# Proteomic analysis of phenol degradation in soil yeast *Trichosporon cutaneum*

Jun Man

A Thesis

in

The Department

of

Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
For the degree of Master of Science (Chemistry) at
Concordia University
Montreal, Quebec, Canada

July 2004

*Your file Votre référence*
*ISBN: 0-612-94669-X*
*Our file Notre référence*
*ISBN: 0-612-94669-X*

## ABSTRACT

**Proteomic analysis of phenol degradation in soil yeast *Trichosporon cutaneum***

**Jun Man**

*Trichosporon cutaneum* is a soil yeast that utilizes a variety of aromatic compounds as sole carbon and energy sources. *T. cutaneum* is perhaps the best-characterized aromatic-degrading fungus from which further information about biodegradation can be obtained. In this thesis, the global effects on the *T. cutaneum* proteome associated with growth on phenol have been examined. In order to identify all the enzymes involved in the phenol degradation pathway and to find other proteins associated with the degradation process that are not part of the catabolic pathway, proteomic analysis combine high-resolution two-dimensional gel electrophoresis (2-DE) and mass spectrometry. The optimized sample preparation methods and experimental conditions provide a great advantage in quantitative analysis and potential high-throughout applications. Three mass spectrometry methods, matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry (MALDI-QTOF-MS), nano-electrospray mass spectrometry (nanoESI-MS/MS) and capillary liquid chromatography mass spectrometry (CapLC-MS/MS), have been successfully applied to studying proteins of *T. cutaneum* strain ATCC 58094 differentially expressed in response to growth on phenol. Cross-species identification provides a powerful tool in protein identification with an unknown genome and was applied to the mass spectral data obtained. Except for phenol hydroxylase and *cis, cis* - muconate cyclase, the other identified 24 proteins were not previously known to be associated with phenol degradation in *T. cutaneum*.

## ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Justin Powlowski for giving me the opportunity to work on this exciting field and supervising me throughout this project. Inspired by his faith and passion in science, I learn and grow over the past few years that make my graduate study has truly been a valuable journey in my life. I would also like to thank my committee members, Dr. Paul Joyce and Dr. Ann English, for guiding me through this project and demanding my best work. Special thanks go to Dr. Lena Sahlman for the help in obtaining supplemental data and many suggestions for this thesis. My special thanks also go to Dr. Bernard Gibbs and Dr. David Hyndman for their generous supports and efforts in mass spectrometric work. Without all these contributions, my study would never have reached fruition.

I would also like to thank my colleagues and friends, Liz, Nathalie, Roch, Andrew, Yu, Sheehab, Tara, Wensu and Limei, for much help, friendship and fun times. I especially want to thank Liqun and Yixin for making their place a home for me. Without your love and friendship, I would not have survived and would not be who I am.

As for my family, my deepest gratitude goes to my parents, my sister and brother in law --- my biggest supporters and friends for their constant encouragements and un-conditional love. Thank you for have faith in me and giving me the strength to overcome any obstacle.

# TABLE OF CONTENTS

## List of Figures

## List of Tables

**Abbreviations**

| | |
|---|---|
| 2-DE | two dimensional electrophoresis |
| Cap-LC | capillary liquid chromatography |
| CHAPS | 3-(3-Cholamidopropyl)-dimethylammonio-1-propanesulfonate |
| DEAE | diethylaminoethanol |
| DNA | deoxyribonucleic acid |
| DNase | deoxyribonuclease |
| DTT | dithiothreitol |
| ESI | electrospray ionization |
| IEF | isoelectric focusing |
| IPG | immobilized pH-gradient |
| MALDI | matrix-assisted laser desorption ionization |
| MS/MS | tandem mass spectrometry |
| PDA | diacrylylpiperazine |
| PMF | peptide mass fingerprint |
| PTM | post-translational modification |
| RNA | ribonucleic acid |
| RNase | ribonuclease |
| SDS-PAGE | sodium dodecylsulfate-polyacrylamide gel electrophoresis |
| TCA | tricarboxylic acid |
| TEMED | N, N, N', N'-tetramethylethylenediamine |
| TFA | trifluoroacetic acid |
| TOF | time-of-flight |

**Contributions to the thesis**

All experimental work were performed by Jun Man with the exception of the following. Nano-ESI-MS/MS spectra were obtained and analyzed by Dr. Bernard Gibbs at MDS Pharma. Some 2DE-gels (described on page 60) were prepared and run by Dr. Lena Sahlman at Concordia University.

# Introduction

Degradation of aromatic compounds by microbes is widespread in the biosphere. A large variety of naturally-occurring aromatics can be used as sources of carbon and energy, and many microbial catabolic enzymes also can attack man-made aromatics. The biochemistry of these processes has been studied in the most detail in various bacterial species, and to a lesser extent in yeast and fungi. The focus of this thesis is a proteomics-based approach to study aromatic degradation in the soil yeast, *Trichosporon cutaneum*.

## I. *Trichosporon* and aromatic degradation

### 1. Metabolic capabilities of *Trichosporon* species

The yeast *Trichosporon cutaneum* belongs to the genus *Trichosporon* Behrend, which was described as early as 1890 (Behrend, 1890). Strains have been isolated from a number of sources including soil, industrial waste water, wood pulp, sludge, and clinical specimens. *Trichosporon* species have been shown to use a very large variety of carbon sources (Laaser *et al.*, 1989), including pentoses and hexoses, disaccharides, and polysaccharides (Mörtberg and Neujahr, 1986). Members of this genus also have been shown to degrade and utilize cyclohexanecarboxylic acid, lipids, uric acid, and other purine derivatives (Reiser, 1992). Furthermore, *T. cutaneum* and *T. beigelii* are capable of using various aromatic compounds as sole carbon and energy sources (Fig. 1; Dagley, 1985; reviewed by Neujahr, 1990). A number of different *Trichosporon* strains have been used in diverse biotechnology applications (Table 1). The capabilities of these strains indicate the extraordinary potential of these organisms to efficiently convert various cheap carbon sources into biomass and for biodegradation.

