Analysis of the microbial proteome

Michael P Washburn* and John R Yates III†

Proteomics has begun to provide insight into the biology of microorganisms. The combination of proteomics with genetics, molecular biology, protein biochemistry and biophysics is particularly powerful, resulting in novel methods to analyse complex protein mixtures. Emerging proteomic technologies promise to increase the throughput of protein identifications from complex mixtures and allow for the quantification of protein expression levels.

Addresses
*Department of Molecular Biotechnology, University of Washington, Seattle, WA 98195, USA
†Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92937, USA; e-mail: jyates@scripps.edu

Current Opinion in Microbiology 2000, **3**:292–297

1369-5274/00/$ – see front matter

Abbreviations
2D two-dimensional
MS mass spectrometry

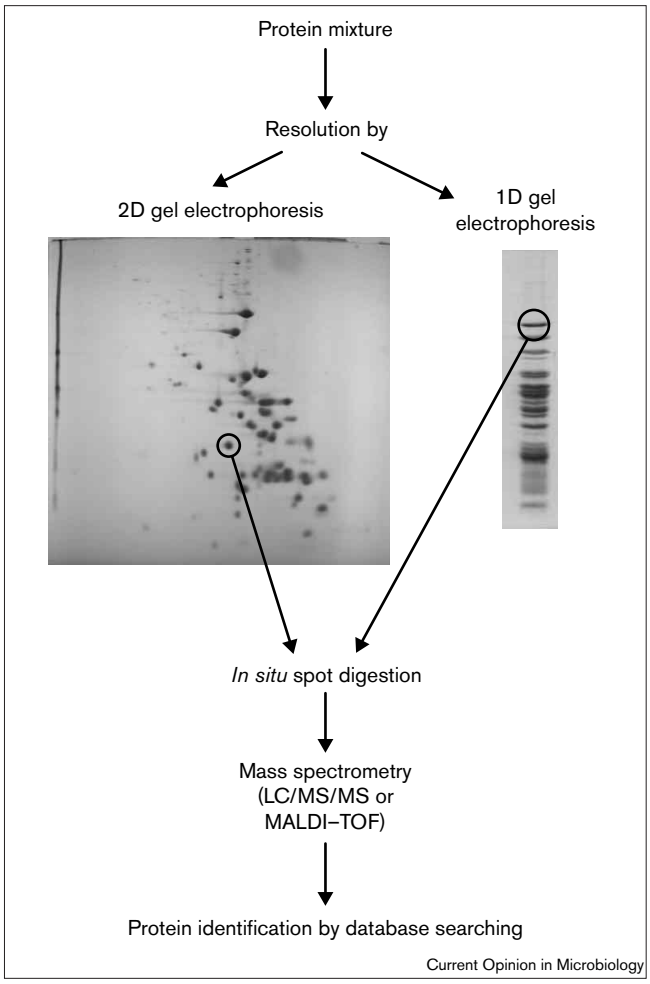
Introduction

In the post-genomic era, innovative methods of analysis are needed to determine protein functional information on a large scale with high throughput. Proteomics evolved from this need, resulting in novel investigations into the expressed proteins of an organism. The characterization of a proteome by itself is merely the first step. Combining proteomic technologies with genetics, molecular biology, protein biochemistry and/or biophysics has resulted in accelerated discovery of protein functional information. A standard proteomic analysis of an organism begins with the isolation of its soluble proteins, followed by the separation and visualization of the protein mixture by two-dimensional (2D) gel electrophoresis (Figure 1). Although powerful, there are several limitations to 2D gel electrophoresis. Besides being time-consuming and laborious, 2D gels rarely identify low-abundance and hydrophobic proteins. Methods of proteomic analysis are being developed to gain access to these classes of proteins and to accelerate the speed of discovery. Finally, quantitative methods are being developed, allowing for relative protein abundances to be determined. This review presents and discusses a wide variety of proteomic applications to microbiology from the past year. In addition, novel high throughput separation and quantitation methods are discussed which may revolutionize proteomics.

2D gel electrophoresis and microbial proteomics

The basic goal of a proteomic study is to identify proteins in an organism involved in a particular process. Generally, an organism is grown under different experimental conditions and abundant proteins of each sample are resolved on separate 2D gels (Figure 1). After comparison of the gels, the protein identities of selected spots are determined, providing insight into the protein expression changes resulting from the experimental conditions (Figure 1). Using this methodology, protein functional groups have been identified from *Bradyrhizobium japonicum* grown under anaerobic and aerobic conditions [1] and stressed by heat shock [2•]. In addition, Guerrerio *et al.* [3] determined the changes in protein expression patterns of

Figure 1



Standard proteomic analysis of a protein mixture. After a complex protein mixture, such as a protein complex or whole-cell lysate, is prepared, the proteins are separated and visualized by either 1D or 2D gel electrophoresis. Individual spots from the gel are excised and digested. After extracting the peptides from the gel matrix, the spots are analysed by matrix-assisted laser desorption ionization–time-of-flight (MALDI–TOF) mass spectrometry or capillary liquid chromatography, followed by tandem mass spectrometry (LC/MS/MS). The protein identities are determined in MALDI–TOF by peptide mass mapping, and in LC/MS/MS by correlating tandem mass spectra of fragmented peptides to protein sequence databases.

