.3 Methodology for Soil Proteomics

Methods for extracting protein molecules from soils are entirely independent of the subsequent methods of protein analysis. If the analytical methods involve assessment of protein function, including enzyme activity, then the extraction methods must preserve function as much as possible. For example, the availability of colorimetric staining techniques for more than 300 enzymes (Manchenko 1994) coupled with nitrocellulose membrane-immobilisation of protein molecules extracted from environmental samples has aided rapid screening of multiple enzymes from a single source, thereby facilitating comparative analysis of specific metabolic activities across samples (Ogunseitan et al. 2000, 2001, 2002). Alternatively, if analytical methods are focused on molecular size comparisons or immunological detection, then extraction methods are generally less conservative.

5.3.1 Extraction Methods

All protein extraction methods aim to achieve quantitative recovery of polypeptides from the complex matrix of soil particles, and to purify the

extract as much as possible by removing unwanted molecules that may interfere with further analysis. However, many protein molecules exist in nature as complexes with other kinds of molecules, including colloids, nucleoproteins, glycoproteins, and metalloproteins. In some cases, these complexes interfere with protein detection and quantification techniques, and, importantly, methods which aim at detecting protein function should not strip polypeptides of the essential non-proteinous components that are essential for optimum activity.

Sonication is one of the most common methods for breaking apart aggregated soil particles and for lysing cells to release protein molecules. However, sonication may also cause the denaturation of proteins (sometimes visible as foam) under certain conditions, depending on soil characteristics; but working at low temperatures ameliorates this detrimental effect. Following sonication, centrifugation is typically used to recover protein fractions which may be further purified by dialysis or salt precipitation (Ogunseitan and LeBlanc 2004). In certain cases, "soil washing" is used to first release cells from the soil matrix prior to protein extraction, although this step introduces another level of uncertainty that compromises the exhaustive recovery of the protein content of soils (Lindahl and Bakken 1995).

Many techniques have been published for extracting protein molecules from bulk soil, and some case-by-case evaluation is still necessary. In the simplest case, briefly, the extraction of proteins from bulk soils usually involves the suspension of 1–100 g of fresh sample in 1.5 volumes (w/v) of cold 0.3 M K_2HPO_4, plus 0.5 volume of 0.3 M EDTA (pH 8.0; Singleton et al. 2003; Harner et al. 2004; Ogunseitan and LeBlanc 2004). The suspension is subjected to sonication using a medium-tipped probe such as the 1 cm tip accompanying the sonic dismembrator model 550 (Fisher Scientific, Tustin, CA, USA) at ~20 kHz for 30–60 min (in 10 s pulses) in an ice bucket. The samples are then agitated on a rotary shaker at 100 rpm in a cold room (0–4 °C) for 4 h to overnight. To recover the protein fraction, the samples are centrifuged at 3,000 *g* for 10 min to remove large particles. The supernatant is collected for repeated centrifugation at 25,000 *g* for 25 min at 4 °C to completely remove all suspended particulate matter. Protein molecules can be precipitated from the supernatant by adding 60% (w/v) $(NH_4)_2SO_4$, adjusting the pH to 7.0 by adding 1 M KH_2PO_4, and incubating at 0–4 °C for 2–6 h. Precipitated proteins are then recovered by centrifugation at 25,000 *g* for 25 min at 4 °C, followed by resuspension in an appropriate volume (e.g. 10 ml) of 0.2 M K_2HPO_4 per gram of wet precipitate, with adjustment of pH to 7.0.

Further purification of soil protein extracts can be achieved by dialysis against 0.02 M phosphate buffer overnight in a cold room (0–4 °C; Ogunseitan 1993). The dialysate is centrifuged at 25,000 *g* for 25 min at 4 °C to

collect supernatant which is now ready for the determination of protein concentration, enzyme assays, electrophoresis, and other analytical methods (Schulze et al. 2005). Soil protein extracts may also be filtered through a 0.2 μm Nalgene syringe filter (Nalge, Rochester, NY, USA) and stored at −20 °C until further analysis.

Several variations to the general protocol described in the previous paragraphs have been reported in the literature, with the most notable alternatives to sonication being the "freeze/thaw", "bead-beating", and "autoclaving" methods (Wright and Upadhyaya 1996; Bolygo et al. 2003; Barzaghi et al. 2004). The former two methods begin by suspending samples of soil (e.g. 1 g at 50% water-holding capacity) in 1 ml of an extraction buffer consisting of 50 mM Tris-HCl, 10% sucrose, 2 mM dithiothreitol, 4 mM EDTA, and 0.1% Brij 58. The final pH of the buffer is adjusted to 7.5 with ammonia solution, and 100 μl of protease inhibitor cocktail (e.g. Sigma P2714) are added to the sample plus buffer (Eprogen Corporation 2004). For the freeze/thaw approach, the soil suspension is vortexed briefly (10 s) and subjected to four cycles of freezing in liquid nitrogen and thawing at 25 °C in a vented container (Ogunseitan 1993). The supernatant is then collected by sequential centrifugation to ensure the removal of all soil particles and cell debris. Sample preparation is similar for the bead-beating approach except that, instead of freeze/thawing, the soil suspension is shaken at 5.5 m s^{-1} for 30 s in a vessel containing 0.2 g of sterile glass beads (150–212 μm diameter available from Sigma, St. Louis, MO, USA). Both the freeze/thaw and bead-beating approaches have been used to monitor the response of soil proteins to toxic pollutants (Singleton et al. 2003).