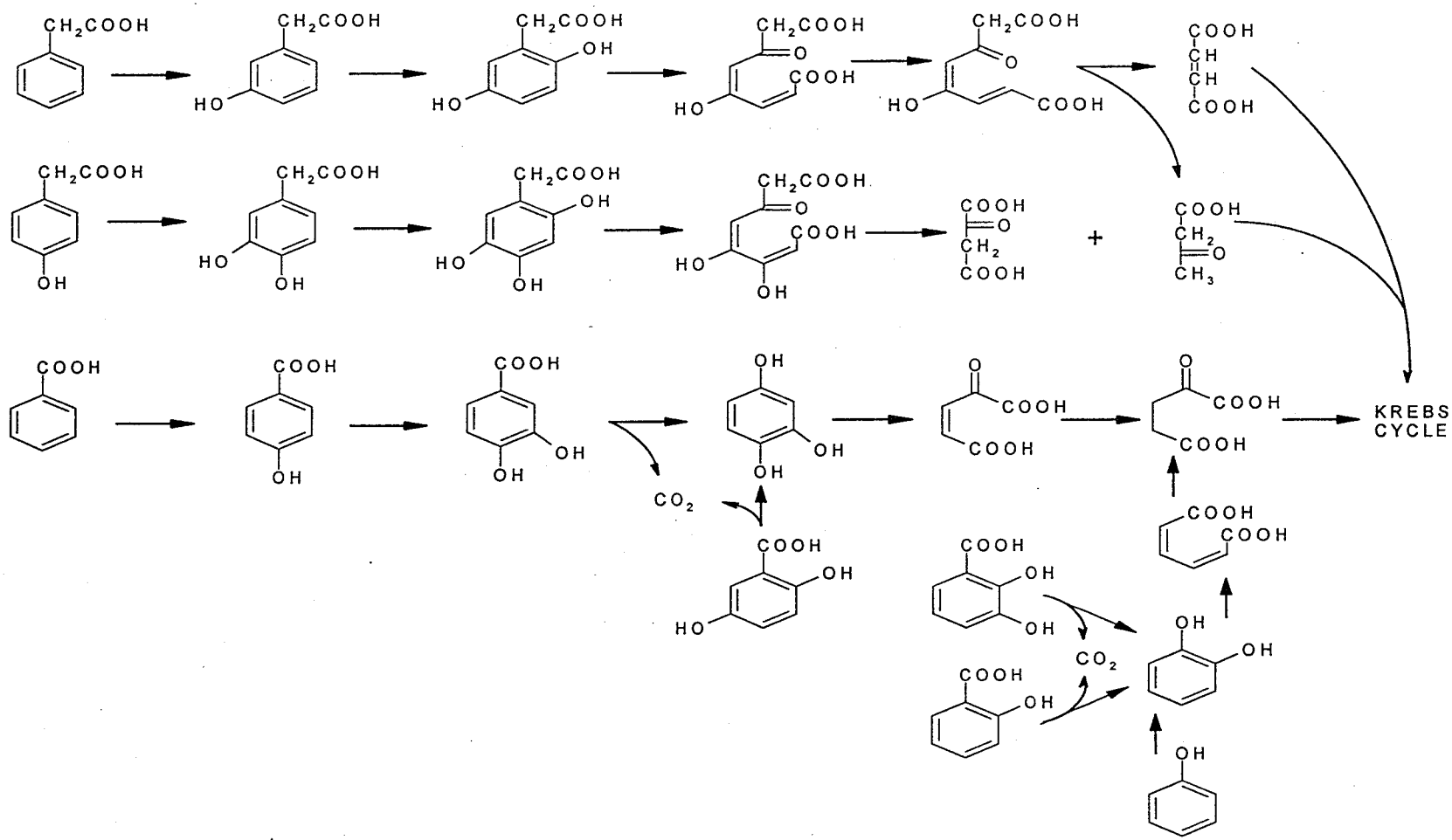


Figure 1. Metabolism of aromatic compounds by *T. cutaneum* (adapted from Reiser *et al.*, 1996). The pathway shown are simplified and do not show all cofactors and substrates required for the individual steps.

**Table 1. Some *Trichosporon* strains and their biotechnological applications**

| Strains | Collections number | Use | Reference |
|---|---|---|---|
| *T. cutaneum* | DSM 70698 | Bioreactor studies<br>Molecular genetics | Käppeli and Fiechter (1982) |
| *T. cutaneum* | ATCC46490 | Phenol degradation<br>Phenol biosensor | Neujahr and Varga (1970) |
| *T. cutaneum* | ATCC58094 | Degradation of aromatic compounds | Sze and Dagley (1984) |
| *T. cutaneum* | ATCC 20509 | Lipid production from whey | West *et al.* (1990) |

Phenol metabolism is one of the best-studied aromatic degradation pathways in *Trichosporon cutaneum* (Fig. 2). The first three enzymes involved, phenol hydroxylase, catechol 1,2-oxygenase and *cis, cis*-muconate cyclase have been isolated from *Trichosporon cutaneum* ATCC 46490. The activities of these enzymes were found to be some 50-400 times higher in phenol-grown than in glucose-grown cells and phenol hydroxylase comprises 2-5% of the total cell protein in fully induced cells (Gaal and Neujahr, 1981; Neujahr and Gaal, 1973). In addition to phenol and resorcinol, growth on catechol, cresols and fluorophenols were found to induce phenol hydroxylase (Gaal and Neujahr, 1981).

Phenol hydroxylase from *Trichospron cutaneum* has been studied extensively. Neujahr and colleagues showed that the enzyme is a flavoprotein (Neujahr and Gaal, 1973; Enroth *et al.*, 1998), and studies by Massey's group demonstrated that, in common with other flavoprotein hydroxylases, enzyme-bound FAD is involved in the oxygen activation and insertion process (Detmer and Massey, 1985; Taylor and Massey, 1990). A full-length phenol hydroxylase cDNA was sequenced and the enzyme was over expressed in an enzymatically active form in *E. coli.*, and purified to homogeneity (Kälin *et al.*, 1992; Waters and Neujahr, 1994). Recombinant phenol hydroxylase was crystallized and a high-resolution crystal structure and a correction of sequencing errors was reported (Enroth *et al.*, 1994; 1998; Enroth., 2003).

In addition to phenol hydroxylase, the two subsequent phenol catabolic pathway enzymes from *Trichosporon cutaneum* also have been studied. Catechol 1,2-oxygenase was