Sinorrhizobium meliloti from early exponential phase to late exponential phase, and Vasseur *et al.* [4] investigated the alterations in protein expression of *Pseudomonas fragi* grown under a variety of stressful conditions, including osmotic shock, pH, biocide and combined treatments. Successfully yielding novel insights into the biology of microorganisms, a review of the applications of this methodology to several microbial organisms has recently been published [5•].

The emergence of antibiotic resistance in an alarming number of human pathogens has necessitated research into microbial pathogenicity. Proteomics has facilitated the determination of changes in an organism that lead to pathogenicity, antibiotic resistance and/or an immune response. Proteomic analyses of *Candida albicans* pathogenicity have recently been reviewed [6] and additional research has identified several immunogenic proteins of *C. albicans* in human sera [7]. Other investigators have identified immunoreactive proteins of *Chlamydia trachomatis* [8] and alterations in protein expression levels in *Streptococcus pneumoniae* resulting from erythromycin resistance [9]. In comparative proteome analysis, the proteomes of different strains of micro-organisms are resolved and compared. When applied to *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains, proteins differing in intensity or position were identified, providing insight into the differences in pathogenicity of these two *Mycobaterium* strains [10••]. The goal of these applications of proteomics is to identify proteins involved in the disease process, hopefully leading to the development of medical intervention.

Proteomics combined with modern biological tools

The real power of proteomics arises when proteomics is combined with genetics, molecular biology, protein biochemistry and/or biophysics. Combining proteomics with molecular biology and yeast genetics, Lee *et al.* [11••] deciphered the separate roles Yap1 and Skn7 play in the oxidative stress response, and Alms *et al.* [12•] identified physiological phosphosubstrates of the phosphorylated protein phosphatase 1-phosphorylated Reg1 complex, a complex involved in the regulation of glucose repression pathways. Additional studies utilizing proteomic technologies, protein complex purification and yeast genetics identified proteins involved in actin assembly and cross-linking [13], a novel component of the mitochondrial nucleoli required for mitochondrial DNA repair [14], and components of a yeast histone acetyltransferase complex [15]. Rigaut *et al.* [16•] developed a generic protein complex purification method in yeast that they have applied to characterize yeast small nuclear ribonucleic acid particles, complexes involved in mRNA splicing [17,18] and the DEAD-box protein Dbp5 [19]. By combining proteomics and structural biology, Houry *et al.* [20••] identified 52 of the *in vivo* substrates of the *E. coli* chaperonin GroEL, a protein involved in the folding of newly translated polypeptides, and determined the common structural motifs of 24 of the protein substrates. Without proteomics, studies of large protein complexes on the scale undertaken by Houry *et al.* are not possible.

Protein databases and proteomics

Perhaps the most detailed proteomic analysis has been applied to *Saccharomyces cerevisiae*. The most recent update of its proteome stands at 279 different gene products from 401 different spots on a 2D gel [21]. Because of the amount of information known about *S. cerevisiae*, there are databases cataloging all of the known yeast proteins into functional, protein and subcellular localization classes (MIPS [22•] and the Yeast Proteome Database [23•]). A proteomic analysis of *S. cerevisiae* can then be easily correlated to these classes, rapidly identifying protein classifications.

Utilizing the detailed genomic and proteomic information available regarding *S. cerevisiae*, investigators have analysed the correlation between mRNA and protein abundance. On the basis of a logarithmic analysis, Futcher *et al.* [24••] discovered that the protein abundances of 148 proteins correlated with mRNA abundance and their respective codon adaptation index values. Future studies attempting to correlate mRNA, codon adaptation index values and protein expression levels in other organisms will provide valuable insight into the organization and expression of genomes and may result in the elucidation of common rules governing protein expression.