The autoclaving approach has been used most notably in the recovery of abundant recalcitrant proteins such as glomalin from soil (Wright and Upadhyaya 1996, 1998; Rillig et al. 2002a). For example, total soil glomalin can be extracted by suspending 1 g of soil samples in 8 ml of an extraction buffer consisting of 50 mM citrate at pH 8.0 and autoclaving this mixture at 121 °C for 60 min. The supernatant is then recovered by centrifugation at 5,000 *g* for 20 min. The autoclaving and centrifugation cycle is repeated for as many times as necessary until the supernatant is clear of the brown discolouration attributable to glomalin. A final centrifugation at 10,000 *g* is conducted to remove all soil particles and cell debris.

5.3.2 Analytical Methods

The reliability of inferences about the presence or absence of specific polypeptides in a soil protein extract depends not only on the rigour of the extraction procedure, but also on the accuracy of the technique used to

estimate the concentration of protein molecules in the extract, and on the sensitivity of the analytical technique used for protein detection. For most purposes, the determination of protein concentration in a given extract can be achieved through spectrophotometric analysis using various dyes (e.g. Bradford Dye assay; US Biochemicals, Cleveland, OH, USA). The sensitivity of the protein assay can be improved by incubating the dye reaction at 37 °C for 1 h prior to spectrophotometer readings (Bradford 1976; Zor and Selinger 1996). If necessary, protein molecules can be concentrated further in the extract by centrifugation at 2,000 *g* for 10 min in a Centriprep cartridge with a 3 kDa molecular weight cut-off (Ogunseitan and LeBlanc 2004).

Protein Profiling by Electrophoretic Resolution

PAGE in one or two dimensions is a relatively straightforward and quick way of assessing the diversity of polypeptides in a protein extract, although the resolving power and visualisation of protein bands depend on the staining protocol (typically silver-based or Coomassie brilliant blue is used); on the level of interference by the presence of contaminants which may produce low signal-to-background ratios; and on the level of sophistication of software used for comparative assessment of polypeptide banding patterns within gels and across electrophoretic runs (Ogunseitan 1993, 1999; Ogunseitan et al. 2001; Singleton et al. 2003). In cases where specific experiments allow the use of radioactive tracers, the identity of newly synthesised proteins in a microbial community can be determined by autoradiographical analysis of gels used to resolve complex protein extracts (Ogunseitan 1996). In these cases, radioisotopes are typically added to soil microcosms in the form of inorganic sulphate or labelled methionine. Alternately, depending on the objectives of a given experiment, radioactive substrates may be provided to monitor the outcome of catabolic reactions conducted by specific enzymes.

Enzyme Assays

When enzymes are targeted for monitoring, starch gel electrophoresis is useful for evaluating protein extracts for immobilised enzyme activities (Ogunseitan 1997). Enzyme assays can also be performed on nitrocellulose membranes, or ultrathin polyacrylamide gels (Jehr and Hussain 1994; Ogunseitan 1998; Ogunseitan 2002). Based on these approaches, it is conceivable that if sufficient level of purity is achieved, soil protein extracts can eventually be used for multilocus enzyme electrophoresis (MLEE) to reveal polymorphisms at specific genetic loci in the various populations of a microbial community (Shaw and Prasad 1970; Selander et al. 1986). These polymorphisms are expressed as redundancies in metabolic capac-

ity which contributes to the resilience of ecological functions in stressful environments. For example, the occurrence and distribution of various forms of enzymes which express the same activity but differ in amino acid sequence (allozymes are produced by different alleles at a single locus) can reveal meaningful information on evolutionary divergence of different population members which colonise the same niches in an ecological community. Such analyses have been conducted for aquatic systems (Ogunseitan 1999; Ogunseitan et al. 2002). The advantage of using nitrocellulose membranes in a slot-blot apparatus is that larger volumes of protein extract can be immobilised than when using gel-based techniques. However, with the membrane technique it is not possible to detect enzyme polymorphisms at the molecular level. Soil protein extracts can also be used for immunological detection of specific antigens (Wright 1992, 1994).

Microarray Analysis

Rapid progress in genome sequencing efforts has engendered technological advances in bioinformatics through the accumulations of large databases aimed at unraveling the linkages between genetic potential and phenotypes. One of the major developments in this direction is the production of DNA microarrays for functional genomics analysis of gene expression. Similarly, innovations in instrumentation and experimental techniques have provided new ways for analysing the protein content of cells through differential display arrays (Barzaghi et al. 2004). Such innovations have greatly facilitated our ability to discover polypeptides that mediate cellular responses to various environmental stimuli, although major challenges still exist in applying these innovations to complex systems such as soils. Recently, the development of microarray systems based on miniature chips and mass spectrometry has facilitated the analysis of large numbers of protein molecules, under different experimental conditions, to provide relatively precise data on protein profiles (Ogunseitan et al. 2001; Schulze et al. 2005). The 2002 Nobel Prize in Chemistry was awarded jointly to John Fenn, Koichi Tanaka, and Kurt Wüthrich for developing sophisticated instrumental and methodological approaches for analysing proteins, including spectrometry and magnetic resonance imaging. These achievements have enhanced the recognition that proteome-based strategies will provide significant contributions in the post-genomic research era (Wüthrich 2001; The Nobel Prize in Chemistry 2002).