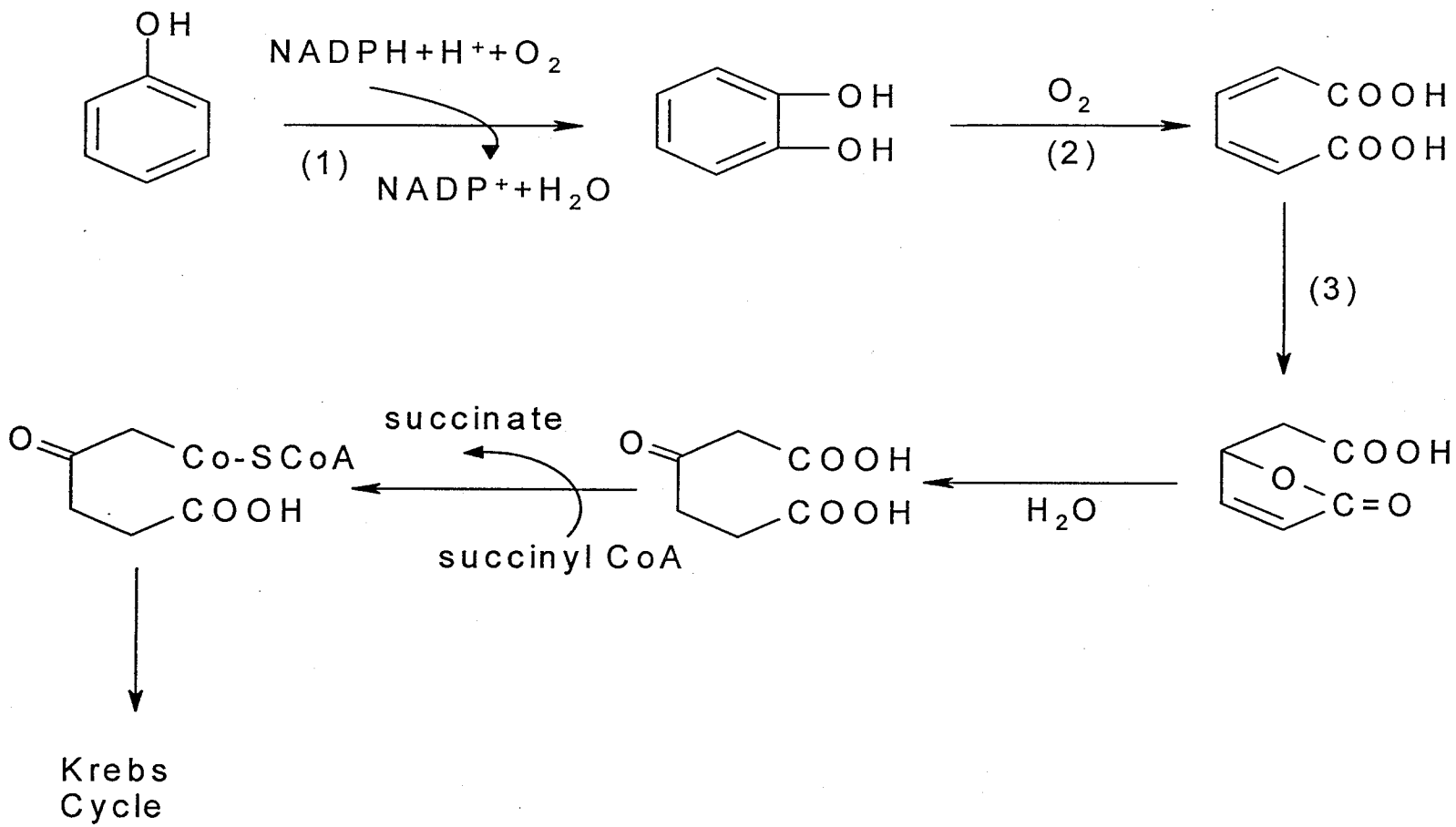


Figure 2. Phenol degradation pathway of *T. cutaneum.* The numbers indicated the sequenced pathway enzymes (1) phenol hydroxylase (2) catechol 1,2-oxygenase (3) *cis, cis*-muconate cyclase.

isolated and purified by Neujahr's group (Varga and Neujahr, 1970; 1972). Another *Trichosporon cutaneum* strain, POB 14, was found to be partially constitutive for catechol 1, 2-oxygenase and preferentially utilized phenol rather than glucose under ordinary growth conditions (Shoda and Udaka, 1980). The *cis, cis*-muconate cyclase from *Trichosporon cutaneum* separated into two bands during isoelectric focusing, designated *cis, cis*-muconate cyclase I (pI 4.58) and II (pI 4.74), respectively (Gaal and Neujahr, 1980). These isozymes differ with respect to the content of free thiol groups: *cis, cis*-muconate cyclase I contains one thiol group that is essential for activity and is stable upon storage, whereas *cis, cis*-muconate cyclase II contains two thiol groups that are readily oxidized during storage with concomitant loss of activity. The absolute stereochemical course of *cis, cis*-muconate lactonizing enzyme from *Trichosporon cutaneum* was determined from $^{1}$H-NMR measurements (Mazur *et al.*, 1994). A cDNA encoding *cis, cis*-muconate lactonizing enzyme was isolated from phenol-induced *T. cutaneum,* and the deduced amino acid sequence exhibited moderate sequence similarity with 3-carboxy-*cis, cis*-muconate lactonizing enzyme from *Neurospora crassa* (Mazur *et al.*, 1994).

A variety of amino and aromatic acids also is degraded by *T. cutaneum* (Anderson and Dagley, 1980; Anderson and Dagley, 1981), as are methyl-substituted phenols (Powlowski and Dagley, 1985). From these studies, and those on phenol degradation, a number of generalizations can be made regarding the strategies used by *T. cutaneum* in comparison with bacterial aromatic degraders. In common with bacteria that degrade aromatic compounds, hydroxylases convert aromatic compounds into di-hydroxylated (catecholic) intermediates, which are substrates for a limited set of ring cleavage

dioxygenases. However, there are some interesting differences between the ring cleavage dioxygenases of *T. cutaneum* and its bacterial competitors. *T. cutaneum* only possesses *ortho*-cleavage enzymes (Neujahr, 1990), which cleave between the two hydroxyl groups of the catecholic substrates, whereas some species of bacteria can use *meta*-cleavage dioxygenases that cleave adjacent to the catechol hydroxyl groups (Fig. 3). *T. cutaneum* also appears to lack a number of common bacterial ring fission dioxygenases, such as protocatechuate dioxygenases (Anderson and Dagley, 1980). Some intermediates in *T. cutaneum* are metabolized differently from those in bacteria due to the pathway enzyme specificities. Thus, for example, *cis, cis*-muconate lactonizing enzymes from bacteria lactonize methyl-*cis, cis*-muconate to a dead-end intermediate, whereas the *T. cutaneum* enzyme lactonizes that to an intermediate that can be further degradaed. This allows *T. cutaneum* to use the *ortho*-cleavage pathway for degradation of methyl-substituted compounds, whereas bacteria are forced to use a completely different pathway involving *meta*-cleavage (Powlowski and Dagley, 1985; Powlowski *et al.*, 1985).

2. Aromatic degradation in other fungi

In addition to *Trichosporon cutaneum,* some other fungi have been studied with respect to the biochemistry of aromatic degrading abilities.

*Aspergilli* are metabolically versatile ascomycetes that are able to rapidly transform a wide spectrum of lignin-related aromatic compounds (Hara *et al.*, 1971). *Aspergillus japonicus* has been described as being a more efficient degrader than other related strains such as *A. terreus*, *A. niger* or *A. flavus*. It is capable of a wide variety of reactions,