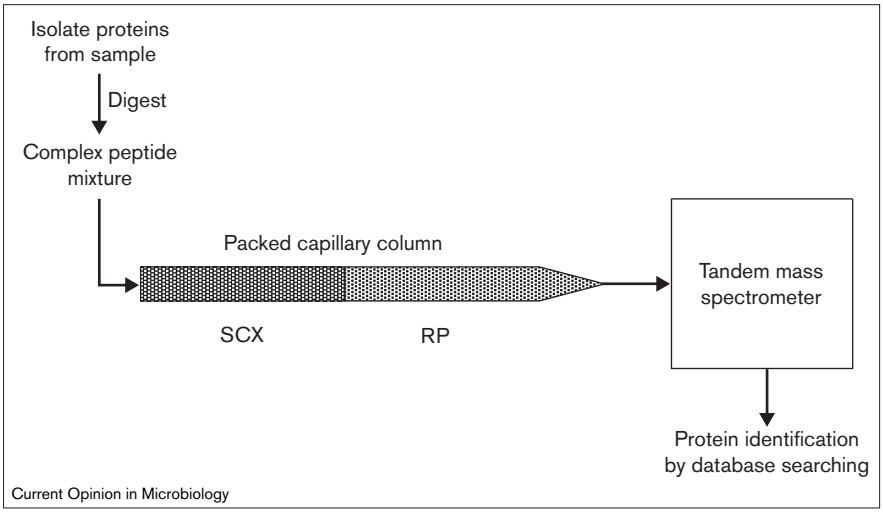
Proteomics on the internet

As proteomes of organisms continue to be characterized, there is a major effort to make this information readily available on the Internet. Recent updates of many microbial proteomes, including those of the cyanobacterium *Synechocystis* [25], *S. cerevisiae* [21], *M. tuberculosis* and *M. bovis* [26,10••], are all available as interactive web databases. Cordwell *et al.* [27] have recently described an approach to integrate all the information generated at the Australian Protoeme Analysis Facility for *E. coli*, *Pseudomonas aeruginosa* and *Staphylococcus aureus*. After many proteomic analyses not all of the information can be presented in a publication and must be available in some other form — the Internet is an ideal tool for this purpose.

Shortcomings of 2D gel electrophoresis

In spite of the enormous value and amount of information that 2D gel electrophoresis has generated in the field of proteomics, there exist many shortcomings of this technology. Even though up to 1400 spots can be resolved on a gel, in every study to date many more spots are resolved than are actually identified. For example, in the study correlating mRNA abundance and protein expression in *S. cerevisiae*, Futcher *et al.* [24••] resolved 1400 spots, but only identified 148 different proteins from 169 spots. With a genome of approximately 6000 proteins, their conclusions were based on 2.5% of the yeast proteome. Analysis of 1400 spots is time-consuming because each spot must be extracted, digested and analysed individually.

Figure 2



Proteomic analysis by multidimensional protein identification technology [37••]. A complex mixture of proteins is digested and directly loaded onto a capillary column packed with strong cation exchange (SCX) and reverse phase (RP) column matrix material. Peptides are sequentially displaced form the SCX phase into the RP, where they are eluted directly into a mass spectrometer. Protein identities are determined by correlating the tandem mass spectra of fragmented peptides to protein sequence databases.

Significant portions of proteomes, especially low-abundance proteins and membrane-associated or -bound proteins, are rarely seen in a 2D gel electrophoresis study. Several attempts have been made to enrich a fraction of a proteome using pre-gel chromatography, to identify low-abundance proteins. Fountoulakis *et al.* [28] loaded a hydroxyapatite column with *E. coli* cell lysate and identified 269 proteins from 800 spots from the column eluate. Of these, 130 proteins, including several low-abundance proteins, had not been previously detected by 2D gel electrophoresis [28]. In an additional study, Fountoulakis *et al.* [29] enriched a *Haemophilus influenzae* sample for low-abundance proteins using hydrophobic interaction chromatography. In both cases, only proteins that bound to the column matrix were enriched, and this included both high- and low-abundance proteins. Although targeted enrichment is promising, a general methodology should be able to detect and identify larger numbers of low-abundance proteins.

Problems with solubilization of entire membrane proteins prior to loading onto a 2D gel prevent their identification, and recent advances in membrane-protein solubilization for proteomics have been reviewed [30]. Promising methods include extracting *E. coli* membrane proteins with organic solvents, followed by solubilization with detergents prior to loading onto a 2D gel [31] and synthesizing novel zwitterionic detergents designed to work in conjunction with strong chaotropic agents [32]. Very few membrane proteins have been identified on 2D gels to date, and this is an area in need of further development.