Technologies for proteome analysis are intrinsically more complex than those used for nucleic acid analysis. Although this is partly due to the larger number of amino acids (20) as compared to the number of nucleic acid bases (5), the greatest complication comes from the post-translational modifications which endow each category of polypeptide chain with a unique

set of characteristics (Smith et al. 2002). Microarray systems combine multiple technologies in a unit platform that facilitates protein immobilisation, purification, analysis, and processing from complex biological mixtures. In particular, the development of surface-enhanced laser desorption/ionisation time-of-flight (SELDI-TOF) mass analysis has increased experimental options for differential display analysis of proteomes for organisms cultivated under different environmental conditions. Developments in coupled mass spectrometric analysis of proteomes in microarray formats provide new opportunities for ecologically relevant soil proteome assessment, including spatial-temporal and quantitative mapping of protein involvement in organism–environment interactions, and in cell–cell communication (Ogunseitan et al. 2002; Ogunseitan and LeBlanc 2004).

The SELDI-TOF process has four major components. The first component requires the capturing of target molecules from the protein extract onto the microarray with minimal sample preparation. The second component involves the enhancement of the signal-to-background ratio through the use of selective "washing" protocols. The third component deals with the identification of remainder protein molecules on the array by means of a laser-induced process to generate a protein mass profile. Finally, target protein molecules can be characterised directly through modification reactions that provide information about protein structure and function. The information gained can thus be verified through comparison to an extensive database of protein sequence, size, conformation, and activity to identify known proteins or characterise entirely novel discoveries. In general, the arrays contain chemically treated surfaces that facilitate specific interactions with protein molecules, including cationic, anionic, hydrophobic, and hydrophilic reactions. Similarly, arrays may be biochemically configured with specific antibodies, cell-surface receptors, or nucleic acids, to identify heterogeneous molecular interactions (Ogunseitan and LeBlanc 2004). Here again, it is important to emphasise that successful application of these approaches depends on the quality of protein extracts derived from soils. Different soil types will invariably require different protein purification strategies to remove contaminants and co-extracting biochemicals that may interfere with molecular resolution and quantification procedures.

5.4 Case Studies and Emerging Issues in Soil Proteomics

The appeal of molecular techniques in soil science is attributable to their power in dissecting complex features to reveal component structures and functions that facilitate better understanding of how these components are interrelated; and how the relationships can be exploited to craft sustain-

able solutions to persistent environmental problems. This section focuses on two case studies where the use of protein extraction and analysis methods has led to new understanding of soil properties and how soil organisms cope with stressful conditions. The first case explores the role of glomalin in creating the features that support soil particle aggregation and carbon (C) storage. The second case involves the role of protein extraction and analytical methods in understanding how soil microbial communities respond to pollutants such as toxic metals. These cases are selected as possible extremes out of several possible general cases where protein analyses in soils have been successfully achieved because of the strengths of the lessons they bear for innovation in the midst of challenging methodology.

5.4.1 Glomalin

In 1996, Sara Wright and colleagues discovered an abundant soil protein that has been associated with several important soil properties and biogeochemical functions (Wright and Upadhyaya 1996). The protein, named glomalin for its association with arbuscular mycorrhizal fungi belonging to the order *Glomales*, is extremely stable in soils where it accumulates to several mg per g of soil, reaching up to 100 mg per g in some Hawaiian soils (Fig. 5.1). Two fractions of glomalin are known to be present in soils: the easily extractable glomalin (EEG), and total extractable glomalin (TEG). The extraction of glomalin from soils requires the unusual step of autoclaving. Typically for EEG, 1 g of soil is extracted with 8 ml aliquots of 20 mM sodium citrate under autoclaving conditions (15 psi pressure and 121 °C) for 30 min. For TEG, 50 mM sodium citrate is used, and the autoclaving time is doubled. It is noteworthy that citrate will also cause the co-extraction of other non-proteinous soil components that may interfere with the quantitative analysis of glomalin further on in the process. The extraction is repeated as long as no more brown-coloured glomalin is visible in the remaining particulate soil fraction. Centrifugation-clarified extracts of glomalin are stable as freeze-dried protein (Fig. 5.2).

Glomalin is known to be associated with carbohydrates (glycoproteinous) containing between 30–40% C by weight, and 1–9% of tightly bound iron, but, beyond that, not much else is known about its structure (Gonzalez-Chavez et al. 2004). In addition, neither the environmental cues for the production of glomalin nor its specific cellular biological role and processing in the fungi that produce it are well understood. However, because it is detectable by means of immunological techniques, several studies have demonstrated the ubiquitous occurrence of glomalin in soils