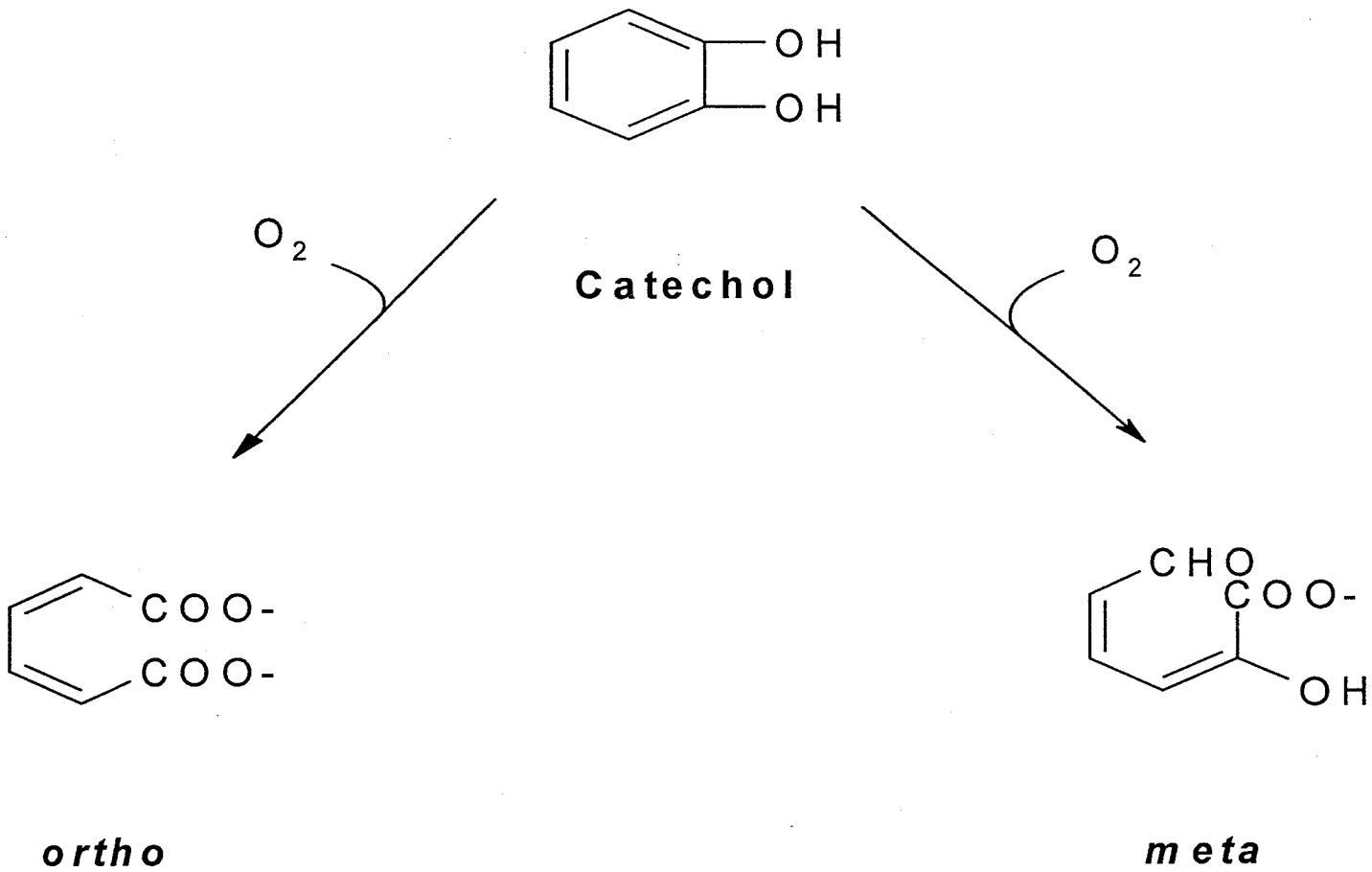


Figure 3. *Ortho* fission and *meta* fission of catechol.

including non-oxidative and oxidative decarboxylations, aromatic alcohol and aldehyde oxidation, aromatic acid reduction, side- chain oxidation, demethylation of aromatic compounds, aromatic ring hydroxylation and aromatic ring cleavage (Milstein *et al.*, 1983). The catabolism of single ring aromatic acids by four species of the genus *Aspergillus*, *A. fumigatus*, *A japonicus*, *A. niger* and *A. terreus*, also has been reported (Milstein *et al.*, 1988). *Aspergillus fumigatus* is capable of growth on phenol as the sole carbon and energy source and its metabolic pathways also have been studied (Jones *et al.*, 1995). The enzymology of phenol degradation in *Aspergillus niger* has been studied in some detail. Thus, β-carboxymuconolactone has been identified as a catabolic intermediate (Thatcher and Cain, 1970). The pathway enzyme, 3-carboxy-*cis*-*cis*-muconate cycloisomerase, has been purified and characterized (Thatcher and Cain, 1974a; 1975), and its subunit structure has been reported (Thatcher and Cain, 1974b). The mechanisms of aromatic decarboxylation and related enzymes also have been described (Kamath *et al.*, 1987, 1989; Santha *et al.*, 1995)).

Another soil yeast, *Candida tropicalis,* that can be induced to metabolize phenol as a sole carbon and energy source has been described (Klein *et al.*, 1979; Neujahr and Gaal, 1973). In addition to phenol, resorcinol, quinol, hydroxyquinol, catechol, and to a lesser extend 4-chlorocatechol, protocatechuate, *p*-cresol, *m*-chlorophenol, and *p*-chlorophenol were found to be metabolized by this yeast (Krug *et al.*, 1985). Some properties of the first two enzymes of the degradation pathway from the strain HP15 have been determined (Krug and Straube, 1986). The kinetics of phenol oxidation by *Candida tropicalis*

indicate that the presence of essential nutrients increases the specific phenol degradation rate and lead to complete phenol oxidation (Paca *et al.*, 2002).

A model filamentous fungus, *Neurospora crassa,* has been reported to degrade some aromatic amino acids via the protocatechuate pathway (Cain *et al.*, 1968; Bilton and Cain, 1968) and the regulation of this pathway has been described (Cain, 1969). Furthermore, there are many fungal species, such as *Exophiala jeanselmei*, *Cryptococcus spp.*, *Rhodotorula spp.*, and *Dipodascus spp.* where growth on aromatics has been shown (Middelhoven, 1993), although details of the biochemistry of aromatic degradation in these species are lacking. Comparatively, the studies of aromatic degradation in *T. cutaneum* are far more extensive than those in other fungi having such degradative capability.

## II. Proteomic analysis and its technologies

A large-scale analysis of proteins will contribute greatly to the better understanding of the biodegradation process in *T. cutaneum*. Proteomics provides a rapid way to achieve the task. A short review of proteomics technologies is presented below.

### 1. The advent of proteomics

Progress in fungal genome sequencing projects has been dramatic and widely publicized (Table 2). The knowledge of full genome sequence information provides a rich opportunity for advancing our understanding of biological processes in these organisms.

**Table 2. Some online sources of fungal genome databases and ongoing projects**

| | |
|---|---|
| *Ashbya* Genome Database | http://agd.unibas.ch/ |
| *Aspergillus nidulans* Database | http://www.broad.mit.edu/annotation/fungi/aspergillus/ |
| *Candida albicans* information | http://alces.med.umn.edu/Candida.html |
| *Candida dubliniensis* Sequencing | http://www.sanger.ac.uk/Projects/C_dubliniensis/ |
| *Cryptococcus neoformnas* Serotype A Database | http://www.broad.mit.edu/annotation/fungi/cryptococcus_neoformans/ |
| *Fusarium graminearum* | http://www.broad.mit.edu/annotation/fungi/fusarium/ |
| *Magnaporthe grisea* Database | http://www.broad.mit.edu/annotation/fungi/magnaporthe/ |
| *Neurospora crassa* Database | http://www.broad.mit.edu/annotation/fungi/neurospora/ |
| *Phanerochaete chrysosporium* | http://genome.jgi-psf.org/whiterot1/whiterot1.home.html |
| *Saccharomyces* Genome Database | http://genome-www.standford.edu/Saccharomyces/ |
| The *S. pombe* Genome Sequencing Project | http://www.sanger.ac.uk/Projects/S_pombe/ |
| *Ustilago maydis* Database | http://www.broad.mit.edu/annotation/fungi/ustilago_maydis/ |
| *Yarrowia lipolytica* | http://cbi.labri.fr/Genolevures/Y_lipolytica.php |

The accumulation of vast amounts of DNA sequence in databases is helpful for predicting the proteins that potentially can be generated, and may be sufficient to ascertain physiological functions. In addition, transcriptional analysis provides important information about expression levels, which may vary under different growth conditions. However, a complicating factor is that after transcription from DNA to RNA, mRNA can be spliced in different ways prior to translation into protein. In addition, proteins synthesized from transcripts may be turned over rapidly. Thus, the expression level of individual mRNAs may not necessarily reflect the level of their protein products *in vivo*. The first multigene comparison plot of mRNA vs. protein abundance for cellular gene products was found to have a correlation coefficient of 0.43 (Anderson and Seilhamer, 1997). In the yeast *Saccharomyces cerevisae*, invariant steady-state levels of certain proteins were observed with respective mRNA transcript levels that varied by as much as 30-fold (Gygi *et al.*, 1999). In addition, levels of active proteins can be altered by post-translational modifications, such as phosphorylation, glycosylation, or protein processing. Therefore, even if all of the genes of a genome have been sequenced and transcription levels have been analyzed, it remains a major task to identify and characterize the proteins in a particular cell type or under a certain condition.