Emerging separation methods for proteins

The difficulties of detecting and identifying low-abundance and membrane-associated proteins have necessitated the development of novel technologies for the resolution of these proteins. These methodologies are proving to be both reproducible and higher in throughput than 2D gel electrophoresis. Capillary isoelectric focusing (for a review, see [33]) coupled with mass spectrometry (MS) has been of limited use in separating complex mixtures of the order of entire proteomes [34] but may potentially allow for rapid identification of up to 1000 proteins [35]. When applied to smaller complex protein mixtures, such as the ribosome, capillary electrophoresis with solid-phase extraction, coupled with MS, detected and identified 80–90% of the *S. cerevisiae* ribosome [36].

In an alternate approach, a digested complex is loaded directly onto packed capillary columns that elute directly into an electrospray ionization ion-trap mass spectrometer (Figure 2). The capillary columns are packed with strong cation exchange and reverse phase matrix material (Figure 2). The peptides are sequentially eluted off to the capillary column and fragmented in the mass spectrometer, and advanced search algorithms then match the fragmented peptides to their respective proteins in a database. Using this approach, Link *et al.* [37••] identified 75 of the 78 predicted proteins from purified yeast ribosomes and discovered a novel protein in the yeast ribosome. Furthermore, they identified 189 unique proteins from a yeast whole cell lysate [37••]. Already less time-consuming than spot-by-spot identification, further optimization of this system may allow for up to 2000 proteins to be detected and identified in a single experiment.

Quantitative proteomics

Two new quantitative proteomic methods have been developed in the past year (reviewed in [38•]). Pása-Tolic′ *et al.* [39] and Oda *et al.* [40•] developed methods to metabolically label proteins by growing cells with either ^{15}N or ^{14}N as the sole nitrogen source. Identical proteins from each sample are either 'heavy' or 'light', allowing for the relative quantity of a protein between the two samples to

be determined by MS. Oda *et al.* [40•] grew two yeast strains, one knocked out for the G1 cyclin Cln2, under identical conditions (except that the only nitrogen source for one of the strains was ^{15}N). After pooling the two samples, the whole-cell extract was fractionated by high-pressure liquid chromatography and resolved by sodium dodecye sulfate/polyacrylamide gel electrophoresis. Bands from the gel were excised and digested, and the protein's identities were determined by MS. The mass differences of identical peptides from each sample allowed Oda *et al.* to determine the relative abundance of 12 proteins in each sample [40•]. Although powerful, the shortcoming of metabolic labeling is that it is limited to species that can be grown under such conditions.

Gygi *et al.* [41••] have developed a general quantitation method for proteomics in which cysteine residue of proteins were isotopically labeled after cell growth (Figure 3, [41••]). After growing yeast on either ethanol or galactose (as the carbon source), the proteins from each growth condition were isolated and the cysteine residue were modified with either a 'heavy' or 'light' reagent (Figure 3). Gygi *et al.* [41••] then determined the percentage changes in specific protein levels resulting from the different carbon sources. Unlike metabolic labeling strategies, this method can be applied to any system because the proteins are modified after cell growth.