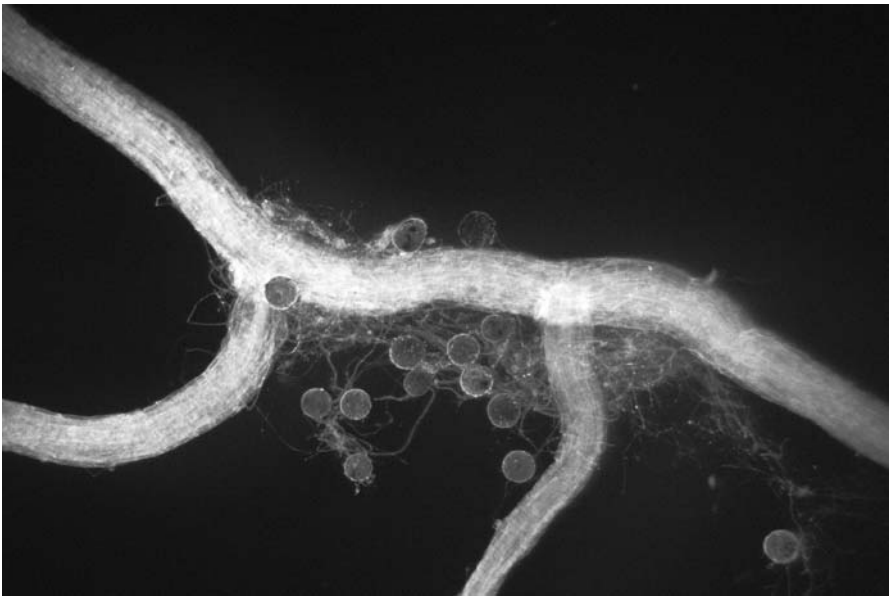

Fig. 5.1. View of glomalin revealed by immunological reactions using a green dye tagged to an antibody against glomalin on an arbuscular mycorrhizal fungus growing on a corn root. The *round structures* are spores, and the *threadlike filaments* are hyphae. Photograph by Sara Wright, United States Department of Agriculture

across many geographical zones and land-use patterns (Wright and Upadhyaya 1998; Rillig et al. 2002b; Knorr et al. 2003; Steinberg and Rillig 2003).

Several interesting ecological roles have been attributed to glomalin. These ecological functions include its role as "glue" between soil particles thereby explaining many features regarding soil texture, soil aggregation, water-holding capacity, autochthonous microbial diversity, and the stability or turnover rate of the soil C pool. Glomalin concentrations in soil also correlate with land-use activities, and differ between tilled or agricultural soils and forested regions, and between tropical and temperate climate soils (Rillig et al. 2003). Glomalin has been shown to account for a great proportion of the organic C in soils. Therefore, this protein is also a major component of soil organic matter, weighing 2 to 24 times more than humic acid, which was previously thought to be the main contributor (approximately 8%) to soil C. Rillig and colleagues (2003) have estimated that glomalin molecules survive in soils for 7–42 years, depending on soil temperature, humidity conditions, and other physicochemical properties of soils. Apparently, higher levels of atmospheric carbon dioxide stimulate the growth of glomalin-producing fungi, and consequently the level of glo-

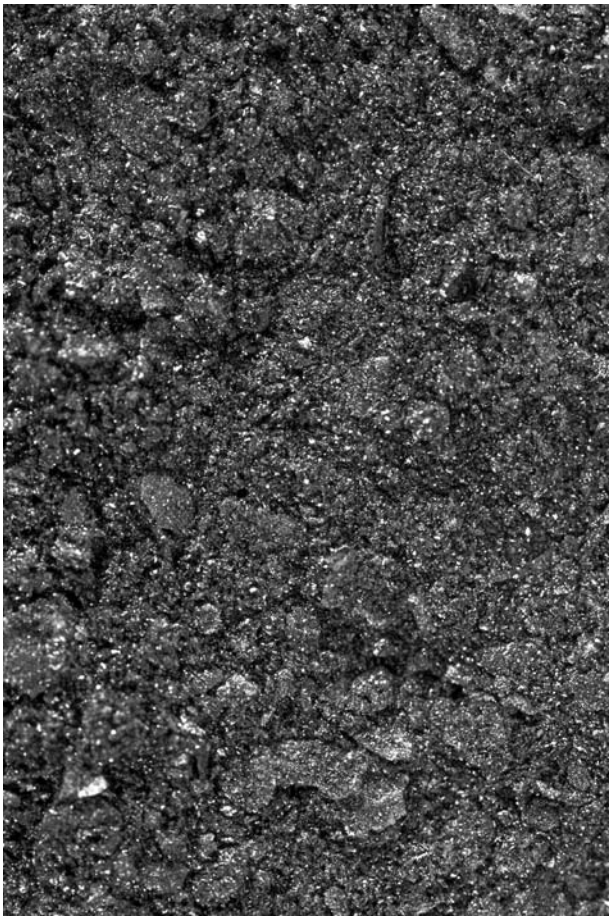

Fig. 5.2. Freeze-dried glomalin extracted from undisturbed soil, where it is present in concentrations of several milligrams per gram of soil. Photograph by Keith Weller, United States Department of Agriculture

malin found in soils. This and other observations have raised the profile of glomalin management in soils as a potential indicator of the ecological effects of global climate change (Rillig et al. 2002a).

Apparently much remains to be learned about the molecular details of the glomalin structure, the diversity of organisms capable of producing glomalin, and the diversity of organisms capable of degrading the protein. These challenges will require the orchestration of a wide variety of the proteomics techniques discussed in the first sections of this chapter. In particular, glomalin extracts need to be subjected to molecular mass fingerprinting techniques such as the protein-chip system based on SELDI-TOF

technology. This approach could reveal the fine-scale molecular diversity expected by the differences in the extraction procedure attributable to the easily extractable and total extractable glomalin.

5.4.2 Soil Proteins as Metal Biosensors

It has been estimated that at least 30% of all proteins are metalloproteins, requiring one or more metal ions for proper functioning (Andreini et al. 2004; for examples, see Table 5.1). The specificity of protein–metal complexes and the affinity (and stoichiometric interaction) of proteins for metals in general provide an opportunity to detect or to determine biological availability of metal ions by assaying protein functions including enzymatic activity (Haraguchi 2004). Theoretically, two major approaches are possible for developing protein-based strategies for indicating the biological availability of metallic contaminants in soil. The first approach relies on the comparative assessment of the profiles of polypeptide molecules extracted directly from soils in the presence and absence of the pollutant in question. The robustness of such differential displays depends on pattern recognition techniques, and is often difficult to interpret because of the labile nature of protein expression in organisms. The second, rather traditional approach is more narrowly defined, seeking to detect changes in the expression and activity of a particular soil protein in response to metal exposure (Tyler 1974). For example, it is possible to correlate delta-aminolevulinate dehydratase activity with the concentration of biologically available lead (Pb) in contaminated systems (Ogunseitan et al. 2000).