Traditionally, the study of proteins has been very much a one-protein-at-a-time science. Individual proteins are purified to homogeneity, and often are studied in isolation from other cellular proteins. Since individual proteins often need to associate with other proteins in the cell, more global methods of analysis are needed. In the late 1990s, it became possible for protein researchers to purify protein complexes expressed at their

natural level under native conditions from a relatively small number of cells without prior knowledge of the complex composition, activity, or function (Rigaut *et al.*, 1999), and to identify all proteins expressed in an organism (Fountoulakis *et al.*, 1997, 1998a, b; Langen *et al.*, 1997). These advances required improvements in the speed, sensitivity and precision of protein identification and characterization technology.

At the first two-dimensional electrophoresis (2-DE) meeting in Siena, Italy, in 1994, the term *proteome* was first defined by analogy with the term *genome*, and the study of the *proteome* was called *proteomics* (Wilkins *et al.*, 1995). With the rapid developments in genomics and protein chemistry, the term *proteomics* may be defined as "the systematic study of the many and diverse properties of proteins in a parallel manner with the aim of providing detailed descriptions of the structure, function and control of biological systems in health and disease" (Patterson and Aebersold, 2003).

2. Two-dimensional Electrophoresis (2-DE) in proteomics

Major requirements for proteomic analysis are the separation, visualization and analysis of complex mixtures containing as many as several thousand proteins obtained from whole cells, tissues or organisms. Although other technologies exist, two-dimensional electrophoresis (2-DE) remains a powerful technology of choice for separating complex protein mixtures in the majority of proteome projects.

The origin of 2-DE experiments can be traced to 1975, when 2-DE was first introduced (Klose, 1975; O'Farrell, 1975). 2-DE employs isoelectric focusing (IEF) followed by

sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to resolve protein mixtures by charge and by molecular weight. This method can easily resolve 1000, and in some cases as may as 10 000, different proteins at one time (Klose, 1999; Klose and Kobalz, 1995).

Although the basic separation principles of 2-DE have not been changed since 1975, the recent introduction of immobilized pH-gradient (IPG) strips (Bjellqvist *et al.*, 1993) has made IEF more consistent and convenient. IPG strips are available commercially, and the reproducibility is sufficiently good between different laboratories to allow establishment and standardization of two-dimensional gel databases (Blomberg *et al.*, 1995). Furthermore, 2D gel analysis has been improved by the introduction of more sensitive stains for protein detection (Patton, 2000), and software packages to help compare and quantify expression from 2-DE gel images. Two publicly available 2D-PAGE databases have been constructed: SWISS-2DPAGE (Hoogland *et al.*, 2000) and GELBANK (Babnigg and Giometti, 2004). SWISS-2DPAGE contains more than 30 reference maps with more than 1000 proteins identified, mostly for eukaryotes (human, mouse, *Arabidopsis thaliana*, *Dictyostelium discoldeum* and *Saccharomyces cerevisiae*), whereas GELBANK has 81 gel patterns with 233 proteins identified, for mostly prokaryotes (*Methanococcus jannaschii*, *Pyrococcus furiosus*, *Shewanella oneidensis*, *Escherichia coli* and *Deinoccoccus radiodurans*). Thus, 2-DE has maintained its place as the most widely-used method to simultaneously separate and display several thousand proteins from a complex mixture.

Despite its widespread use, there are several limitations inherent in the 2-DE approach. The major problem is that the current dynamic range of 2-DE is not sufficient to reliably profile low-abundance proteins in complex proteomes. In one example, Gygi *et al.* calculated that proteins present at less than 100 copies/cell of a 2-DE gel run with 0.5 mg sample could not be visualized by silver staining (Gygi *et al.*, 2000). Besides that, complete protein solubilization is a well-known potential problem in IEF step, especially for poorly water-soluble proteins, such as membrane proteins and nuclear proteins (Santoni *et al.*, 2000). Finally, due to the inherently complex nature of silver staining procedures, spot intensities may easily vary by as much as 20% between electrophoretic separations (Quadroni and James, 1999). In a recent study, 20% of all spots differed by a factor of 2-5, while an additional 2.3% differed by 5-10 fold and 0.7% by more than one order of magnitude (Voss and Haberl, 2000).

All of these limitations have driven the development of several improvements and alternatives. Strategies for pre-fractionation of samples prior to 2-DE appear to be highly promising to increase the number of proteins that can be visualized in a complex proteome. Two main approaches have been used for enrichment or pre-fractionation of protein mixtures. Fountoulakis and his group have developed extensively the chromatographic approach. This group adopted affinity chromatography on heparin gels as a pre-fractionation step then varied this method by chromatofocusing on Polybuffer Exchanger (Fountoulakis *et al.*, 1997; Fountoulakis and Takacs, 1998a; 1998b). Later, the same group reported another enrichment method involving hydroxyapatite chromatography (Fountoulakis *et al.*, 1999). These different chromatographic steps

allowed the discovery and characterization of several hundred new polypeptide chains, which could not be detected in the unfractionated lysates.

The work approach used by Fountoulakis' group does have some inherent drawbacks. In general, a high concentration of salt (2.5 M NaCl) is needed for complete elution from the columns, especially ion-exchangers. As a consequence, some proteins could be lost during dialysis for salt removal which also will dilute fraction concentrations. Furthermore, the eluted fractions do not represent narrow pI cuts, but generally are constituted by proteins with pIs in the pH 3-10 range.

Two alternative electrophoretic prefractionation methods have also been reported. Hochstrasser's group used a multichamber device developed by Bier (Bier, 1998) to obtain narrow pI cuts (Hochstrasser *et al.*, 1991). However, the pI accuracy of this methodology is quite poor: it ranges from ± 0.65 to ± 1.73 pI units (Wall *et al.*, 2000). Pedersen's group (Pedersen *et al.*, 2003) exploited a multi-compartment electrolyser (Righetti *et al.*, 1989; 1990; 2001), which was able to capture and to detect much more of the "unseen" yeast membrane proteome. It also revealed two isoforms of NADH-cytochrome b5 reductase (pI 8.7 and pI 9.1), which were not detected by chromatographic techniques.

Recent studies have shown that Sypro post-electrophoretic fluorescent stains are more sensitive than silver staining (Lopez *et al.*, 2000). Fluorescent 2-D differential gel electrophoresis (DIGE) has made it possible to detect and quantify differences between

experimental pairs of samples that were resolved on the same 2-D gel, and this was more rapid than using multiple runs for statistical certainty (Tonge *et al.*, 2001). However, the cost of Sypro Ruby from commercial suppliers has been a drawback for extensive application.