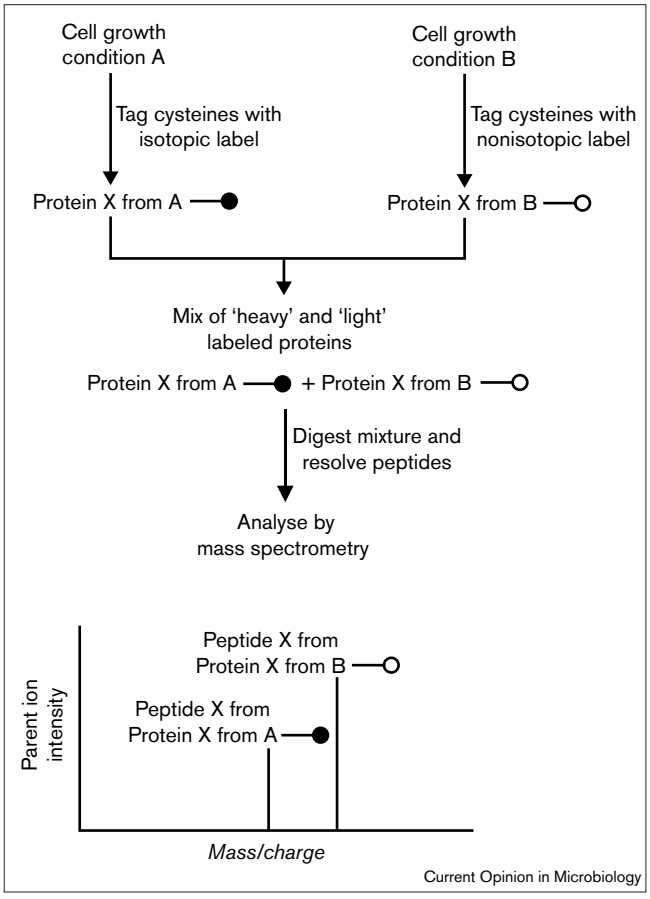
Conclusions

Proteomics has allowed novel investigation of the biology of micro-organisms. The combination of proteomics with other tools available to the modern biologist is an especially powerful approach. However, there is still a need for further technological development, and the current technology is under utilized. For example, proteomic approaches to detecting and identifying post-translational modifications are just beginning to emerge with new software developments [42•]. Two proteomic analyses of phosphorylation in mouse fibroblasts have been carried out [43,44], but this is a dramatically under-represented area, especially in microbial proteomics. Owing to the shortcomings of 2D gel electrophoresis, technological advances, such as capillary electrophoresis coupled to MS, are being pursued to dramatically increase the throughput of a proteome analysis. Finally, with the advent of quantitative proteomics, novel research analysing the alterations in cellular protein levels is possible. In the future, when the proteomics researcher is able to determine the relative quantity of 2000 proteins in a sample, proteomics will become as important to biology as quantitative analysis of gene expression.

Update

See [45••, 46•] for two recently published papers. Rout *et al.* [45••] combine proteomic methods with molecular biology and immunolocalization to map the nuclear pore complex in yeast and propose a detailed transport mechanism. Langen *et al.* [46•] use a comprehensive 2D gel/proteomics project to detect and identify 502 proteins from the proteome of *Haemophilus influenzae*.

Figure 3



Quantification of differential protein expression by isotopic labeling [41••]. Cells are grown under two different conditions expected to alter protein expression levels. The cysteine residues of the proteins from each cell growth condition are tagged with either an isotopic or nonisotopic label. After the tagging, the two fractions are mixed, digested and analysed by MS. The peak ratios of identical peptides from each growth condition yield relative quantification information. The tandem mass spectra of the peptides are matched against predicted spectra from protein sequence databases yielding the protein's identity.

Acknowledgements

The authors would like to thank David Tabb and David Schieltz for critical review of the manuscript. The authors would like to acknowledge funding from NIH grants R33CA81665-01 and RR11823-03. MP Washburn would like to acknowledge support from a NIH genome training grant (T32HG000035-05).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Dainese-Hatt P, Fischer H-M, Hennecke H, James P: **Classifying symbiotic proteins from *Bradyrhizobium japonicum* into functional groups by proteome analysis of altered gene expression levels.** *Electrophoresis* 1999, **20**:3514-3520.

2. Münchbach M, Dainese P, Staudenmann W, Narberhaus F, James P:
• **Proteome analysis of heat shock expression in *Bradyrhizobium japonicum*.** *Eur J Biochem* 1999, **263**:39-48.
By subtractive two-dimensional gel electrophoresis, the authors identified 19 heat shock proteins in *B. japonicum*. Of the 19 proteins, four were novel. This paper is a good model for studies of a similar nature.

3. Guerreiro N, Djordjevic M, Rolfe B: **Proteome analysis of the model microsymbiont *Sinorhizobium meliloti*: isolation and characterization of novel proteins.** *Electrophoresis* 1999, **20**:818-825.

4. Vasseur C, Labadie J, Hébraud M: **Differential protein expression by *Pseudomonas fragi* submitted to various stresses.** *Electrophoresis* 1999, **20**:2204-2213.

5. VanBogelen R, Schiller E, Thomas J, Neidhardt F: **Diagnosis of**
• **cellular states of microbial organisms using proteomics.** *Electrophoresis* 1999, **20**:2149-2159.
Written in the past year, this a good review to learn about the wide range of potential applications of two-dimensional gel electrophoresis to microbiology.

6. Niimi M, Cannon R, Monk B: ***Candida albicans* pathogenicity: a proteomic perspective.** *Electrophoresis* 1999, **20**:2299-2308.

7. Pitarch A, Pardo M, Jiménez A, Pia J, Gil C, Sánchez M, Nombela C: **Two-dimensional gel electrophoresis as analytical tool for identifying *Candida albicans* immunogenic proteins.** *Electrophoresis* 1999, **20**:1001-1010.

8. Sanchez-Campillo M, Bini L, Comanducci M, Raggiaschi R, Marzocchi B, Pallini V, Ratti G: **Identification of immunoreactive proteins of *Chlamydia trachomatis* by Western blot analysis of a two-dimensional electrophoresis map with patient sera.** *Electrophoresis* 1999, **20**:2269-2279.

9. Cash P, Argo E, Ford L, Lawrie L, McKenzie H: **A proteomic analysis of erythromycin resistance in *Streptococcus pneumoniae*.** *Electrophoresis* 1999, **20**:2259-2268.