As described in Sect. 5.4.1, direct extraction methods have been used to investigate the contribution of extracellular fungal proteins to soil stabilisation (Wright and Upadhyaya 1996). Yields between 1 and 100 mg protein per g of soil have been reported by investigators employing prolonged heat extraction to generate proteins detectable by enzyme-linked immunosorbent assay (ELISA) from agricultural soils (Rillig et al. 2003). Conversely, methods aimed at protecting protein function clearly cannot employ the autoclave, and consequently these methods typically achieve lower protein yields. In situations where a specific group of proteins is targeted, it is worthwhile to use extraction conditions such as buffer-pH and temperature that maximise protein stability, although it must be emphasised that most extraction techniques are only relatively successful with respect to soil chemistry and biological composition. Consequently, decisions about trade-offs must be made between optimising extraction effectiveness and representativeness. For example, in vitro studies have used deoxycholate to extract ribosomal proteins, and urea pre-extraction has been used for

Table 5.1. Examples of metals required by key enzyme functions. The biological availability of metal ions affects enzyme function, potentially constituting the basis for the development of protein-based biosensors

Metal	Enzyme examples	Relevant organisms and ecologicalfunction
Calcium	Collagenase Calpain protease	*Clostridium histolyticum* Pathogenesis
Cobalt	Halomethane methyltransferase	Facultative methylotrophs Chloromethane degradation
Copper	Copper amine oxidases	*Arthrobacter globoformis* Deamination of primary amines to corresponding aldehydes
Iron	Cytochrome P450 Soluble Methane Monooxygenase	*Pseudomonas putida.* P450cam Polyaromatic hydrocarbon biodegradation
Magnesium	Magnesium chelatase	*Rhodobacter sphaeroides* Chlorophyll synthesis
Manganese	Manganese-dependent peroxidases	*Phanerochaete chrysosporium* Carbon cycling/lignin degradation
Molybdenum	Xanthine dehydrogenase Nitrate reductase; nitrogenase	*Pseudomonas putida* *Rhizobium* Purine (caffeine) degradation Nitrogen fixation
Selenium	Selenocysteine (UGA codon) enzymes; glutathione peroxidase; hydrogenases	*Escherichia coli* *Methanococcus voltae* Nitrogen respiration
Tungsten	"True W-enzymes" Aldehyde ferredoxin oxidoreductase	*Thermophilic Archaea*; *Pyrococcus furiosus* Molybdenum antagonist
Zinc	Aminolevulinate dehydratase (inhibited by Pb)	Several bacteria and nearly all archaea Porphyrin synthesis

bacterial membrane proteins (Deutscher 1990). Systematic studies of such protein fractions, particularly for enzymes in natural environments, have yet to be standardised, but certainly represent a field in microbial ecology that is ripe for further investigation.

Singleton and colleagues (2003) successfully demonstrated that the presence of cadmium (Cd) in soils affects both the concentration of total proteins and the size distribution of proteins extracted from soils. In this regard, soil protein content was found to be as sensitive as measurements of soil microbial biomass in indicating soil ecosystem stress from toxic metals (Sandaa et al. 2001). The increase in the production of small molecular weight proteins (< 21 kDa) in response to Cd exposure may be related

to the observation in laboratory cultures that organisms synthesise metallothionein and similar proteins as a protective measure against metal toxicity. However, a validation of the putative correlation between bulk protein expression and metal concentration will require further research. It is quite possible that such correlations will prove difficult to establish with confidence given the complex variability of physical and chemical parameters of soils. In such cases, establishing and integrating the responses of several key individual soil organisms may provide a feasible strategy for approximating community-level responses. However, challenges posed by the enormity of the statistical processing capacity needed for such computations may prove insurmountable. Instead, a compartmentalised approach where pools of active and inactive microbial populations are evaluated to derive approximations of overall soil processes is perhaps inevitable.

We have used the SELDI-TOF technology to investigate the response of individual microbial protein profiles to toxic metal exposure (Ogunseitan et al. 2002). We focused specifically on mercury and lead because these metals are not only capable of inducing genetic expression, but they also have the capacity to destabilise cellular metabolism by inhibiting the functions of several enzymes (Ogunseitan et al. 2000). Specific heavy metal resistance systems have evolved in the microbial communities (Ogunseitan et al. 1999), but several observations have also demonstrated that genetically encoded metal resistance is not found in all organisms inhabiting a metal-contaminated ecosystem (Ogunseitan 1998). Metal-resistant organisms typically exhibit a concentration threshold depending on whether the resistance mechanism is based on limiting metal ion transportation from outside to inside the cell, or on enzymatic transformation of metals after active transportation into the cell, as mediated by mercuric reductase in the case of mercuric ions (Ogunseitan 1998). These ecological contingencies on microbial interactions with toxic metals in the environment provide a rich case study for analysing the whole proteome in order to understand how organisms cope with environmental stress at the molecular level.