3. Mass spectrometry in proteomics

The characterization of 2-DE resolved proteins using mass spectrometry is currently the method of choice in proteomic investigation. Mass spectrometry was introduced into protein analysis more than 30 years ago (Barber *et al.*, 1965; Shemyakin *et al.*, 1966). Compared with the other protein identification methods, such as Edman degradation, the mass spectrometry techniques require less protein sample, less time and are less labour intensive. With the development of new ionization methods in the late 1980s, two "soft" ionization methods, electrospray ionization (ESI; Fenn *et al.*, 1989) and matrix-assisted laser desorption ionization (MALDI) (Karas and Hillenkamp, 1988), mass spectrometry could be used to ionize proteins and peptides at high sensitivity and without excessive fragmentation. The success of these two methods in protein analysis was rewarded by the 2002 Nobel prize in chemistry.

In proteome analysis, mass spectrometry is coupled with 2-DE. Protein spots of interest are excised from gels and digested into peptide fragments using enzymatic (*e.g.* trypsin) or chemical (CNBr) methods (Aebersold *et al.*, 1987). The resulting peptides may be analyzed using mass spectrometry to create a peptide mass fingerprint (PMF) of the protein spot (Henzel *et al.*, 1993; James *et al.*, 1993; Mann *et al.*, 1993; Pappin *et al.*,

1993; Yates *et al.*, 1993). This type of analysis is carried out using MALDI-TOF mass spectrometry where the collected spectra are used to generate a list of peptide masses, that then are compared against the theoretically predicted tryptic peptide masses for each entry in a database, allowing identification of the target protein. The PMF method works well if the analyzed protein and reference database entry have > 80% sequence identity (Willkins and Williams, 1997) or if the full-length sequence of the target protein is in the database (Mann *et al.*, 2001).

However, many genome and protein sequences are still unknown: such is the case for *T. cutaneum,* for which only four sequences are present in Genbank (May 2004). Even for those that are known, modifications such as post-translational modification (PTM) may provide an additional obstacle to identification of the protein. Thus, protein identification by mass spectrometry often requires determination of the protein sequence with minimal assistance from the database. In these cases, tandem mass spectrometry (MS/MS) becomes a very powerful tool to obtain the structural information necessary to identify correctly a protein.

Two MS/MS methods are particularly useful for proteomics studies, where sample quantities are often limited. In nanospray-ESI, unseparated peptide mixtures are sprayed into the mass spectrometer at low flow rates and detected at very high sensitivities, not achieved by ESI-MS (Wilm and Mann, 1996a; Wilm *et al.*, 1996b). Alternatively, an online, one- or multi-dimensional capillary liquid chromatography (Cap-LC) system is used for initial separation of peptide digests followed by real-time sequencing while they

elute into the electrospray ion source (Washburn *et al.*, 2001). This provides a high-throughput and effective way to generate and process the preliminary mass spectrometric data for both known and unknown peptide identifications. While PMF database searching relies heavily on matching masses of intact peptides, data analysis in the tandem mass spectrometry approach usually starts with the comparison of the acquired MS/MS spectrum with theoretical spectra calculated using common peptide fragmentation rules from database entries. The generated amino acid sequences are used to identify the proteins present in the database (Link *et al.*, 1999). If a protein of interest is not present in a database, the peptide sequences can be deduced by using a BLAST- or FASTA-type sequence similarity searches to infer the corresponding proteins (Taylor and Johnson, 2001) or *de novo* sequencing via MS/MS spectral interpretation (Dancik *et al.*, 1999).

With the development of protein separation technologies and advanced mass spectrometry and its instrumentation, proteome studies have been undertaken at a rapid pace. For example, proteome studies of *B. subtilis* identified glycolysis/TCA cycle (house keeping proteins) and proteins synthesized in non-growing cells, allowing a comprehensive analysis of the heat stress stimulon (Hecker, 2003). In a second example, systematic analysis of multiprotein complexes in *Saccharomyces cerevisiae* showed that conservation between eukayotes extends from single proteins to their molecular environment, indicated that the eukaryotic proteome exists as a network of protein complexes, and allowed functions to be proposed for 231 yeast proteins for which no function was known previously (Gavin *et al.*, 2002). A third study, of the proteome of

the malaria parasite, *Plasmodium falciparum,* functionally profiled over 2 400 proteins that were found to agree with the physiology of each stage of the parasite life cycle, and also identified new potential drug and vaccine targets (Florens *et al.*, 2002). Finally, various organellar proteomes have been examined to help shed light on phenomena such as cellular signaling pathways, protein-protein interactions, and responses to external stimuli (Taylor *et al.*, 2003). From examples such as these, and many others, it is clear that proteomic analysis presents unique opportunities to help catalyze biological discovery.

4. Proteomic analysis in fungal biodegradation

As a developing new field, proteomic work on fungi is comparatively limited. Since the Yeast Proteome Database (YPD) was published (Hodges *et al.*, 1999), the proteomic work on *S. cerevisiae* has been ahead of that of all other fungi. The majority of the proteomic studies on other fungi have focused on species, such as *Aspergillus funigatus* (Bruneau *et al.*, 2001) and *Candida albicans* (Pitarch *et al.*, 2003a, 2003b), that are pathogenic to humans. However, no published work has presented to proteomic analysis of biodegradation in fungi. It is, therefore, of interest to investigate the aromatic degradation activities in the model soil yeast *T. cutaneum* by means of proteomic analysis.

5. Aims of this thesis

Since the strain ATCC 58094 has been used for studying aromatic degradation for many years, in this study, phenol induced *T. cutaneum* ATCC 58094 was used to: (i) attempt to

identify all of the enzymes involved in the phenol degradation pathway; (ii) find other proteins associated with the degradation processes that are not part of the catabolic pathway; and (iii) construct a basic proteome for the phenol degradation response in *T. cutaneum* for further comparison with proteins induced for the degradation of related aromatic compounds.

## Materials and Methods

1. Chemicals

Urea, agarose, glycerol, CHAPS, iodoacetamide, acrylamide, ammonium bicarbonate Tris, glycine, and SDS were purchased from ICN (Irvine, CA, USA). DTT was from BioShop. ReadyStrip IPG strips (pH 4-7, pH 3-10, 11cm), mineral oil, 2-D SDS-PAGE standards, ammonium persulfate, TEMED, Bio-Rad Protein Assay Dye Reagent and PDA were purchased from Bio-Rad (Hercules, CA, USA). IPG buffers (4-7), Protein Molecular Weight Standards, HiTrap DEAE-Sepharose columns and silver staining kits were from Amersham Pharmacia Biotech AB (Uppsala, Sweden). Acetic acid, acetone, acetonitrile and methanol were from Fisher (Fair Lawn, New Jersey, USA). TFA was purchased from Sigma (St. Louis, MO, USA). Trypsin (sequencing grade), protease inhibitor cocktail tablets, DNase I, and RNase A were from Roche Applied Science (Indianapolis, IN, USA).