10. Jungblut P, Schaible U, Mollenkopf H-J, Zimny-Arndt U, Raupach B,
•• Mattow J, Halada P, Lamer S, Hagens K, Kaufmann S: **Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens.** *Mol Microbiol* 1999, **33**:1103-1117.
This article is an excellent example of comparative proteomics. The authors determined the differences in protein expression between a pathogenic and labile strain of *Mycobacterium*.

11. Lee J, Godon C, Lagniel G, Spector D, Garin J, Labarre J, Toledano M:
•• **Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast.** *J Biol Chem* 1999, **274**:16040-16046.
In this work, Lee *et al.* used 2D gel electrophoresis to compare the protein expression differences between two strains of yeast, each, knocked out, for a different transcription factor. This approach allowed them to distinguish the role each of the transcription factors play in response to redox stresses.

12. Alms G, Sanz P, Carlson M, Haystead T: **Reg1p targets protein**
• **phosphatase 1 to dephosphorylate hexokinase II in *Saccharomyces cerevisiae*: characterizing the effects of a phosphatase subunit on the yeast proteome.** *EMBO J* 1999, **18**:4157-4168.
Using 2D gel electrophoresis, Alms *et al.* deleted the gene for the protein phosphatase binding protein, Reg1, from yeast and analysed the effects of the deletion on the yeast phosphoproteome. This allowed them to determine the physiological substrates of the phosphatase complex.

13. Goode B, Wong J, Butty A-C, Peter M, McCormack A, Yates J, Drubin D, Barnes G: **Coronin promotes the rapid assembly and cross-linking of actin filaments and may link the actin and microtubule cytoskeletons in yeast.** *J Cell Biol* 1999, **144**:83-98.

14. Meeusen S, Tieu Q, Wong E, Weiss E, Schieltz D, Yates J, Nunnari J: **Mgm101p is a novel component of the mitochondrial nucleoid that binds DNA and is required for the repair of oxidatively damaged mitochondrial DNA.** *J Cell Biol* 1999, **145**:291-304.

15. Eberharter A, Sterner D, Schieltz D, Hassan A, Yates J, Berger S, Workman J: **The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1999, **19**:6621-6631.

16. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B:
• **A generic protein purification method for protein complex characterization and proteome exploration.** *Nat Biotechnol* 1999, **17**:1030-1032.
By constructing a removable tag, Rigaut *et al.* have devised a generic method to purify any protein and its associated complex from yeast. After purification, the protein of interest and its associated complex can be analysed by mass spectrometry to determine the identity of the proteins in the complex. The application of this tag to protein complex identification is found in [17–19].

17. Fortes P, Bilbao-Cortés D, Fornerod M, Rigaut R, Raymond W, Séraphin B, Mattaj I: **Luc7p, a novel yeast U1 snRNP protein with a role in 5′ splice site recognition.** *Genes Dev* 1999, **13**:2425-2438.

18. Caspary F, Shevchenko A, Wilm M, Séraphin B: **Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly.** *EMBO J* 1999, **18**:3463-3474.

19. Schmitt C, von Kobbe C, Bachi A, Panté N, Rodrigues J, Boscheron C, Rigaut G, Wilm M, Séraphin B, Carmo-Fonseca M, Izaurralde E: **Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p.** *EMBO J* 1999, **18**:4332-4347.

20. Houry W, Frishman D, Eckerskorn C, Lottspeich F, Hartl R:
•• **Identification of *in vivo* substrates of the chaperonin GroEL.** *Nature* 1999, **402**:147-154.
This is a fascinating application of functional proteomics to determine the *in vivo* substrates of GroEL in *E. coli*. Via large-scale immunoprecipitation, GroEL–substrate complexes were isolated. GroEL–substrate complexes were than resolved by two dimensional gel electrophoresis and 52 protein substrates were identified by mass spectrometry. After analysing the three-dimensional structures of 24 of the identified substrates, Houry *et al.* proposed a common structural motif indicative of GroEL protein substrates.

21. Perrot M, Segliocco F, Mini T, Monribot C, Schneider U, Shevchenko A, Mann M, Jenö P, Boucherie H: **Two-dimensional gel protein database of *Saccharomyces cerevisiae* (update 1999).** *Electrophoresis* 1999, **20**:2280-2298.

22. Mewes H, Frishman D, Gruber C, Geier B, Haase D, Kaps A,
• Lemcke K, Mannhaupt G, Pfeiffer F, Schüller C *et al.*: **MIPS: a database for genomes and protein sequences.** *Nucleic Acids Res* 2000, **28**:37-40.
As more and more knowledge concerning individual proteins is gained, there is a need for databases to categorize proteins based on their function, class and subcellular localization. These types of databases are extremely useful for research in proteomics. This and [23•] are two examples of curated databases available via the Internet that classify yeast proteins into several different categories.