5.4.3 Prospects for Proteomic Analysis of Soil Microbial Communities

Research in the newly emerging field of soil proteomics proteins has aligned with three important areas of investigation. The most active of these is focused on cell–environment interactions, including protein-based biosensors, ectoenzyme activities, stress proteins, and the induction of metabolic proteins by fluctuating concentrations of environmental chemicals. The second area involves the analysis of homologous cell-to-cell interactions, including quorum sensing, genetic exchange activities, and secondary

structure or colony formation. The third area of focus is the investigation of heterologous cell-to-cell interactions, involving the production of bioactive peptides that define competition, predation, commensalisms, and symbiosis in the natural microbial community (Ogunseitan et al. 2002).

Although great strides are being made in these areas of focus within microbial environmental proteomics research, there are three major challenges for future developments toward the analysis of soil microbial community functions through protein assessment. The first challenge concerns the exhaustive recovery of proteins from heterogeneous populations inhabiting proximate habitats. Traditionally, protein extraction methods have been developed to optimise the selective recovery of active proteins through non-destructive methods such as sonication or to maximise the recovery of total proteins by harsh cell lysis conditions such as heat treatment with detergents. In addition, methods have been developed to recover extracellular proteins through soils washing, and the use of differential centrifugation, thereby avoiding recovery of particulate proteins (Ogunseitan 1997). A complete proteome assessment should rely on methodological integration that requires the optimisation of protein recovery while preserving functional characteristics. There is no doubt that trade-offs must be made between these two objectives, depending on particular soil characteristics and the stated goal of research topics.

The second challenge concerns the resolution of independent protein molecules extracted from microbial populations in a manner that accommodates the dynamic nature of gene expression (including polymorphisms), protein synthesis and protein degradation. The use of matrix-assisted laser desorption or electrospray ionisation mass spectrometry has achieved enhanced resolution, and, theoretically, up to 30,000 polypeptides can be systematically resolved through these techniques. However, given the size of microbial genomes and the very high diversity that exists in most soil microbial communities, only a small fraction of existing polypeptides can be resolved even by these high-throughput techniques. The methodological matrix of resolving conditions associated with SELDI-TOF can assist in a stepwise protein resolution scheme that will increase the coverage and resolving power of community proteomes, but integrating the data requires extensive computing and numerous controls. Thus, the third challenge lies in the development of new bio-informatics techniques for processing large databases of protein mass and function data. Neural networks developed for pattern recognition of protein profiles show good promise in this direction (Ogunseitan 2002).

Ultimately, as in other fields of molecular biology, there is reasonable optimism that technical developments will catch up with the rapid pace at which key research questions are being framed within soil science that requires the application of proteomics. This chapter explored recent

methodological developments and salient case studies. These and other path-defining explorations in soil proteomics will continue to contribute significantly to our understanding of phenomena such as soil structure, biodiversity, and biogeochemical cycling, including the response of soil systems to toxic pollutants.

Acknowledgements. Research in my laboratory group is supported in part by awards from the National Science Foundation (DEB 95-24481 and DMI 02-23894), the Josiah Macy Jr. Foundation, the University of California Toxic Substances Research and Teaching Program, and the Program in Industrial Ecology at UC-Irvine.

References

Andreini C, Bertini I, Rosato A (2004) A hint to search for metalloproteins in gene banks. Bioinformat 20:1373–1380

Barzaghi D, Isbister JD, Lauer KP, Born TL (2004) Use of surface-enhanced laser desorption/ionization-time of flight to explore bacterial proteomes. Proteomics 4:2624–2628

Blom A, Harder W, Matin A (1992) Unique and overlapping stress proteins of *Escherichia coli*. Appl Environ Microbiol 58:331–334

Bohlool BB, Schmidt EL (1980) The immunofluorescence approach in microbial ecology. Adv Microbial Ecol 4:203–241

Bolygo E, Ricketts D, Moffatt F, Jessop M, Attenborough S (2003) Determination of antifungal proteins in soil by liquid chromatography. Anal Bioanal Chem 376:701–705

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-pye binding. Anal Biochem 72:248–254

Bremner JM, Zantua MI (1975) Enzyme activity in soils at subzero temperatures. Soil Biol Biochem 7:383–387

Briggs MH, Segal J (1963) Preparation and properties of a free soil enzyme. Life Sci 2:69–72

Burns RG (1978) Enzyme activity in soil: some theoretical and practical considerations. In: Burns RG (ed) Soil ernzymes. Academic Press, London, pp 295–340

Burns RG, El-Sayed MH, McLaren AD (1972) Extraction of an urease-active organo-complex from soil. Soil Biol Biochem 4:107–108

Busto MD, Ortega N, Perez-Mateos M (1995) Induction of β-galactosidase in fungal and soil bacterial cultures. Soil Biol Biochem 27:949–954

Cheng C-N (1975) Extracting and desalting amino acids from soils and sediments: evaluation of methods. Soil Biol Biochem 7:319–322

Deutscher MP (ed) (1990) Guide to protein purification. Methods Enzymol 182:68–83 and 285–421

Eprogen Corporation (2004) ProteoSep kit lysis protocol. http://www.eprogen.com/proteomics/literature/ProteoSepLysisProtocol_0816021.pdf

Feinstein RN, Lindahl R (1973) Detection of oxidases on polyacrylamide gels. Anal Biochem 56:353–360

Gonzalez-Chavez MC, Carrillo-Gonzalez R, Wright SF, Nichols KA (2004) The role of glomalin, a protein produced by arbuscular mycorrhizal fungi, in sequestering potentially toxic elements. Environ Pollut 130:317–323