2. Cell strains and media

The strain *T. cutaneum* ATCC 58094 was taken from permanent cultures stored at –80°C streaked on Nutrient Broth-Agar (7.5g agar/500mL Nutrient Broth (Difco)) and incubated for 48 h at 30°C. A single colony then was picked and restreaked on a Nutrient Broth-agar plate and incubated for 24 h at 30°C. One loopful of streaked cultures was used as an inoculum for liquid cultures (Powlowski and Dagley, 1985) that contained (per liter): 1.97 g $Na_2HPO_4$; 4.92 g $KH_2PO_4$; 2.0 g $NH_4Cl$; 0.5 g Casamino Acids; 0.5 g yeast extract; 0.1 g $MgSO_4$; 0.5 mL trace metal solution (0.71 g/L disodium EDTA, 0.23 g/L

$ZnCl_2$, 1.3 g/L $FeSO_4 \cdot 7H_2O$, 0.62 g/L $MnCl_2 \cdot 4H_2O$, 0.10 g/L $CuSO_4 \cdot 5H_2O$, 0.050 g/L $CoNO_3 \cdot 6H_2O$) (Powlowski and Shingler, 1990). The carbon sources were succinic acid (0.35 g/L), or phenol (0.03%) or cresol (0.02%) (Powlowski and Dagley, 1985). The pH of the medium was 6.4. Cultures were grown at 30°C overnight with shaking at 140 rpm/min. Optical density measurements were taken at one-hour intervals. Once the $OD_{600}$ reached approximately 0.6, an additional aliquot of carbon source was added to the culture and cells were harvested when the $OD_{600}$ reached 1.0. The harvested cells were washed in ice-cold 50 mM phosphate buffer (pH 7.0), and pelleted by centrifugation (8000 x g, 20 min). Cell pellets were either used immediately or stored at -80°C until use.

3. Cell lysis

To the cell pellets harvested from 500 mL cell culture was added 1 mL of lysis buffer [1% SDS (w/v), 100 mM Tris-HCl, pH 7.0, 100 µL proteinase inhibitor cocktail, 5 mM DTT] (Harder *et al.*, 1999). The suspended pellets then were divided in 0.5 mL aliquots into individual microfuge tubes. Approximately 100 µL of acid-washed glass beads was added and sonication was performed for 10 x 10 sec (duty cycle 60, level 3, Branson Sonifier 250) with intermediate cooling on ice. The unbroken cells and cell debris were removed by centrifugation at 16 000 x g for 30 min at 4°C. The supernatant was boiled for 5 min (Harder *et al.*, 1999), cooled in an ice-bath then centrifuged for 20 min to remove insoluble material. The clear supernatants were further treated by addition of 0.05 volumes DNase/RNase (1 mg/mL DNase I and 0.25 mg/mL RNase A in 50 mM $MgCl_2$), followed by incubation for 2 h on ice. After 30 min of centrifugation (16 000 x

g, 4°C), supernatants were removed and stored at -80°C prior to electrophoresis. The protein concentrations of the supernatants were determined before loading onto IEF strips.

4. Estimation of protein concentration

All protein concentrations were determined using the BioRad assay kit, based on the Bradford protein assay (Bradford, 1976).

5. DEAE-Sepharose fractionation of crude extracts

The crude protein extract (Section 3) was applied to a HiTrap Fast-Flow DEAE-Sepharose column (Pharmacia, Uppsala, Sweden), which was first washed with 50 mM phosphate buffer (pH 7.4) containing 1 M NaCl, then equilibrated with the same phosphate buffer without salts at a flow rate of 1 mL/min. The sample was applied at the same flow rate, and the column was continuously washed with 10 ml 50 mM phosphate buffer, pH 7.4 containing protease inhibitor cocktail (one tablet/10 mL buffer). The wash step was followed by an elution step with 0.3 M NaCl in the same phosphate buffer. Fractions were collected and tested for phenol hydroxylase activity (Neujahr and Gaal, 1973). Fractions with activity were pooled and dialyzed against 10 mM Tris-HCl buffer (pH 7.4), and the protein concentration was determined.

6. Protein precipitation and solubilization for IEF

Proteins from crude extracts or DEAE fractions were precipitated in 5 volumes of ice-cold acetone and stored at -20°C for at least 2 h (Guy *et al.*, 1994). Precipitated proteins

were collected by centrifugation (16 000 x g, 30 min, 4°C), air dried at room temperature for 1 h, and solubilized in 250 µL rehydration solution [unless stated, otherwise, 9.0 M urea, 100 mM DTT, 4% CHAPS, 0.2% IPG buffer (pH 4-7, pH3-10), and a trace of bromophenol blue] at 30°C for at least one hour. A final 30 min centrifugation (16 000 x g, room temperature) was applied to the rehydrated protein mixture to remove insoluble materials prior to first-dimension IEF.

7. Membrane protein extraction

The following method was adapted from the review of Molloy (Molly, 2000). To the cell pellets after sonication (Section 3) was added 0.5 mL of ice-cold 0.1 M $Na_2CO_3$ with subsequent incubation on ice for one hour. After 30 min of centrifugation (16 000 x g, 4°C), the cell pellets were resuspended in 0.5 mL 20 mM Tris-Cl (pH 7.3) and treated by addition of 0.1 volume DNase/RNase (1 mg/mL DNase I and 0.25 mg/mL RNase A in 50 mM $MgCl_2$), followed by incubation for one hour on ice. The cell pellet was collected as described above and resuspended again in 0.5 mL 20 mM Tris-Cl (pH 7.3). After centrifugation again, the pellet was solubilized in 0.5 mL sample solution [7.0 M urea, 2.0 M thiourea, 100 mM DTT, 4% CHAPS, 1% Trition X-100, 0.2% IPG buffer (pH 4-7), and a trace of bromophenol blue] at 30°C overnight.

8. Two-dimensional SDS-PAGE

8.1 First dimension – Isoelectric focusing on IPG strips

The ReadyStrip™ (Bio-Rad) immobilized pH gradient (IPG) dry strips (pH 4-7, pH3-10, 11 cm) were rehydrated overnight with 185 µL of rehydration solution that contained

about 500 μg solubilized proteins at 50 V using a Bio-Rad PROTEAN IEF Cell. Isoelectric focusing was conducted using the optimized conditions (Table 3.). After IEF,

the IPG strips were equilibrated for 30 min by rocking in an equilibration buffer (50 mM Tris-HCl, pH 8.8, 6 M urea, 30% glycerol, 1% DTT and 2% SDS) at 30°C. The equilibration solution was replaced with fresh equilibration buffer in which DTT was replaced with 2.5% iodoacetamide to alkylate the reduced proteins (Bjellqvist *et al.*, 1993).

8.2 Second dimension – SDS-PAGE

8.2.1 Vertical system SDS-PAGE

The equilibrated IPG strips (Section 8.1) were applied onto 1.5 mm thick 12% SDS polyacrylamide vertical slab gels prepared according to the Laemmli gel method (Laemmli, 1970). A modification was that PDA was used at the gel polymerization step instead of bisacrylamide (Hochstrasser *et al.*, 1988). The Low Molecular Weight protein standards (Amersham Pharmacia Biotech) were applied to a paper IEF sample application piece, then placed on the top of the gel next to one end of the IPG strip. IPG strips and the protein standards piece were then embedded with sealing solution (1% agarose in 25 mM Tris, 192 mM glycine and 0.1% SDS). SDS-PAGE was carried out in a Hoefer™ SE 600 electrophoresis apparatus for one hour at a constant voltage of 15 V/gel, and then switched to 20 V/gel until the dye front reached the bottom of the gel.