23. Costanzo M, Hogan J, Cusick M, Davis B, Fancher A, Hodges P,
• Kondu P, Lengieza C, Lew-Smith J, Linger C *et al.*: **The Yeast Proteome Database (YPD) and *Caenorhabditis elegans* Proteome Database (WormPD): comprehensive resources for the organization and comparison of model organism protein information.** *Nucleic Acids Res* 2000, **28**:73-76.
This is the second of two references (see [22•]) that describe curated databases that catalogue the current knowledge of yeast proteins. The YPD database is searchable and one can obtain information on both individual proteins and protein classes based on any given search term.

24. Futcher B, Latter G, Monardo P, McLaughlin C, Garrels J: **A sampling**
•• **of the yeast proteome.** *Mol Cell Biol* 1999, **19**:7357-7368.
Previously, it was unknown if mRNA abundance and codon bias information were predictors of protein expression levels. Because of the capabilities of two-dimensional gel electrophoresis, Futcher *et al.* determined, in yeast, that there is a correlation between the codon adaptation index value of a gene and its protein expression level. Future studies in other organisms will determine if this is a general trend or if this is unique to yeast.

25. Sazuka T, Yamaguchi M, Ohara O: **Cyano2Dbase updated: linkage of 234 protein spots to corresponding genes through N-terminal microsequencing.** *Electrophoresis* 1999, **20**:2160-2171.

26. Mollenkopf H, Jungblut P, Raupach B, Mattow J, Lamer S, Zimny-Arndt U, Schaible U, Kaufmann S: **A dynamic two-dimensional polyacrylamide gel electrophoresis database: the mycobacterial proteome via Internet.** *Electrophoresis* 1999, **20**:2172-2180.

27. Cordwell S, Nouwens A, Verrills N, McPherson J, Hains P, Van Dyk D, Walsh B: **The microbial proteome database — an automated laboratory catalogue for monitoring protein expression in bacteria.** *Electrophoresis* 1999, **20**:3580-3588.

28. Fountoulakis M, Takács M-F, Berndt P, Langen H, Takács B: **Enrichment of low abundance proteins of *Escherichia coli* by hydroxyapatite chromatography.** *Electrophoresis* 1999, **20**:2181-2195.

29. Fountoulakis M, Takács M-F, Takács B: **Enrichment of low-copy-number gene products by hydrophobic interaction chromatography.** *J Chromatogr A* 1999, **833**:157-168.

30. Herbert B: **Advances in protein solubilisation for two-dimensional electrophoresis.** *Electrophoresis* 1999, **20**:660-663.

31. Molloy M, Herbert B, Williams K, Gooley A: **Extraction of *Escherichia coli* proteins with organic solvents prior to two-dimensional electrophoresis.** *Electrophoresis* 1999, **20**:701-704.

32. Chevallet M, Santoni V, Poinas A, Rouquié D, Fuchs A, Keiffer S, Rossignol M, Lunardi J, Garin J, Rabilloud T: **New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis.** *Electrophoresis* 1998, **19**:1901-1909.

33. Manabe T: **Capillary electrophoresis of proteins for proteomic studies.** *Electrophoresis* 1999, **20**:3116-3121.

34. Shen Y, Xiang F, Veenstra T, Fung E, Smith R: **High-resolution capillary isoelectric focusing of complex protein mixtures from lysates of microorganisms.** *Anal Chem* 1999, **71**:5348-5353.

35. Jensen P, Paπa-Tolic L, Anderson G, Horner J, Lipton M, Bruce J, Smith R: **Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry.** *Anal Chem* 1999, **71**:2076-2084.

36. Tong W, Link A, Eng J, Yates J: **Identification of proteins in complexes by solid-phase microextraction/multistep elution/capillary electrophoresis/tandem mass spectrometry.** *Anal Chem* 1999, **71**:2270-2278.

37. •• Link A, Eng J, Schieltz D, Carmack E, Mize G, Morris D, Garvik B, Yates J: **Direct analysis of protein complexes using mass spectrometry.** *Nat Biotechnol* 1999, **17**:676-682.
In this work, Link *et al.* developed a novel method to determine the proteins in a large complex – that of the yeast ribosome. This technology bypasses two-dimensional gel electrophoresis by combining high-performance liquid chromatography, capillary electrophoresis and mass spectrometry. The technology may allow for the rapid detection and identification of 2000 proteins from any given sample, a drastic improvement over current two-dimensional gel electrophoresis.