Greenfield LJ, Hamilton RD, Weiner C (1970) Nondestructive determination of protein, total amino acids, and ammonia in marine sediments. Bull Mar Sci 20:289–304
Hantula JA, Korhonen TK, Bamford DH (1990) Determination of taxonomic resolution power of SDS-polyacrylamide gel electrophoresis of total cellular proteins using *Enterobacteriaceae*. FEMS Microbiol Lett 70:325–330
Hantula J, Kurki A, Vuoriranta P, Bamford DH (1991) Rapid classification of bacterial strains by SDS-polyacrylamide gel electrophoresis: population dynamics of the dominant dispersed phase bacteria of activated sludge. Appl Microbiol Biotechnol 34:551–555
Haraguchi H (2004) Metallomics as integrated biometal science. J Anal At Spectrom 19:5–14
Harner MJ, Ramsey PW, Rillig MC (2004) Protein accumulation and distribution in floodplain soils and river foam. Ecol Lett 7:829–836
Jackman PJH (1985) Bacterial taxonomy based on electrophoretic whole-cell protein patterns. In: Goodfellow M, Minnikin DE (eds) Chemical methods in bacterial systematics. Academic Press, London, pp 115–129
Jehr MA, Hussain Z (1994) Multilocus enzyme electrophoresis using ultrathin polyacrylamide gels. J Microbiol Methods 19:307–313
Kersters K, De Ley J (1980) Classification and identification of bacteria by electrophoresis of their proteins. In: Goodfellow M, Board RG (eds) Microbiological classification and identification. Academic Press, London, p 273
Knorr MA, Boerner REJ, Rillig MC (2003) Glomalin content of forest soils in relation to fire frequency and landscape position. Mycorrhiza 13:205–210
Lee L, Saxena D, Stotzky G (2003) Activity of free and clay-bound insecticidal proteins from *Bacillus thuringiensis* subsp. *israelensis* against the mosquito *Culex pipiens*. Appl Environ Microbiol 69:4111–4115
Lindahl V, Bakken LR (1995) Evaluation of methods for extraction of bacteria from soil. FEMS Microbiol Ecol 16:135–142
Liu W-T, Stahl DA (2002) Molecular approaches for the measurement of density, diversity, and phylogeny. In: Hurst CJ, Crawford RL, Knudsen R, McInerney MJ, Stetzenbach LD (eds) Manual of environmental microbiology, 2nd edn. ASM Press, Washington, DC, pp 114–134
Manchenko GP (1994) Handbook of detection of enzymes on electrophoretic gels. CRC Press, Boca Raton, FL, 300 pp
Mayaudon J (1986) The role of carbohydrates in the free enzymes in soil. In: Fuchsman CH (ed) Peat and water. Elsevier Applied Science, New York, pp 263–309
Mayaudon J, Sarkar JM (1974) Chromatography and purification of diphenol oxidases from soil. Soil Biol Biochem 6:275–285
Mayaudon J, Batistic L, Sarkar JM (1975) Properties of proteolytic activities extracted from fresh soils. Soil Biol Biochem 7:281–286
McLaren AD, Pukite AH, Barshad I (1975) Isolation of humus with enzymatic activity from soil. Soil Sci 119:178–180
Nannipieri P, Ceccanti B, Cervelli S, Sequi P (1974) Use of 0.1 M pyrophosphate to extract urease from podzol. Soil Biol Biochem 6:359–362
Nannipieri P, Sequi P, Fusi P (1996) Humus and enzyme activity. In: Piccolo A (ed) Humic substances in terrestrial ecosystems. Elsevier, Amsterdam, pp 293–328
Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003) Microbial diversity and soil functions. Eur J Soil Sci 54:655–670
Nybroe O, Johansen A, Laake M (1990) Enzyme-linked immunosorbent assays for detection of *Pseudomonas fluorescens* in sediment samples. Lett Appl Microbiol 11:293–296
Ogunseitan OA (1993) Direct extraction of proteins from environmental samples. J Microbiol Methods 17:273–281

Ogunseitan OA (1996) Protein profile variation in cultivated and native freshwater microorganisms exposed to chemical environmental pollutants. Microb Ecol 31:291–304
Ogunseitan OA (1997) Direct extraction of catalytic proteins from natural microbial communities. J Microbiol Methods 28:55–63
Ogunseitan OA (1998) Protein method for investigating mercuric reductase gene expression in aquatic environments. Appl Environ Microbiol 64:191–204
Ogunseitan OA (1999) Microbial proteins as biomarkers of ecosystem health. In: Scow K et al. (eds) Integrated assessment of ecosystem health. CRC Press, Boca Raton, FL, pp 207–222
Ogunseitan OA (2002) Assessing microbial proteomes in the environment. In: Bitton G (ed) Encyclopedia of environmental microbiology. John Wiley and Sons, Inc., New York, pp 305–312
Ogunseitan OA (2005) Microbial diversity. Blackwell Science, Oxford, UK, pp 292
Ogunseitan OA, LeBlanc JF (2004) Environmental proteomics: methods and applications for aquatic ecosystems. In: Kowalchuk GA, Bruijn FJd, Head IM, Akkermans AD, Van Elsas JD (eds) Molecular microbial ecology manual, 2nd edn. Springer, Berlin Heidelberg New York, pp 1027–1046
Ogunseitan OA, Yang SL, Scheinbach E (1999) The delta-aminolevulinate dehydratase of marine *Vibrio alginolyticus* is resistant to lead (Pb). Biol Bull 197:283–284
Ogunseitan OA, Yang S, Ericson J (2000) Microbial delta-aminolevulinate dehydratase as a biosensor of lead bioavailability in contaminated environments. Soil Biol Biochem 32:1899–1906
Ogunseitan OA, LeBlanc J, Dalmasso E (2001) Microbial community proteomics. In: Rochelle PA (ed) Environmental molecular microbiology. Horizon Scientific Press, Norfolk, UK, pp 125–140
Ogunseitan OA, LeBlanc JF, Noble PA (2002) Ecological dimensions of microbial proteomics. Rec Res Develop Microbiol 6:487–501
Paul JH (1993) The advances and limitations of methodology. In: Ford TE (ed) Aquatic microbiology. Blackwell Scientific Publications, Boston, MA, pp 15–46
Pettit NM, Smith ARJ, Freedman RB, Burns RG (1976) Soil urease: activity, stability and kinetic properties. Soil Biol Biochem 8:479–484
Rillig MC, Wright SF, Shaw MR, Field CB (2002a) Artificial climate warming positively affects arbuscular mycorrhizae but decreases soil aggregate water stability in an annual grassland. Oikos 97:52–58
Rillig MC, Wright SF, Eviner VT (2002b) The role of arbuscular mycorrhizal fungi and glomalin in soil aggregation: comparing effects of five plant species. Plant Soil 238:325–333
Rillig MC, Ramsey PW, Morris S, Paul EA (2003) Glomalin, an arbuscular-mycorrhizal fungal soil protein, responds to land-use change. Plant Soil 253:293–299
Sandaa RA, Torsvik V, Enger O (2001) Influence of long-term heavy-metal contamination on microbial communities in soil. Soil Biol Biochem 33:287–295
Saxena D, Flores S, Stotzky G (2002) Vertical movement in soil of insecticidal Cry1Ab protein from *Bacillus thuringiensis*. Soil Biol Biochem 34:111–120
Schulze WX, Gleixner G, Kaiser K, Guggenberger G, Mann M, Schulze E-D (2005) A proteomic fingerprint of dissolved organic carbon and of soil particles. Oecologia 142:335–343
Selander RK, Caugant DA, Ochman H, Musser JM, Gilmour MN, Whittam TS (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics andsystematics. Appl Environ Microbiol 51:873–884
Shaw CR, Prasad R (1970) Starch gel electrophoresis of enzymes – a compilation of recipes. Biochem Genet 4:297–320

Singleton I, Merrington G, Colvan S, Delahunty JS (2003) The potential of soil protein-based methods to indicate metal contamination. Appl Soil Ecol 23:25–32
Skujins J (1976) Extracellular enzymes in soil. CRC Crit Rev Microbiol 4:383–421
Smith RD, Anderson GA, Lipton MS, Masselon C, Pasa-Tolic L, Shen Y, Udseth HR (2002) The use of accurate mass tags for high-throughput microbial proteomics. Omics J Integr Biol 6:61–90
Steinberg PD, Rillig MC (2003) Differential decomposition of arbuscular mycorrhizal fungal hyphae and glomalin. Soil Biol Biochem 35:191–194
Stemmer M (2004) Multiple-substrate enzyme assays: A useful approach for profiling enzyme activity in soils? Soil Biol Biochem 36:519–527
Tabatabai A, Fu M (1992) Extraction of enzymes from soils. Soil Biochem 7:197–227
The Nobel Prize in Chemistry (2002) http://www.nobel.se/chemistry/laureates/2002/index.html
Thornton JI, McLaren AD (1975) Enzymatic characterization of soil evidence. J Forensic Sci 20:674–700
Torsvik V, Sorheim R, Goksoyr J (1996) Total bacterial diversity in soil and sediment communities – a review. J Indust Microbiol 17:170–178
Torsvik V, Daae FL, Sandaa R-A, Ovreas L (1998) Novel techniques for analysing microbial diversity in natural and perturbed environments. J Biotechnol 64:53–62
Tyler G (1974) Heavy metal pollution and soil enzymatic activity. Plant Soil 41:303–308
Vettori C, Gallori E, Stotzky G (2000) Clay minerals protect bacteriophage PBS1 of *Bacillus subtilis* against inactivation and loss of transducing ability by UV radiation. Can J Microbiol 46:770–773
Weetall HH, Welicky N, Vango SP (1965) Detection of microorganisms in soil by their catalytic activity. Nature 206:1019
Wright SF (1992) Immunological techniques for detection, identification, and enumeration of microorganisms in the environment. In: Levin MA, Seidler RJ, Rogul M (eds) Microbial ecology principles, methods, and applications. McGraw-Hill, New York, pp 45–60
Wright SF (1994) Serology and conjugation of antibodies. In: Weaver RW, Angle S, Bottomley P et al. (eds) Methods of soil analysis, part 2. Microbiological and biochemical properties. Soil Science Society of America, Madison, WI, pp 593–617
Wright SF, Upadhyaya A (1996) Extraction of an abundant and unusual protein from soil and comparison with hyphal protein of arbuscular mycorrhizal fungi. Soil Sci 161:575–586
Wright SF, Upadhyaya A (1998) A survey of soils for aggregate stability and glomalin, a glycoprotein produced by hyphae of arbuscular mycorrhizal fungi. Plant Soil 198:97–107
Wright SF, Franke-Snyder M, Morton JB, Upadhyaya A (1996) Time-course study and partial characterization of a protein on hyphae of arbuscular mycorrhizal fungi during active colonization of roots. Plant Soil 181:193–203
Wüthrich K (2001) The way to NMR structures of proteins. Nat Struct Biol 8:923–925
Zor T, Selinger Z (1996) Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. Anal Biochem 236:302–308