38. • Mann M: **Quantitative proteomics?** *Nat Biotechnol* 1999, **17**:954-955.
This brief paper reviews the methods being developed to determine the relative quantity of proteins in two samples. In particular, Mann provides an analysis of the method developed by Gygi *et al.* [41••].

39. Pása-Tolić L, Jensen P, Anderson G, Lipton M, Peden K, Martinović S, Tolić N, Bruce J, Smith R: **High throughput proteome-wide precision measurements of protein expression using mass spectrometry.** *J Am Chem Soc* 1999, **121**:7949-7950.

40. • Oda Y, Huang K, Cross F, Cowburn D, Chait B: **Accurate quantitation of protein expression and site-specific phosphorylation.** *Proc Natl Acad Sci USA* 1999, **96**:6591-6596.
In this work, Oda *et al.* developed and applied a novel quantitative proteomic method involving metabolic labeling. Two strains of yeast, one knocked out for the *Cln2* gene involved in the cell cycle, were grown under identical conditions except that one strain had ^{15}N as its nitrogen source and the other, ^{14}N. Via mass spectrometry, the resulting difference in mass of identical peptides from each sample allowed the authors to determine the relative abundance of proteins whose expression levels changed by knocking out *Cln2*. Although powerful, this method can only be applied to organisms that can be grown under the defined conditions.

41. •• Gygi S, Rist B, Gerber S, Turecek F, Gelb M, Aebersold R: **Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.** *Nat Biotechnol* 1999, **17**:994-999.
Again using isotopic labeling, Gygi *et al.* developed a general method of quantitative proteomics. After growing yeast on either ethanol or galactose (a carbon source), the cysteine residues of each sample were modified by a 'heavy' or a 'light' reagent. Again, the resulting mass difference of identical peptides from each sample allowed for the relative abundance of proteins in each sample to be determined by mass spectrometry. Because the isotopically labeled reagent is added after cell growth, this method can be applied to any system.

42. • Wilkins M, Gasteiger E, Gooley A, Herbert B, Molly M, Binz P-A, Ou K, Sanchez J-C, Bairoch A, Williams K, Hochstrasser D: **High-throughput mass spectrometric discovery of protein post-translational modifications.** *J Mol Biol* 1999, **289**:645-657.
Few proteomic approaches to post-translational modifications have been carried out. Part of the problem has been the lack of high-throughput software tools to interpret mass spectrometry data. Wilkins *et al.* have designed a software tool – FindMod – that can simultaneously search for 22 different post-translational modifications. Although it has only been applied to individual proteins, FindMod may allow for high-throughput identification of post-translational modifications.

43. Soskic V, Görlach M, Poznanovic S, Boehmer F, Godovac-Zimmermann J: **Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor β receptor.** *Biochemistry* 1999, **38**:1757-1764.

44. Godovac-Zimmermann J, Soskic V, Poznanovic S, Brianza F: **Functional proteomics of signal transduction by membrane receptors.** *Electrophoresis* 1999, **20**:952-961.

45. •• Rout M, Aitchison J, Suprapto A, Hjertaas K, Zhao Y, Chait B: **The yeast nuclear pore complex: composition, architecture, and transport mechansim.** *J Cell Biol* 2000, **148**:635-651.
By combining proteomic methods with molecular biology, and immunolocalization, Rout *et al*. mapped the nuclear pore complex in yeast and proposed a detailed transport mechanism. They first separated proteins from a highly enriched nuclear pore complex fraction by HPLC and SDS–PAGE. Proteins were digested and their identities determined by mass spectrometry. From an initial set of 174 identified proteins, they searched databases to determine the function and localization of as many proteins in this set as possible. Several proteins were obviously contaminants, but they carried out classification assays to determine that about 30 of the remaining proteins were part of the nuclear pore complex. Tagged proteins were localized by immunoelectron microscopy and they generated a stoichometric map of the nuclear pore complex. This is a remarkable study that demonstrates the true power of proteomics when it is combined with other methods available to modern biologists.

46 • Langen H, Takacs B, Evers S, Berndt P, Lahm H, Wipf B, Gray C, Fountoulakis M: **Two-dimensional map of the proteome of *Haemophilus influenzae.*** *Electrophoresis* 2000, **21**:411-429.
In an impressive long-term study, Langen *et al*. detected and identified 502 proteins from the proteome of *Haemophilus influenzae*. They utilized several different types of 2D gels to visualise large amounts of proteins from specific pH regions and proteins from the cell envelope. Furthermore, Langen *et al.* detected and identified many low abundance proteins by enriching fractions via several chromatographic methods. This is one of the most comprehensive 2D gel/proteomics projects to date, and an excellent model for future comprehensive analysis of proteomes.