**Table 3. First dimension conditions for crude extract and DEAE enrichment.** Samples were loaded onto IPG strips during an initial rehydration step. Subsequently focusing was performed at the voltages and times indicated.

| Step/Voltage | Running Time/Conditions | |
|---|---|---|
| | Crude extract | DEAE enrichment |
| Rehydration | active (50 V) | passive (no voltage) |
| Focussing/50 V | 4 h | 6 h |
| 150 V | 30 min | 2.5 h |
| 300 V | 1 h | 2.5 h |
| 600 V | 30 min | 2.5 h |
| 1200 V | N/A | 1 h |
| 2500 V | N/A | 1 h |
| 5000 V | 2.5 h | 2.5 h |
| 8000 V | 45000 Vh | 45000 Vh |

8.2.2 Horizontal system SDS-PAGE

A precast ExcelGel™ XL SDS 12-14% (180 x 245 x 0.5 mm) and Excel Gel™ SDS buffer strip was placed on the bed of a Multiphor II electrophoresis apparatus (Amersham Pharmacia Biotech). The equilibrated IPG strip was laid gel-side down on the top of the precast SDS gel. The protein standards (Amersham Pharmacia Biotech) were applied as described above for the vertical system. The second dimensional electrophoresis was performed under the following conditions: 1000 V, 40 mA, 40 W for 45 min; remove the IPG strip; 1000 V, 40 mA, 40 W for 5 min; move the cathode buffer strip to cover the place of the IPG strip; 1000 V, 40 mA, 40 W for 3 h.

9. Silver staining

A modified silver staining method was based on a commercial kit (Silver Stain PlusOne; Amersham Pharmacia Biotech) but omitted the use of glutaraldehyde in the sensitization step and formaldehyde in the silver impregnation step (Yan *et al.*, 2000). After staining, all the gels were preserved in 1% acetic acid at 4°C (Mortz *et al.*, 2001).

10. Image acquisition and analysis

The silver stained 2-D gels were scanned (400 dpi resolution) using a UMAX PowerLook III scanner (UMAX) and Photoshop (Version 6.0, Adobe system) software. The gel images were saved as tiff files. Gel images were analyzed, including spot detection and matching, with ImageMaster™ (Version 4.10) 2D Elite software (Amersham Pharmacia Biotech). The size and orientation of each gel image was adjusted with the cropping and rotating tools in the image menu.

11. In-gel digestion

11.1 Destaining

The gels were washed twice for 10 min with distilled water and the gel spots were excised as closely as possible from the silver-stained gels. A same-sized piece of blank gel was excised as a control. The gels were destained further with chemical reducers to remove the silver, as described with the following critical steps (Gharahdaghi *et al.*, 1999). The reactive substances of the chemical reducers are potassium ferricyanide and sodium thiosulfate. These chemical agents were prepared prior to digestion as two stock solutions of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate, both dissolved in water. A working solution was prepared by mixing a 1:1 ratio of the two stock solutions. As this working solution is unstable, it was prepared fresh for each reaction. Working solution (30 to 50 μL) was added to cover the gel pieces and vortexed occasionally until the brownish color disappeared, and then the gel spots were rinsed a few times with water to stop the reaction. Next, 25 mM $NH_4HCO_3$ was added to cover the gel pieces and after 20 min was removed and discarded.

11.2 Excision and washing

The following method was adapted from a previous report (Shevchenko *et al.*, 1996). The destained gel spots were cut into 1 $mm^2$ pieces and transferred into a 1.5 mL microfuge tube. The gel pieces were washed twice for 15 min each with 100 μL distilled $H_2O$ and then dehydrated repeatedly with several changes of 40 μL of 25 mM $NH_4HCO_3$/acetonitrile (1:1) until the gel pieces turned opaque white and sticky. The gel pieces were then dried in a Speed-Vac concentrator for 30 min at 60°C.

11.3 Reduction and alkylation

Prior to enzymatic digestion, reduction and alkylation steps were carried out according to a procedure modified somewhat from a previous report (Shevchenko *et al.*, 1996). The gel particles from the previous step (Section 11.2) were rehydrated in 10 mM DTT in 25 mM $NH_4HCO_3$, and incubated for 45 min at 56°C. After the tubes were cooled to room temperature, the supernatant was replaced by the same volume of fresh 55 mM iodoacetamide in 25 mM $NH_4HCO_3$ at room temperature in the dark for 30 min, and the gel pieces were washed with a 1:1 solution of 25 mM $NH_4HCO_3$/acetonitrile as described above. The gel pieces then were dried in a Speed-Vac Concentrator for 30 min at 60°C.

11.4 In-gel tryptic digestion

Tryptic digestion reactions on excised gel pieces were carried out with 0.1 μg/μL trypsin (modified sequencing grade, Roche) in 25 mM $NH_4HCO_3$ incubated 12 to 16 h at 37°C (Jiménez *et al.*, 1998). The digestion then was stopped by adding 100 μL $H_2O$ and vortexing for 1-2 min. Then 100 μL 2% TFA was added and the peptide solution was removed and transferred to a clean microfuge tube on ice. The gel pieces then were covered with 50 μL 0.1% TFA and placed in cold-water bath with sonication for 30 min. The supernatant was collected and combined with the previous collected peptide solution. An additional 50 μL of 30% acetonitrile/70% 0.1% TFA was added and sonication was repeated as above. This step was repeated with 60% acetonitrile/40% 0.1% TFA and sonication. The supernatant was removed and added to the previous supernatants. The pooled supernatants were concentrated to 10 μL in a Speed-Vac concentrator.

12. Mass spectrometry

12.1 Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry (MALDI-QTOF-MS)

Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry (MALDI-QTOF-MS) was conducted using a hybrid quadrupole time-of-flight mass spectrometer (Q-TOF Ultima™ Global, Micromass, Manchester, UK). The concentrated peptide mixtures (Section 11.4) were prepared according to a thin-layer method (Bacher *et al.*, 2001). Briefly, 1 μL of Alpha-CHC-Matrix™ (Agilent Techonologies, Santa Clara, CA, USA) was mixed with 1 μL of peptide mixtures and spotted onto a MALDI target plate (Micromass, Manchester, UK), then allowed crystallization. Peptide mass fingerprints were generated and were searched against proteins from all fungal species in the Swiss-Prot database by using the software Proteinlynx™ Global Server (Version 2.0.5) (Micromass, Manchester, UK).

12.2 Nano-electrospray tandem mass spectrometry (nanoESI-MS/MS)

Nano-electrospray tandem mass spectrometry (nanoESI-MS/MS) was performed by using an electrospray quadrupole time-of-flight tandem mass spectrometer (MDS Sciex QSTAR; Concord, ON), equipped with a nanospray interface. The concentrated peptide mixtures (Section 11.4) were redissolved in 10 μL of 50% methanol/45% water/5% formic acid and applied to a ZipTip (μC18 column, Millipore, Bredford, MA, USA) for final desalting, purification and concentration. The peptide fragmentation spectra were obtained and were interpreted with the accompanying software (Analyst QS, MDS Sciex QSTAR; Concord, ON). The voltages were adjusted for maximum fragmentation with

the ion spray and orifice potential at 5000 V and 80 V respectively. The collision energy was approximately 30eV. Multiple charged species were selected for collision induced dissociation (CID) and the product ion spectra were submitted to Mascot (Matrix Science, UK) for possible identification. If no positive identities were obtained, the data were manually interpreted with the aid of Analyst QS tools.

12.3 Capillary liquid chromatography mass spectrometry (Cap-LC-MS/MS)

Capillary liquid chromatography mass spectrometry (Cap-LC-MS/MS) was carried out by using a capillary liquid chromatography (CapLC) system (Micromass® CapLC™ pump, Micromass, Manchester, UK) coupled to a hybrid quadrupole time-of-flight mass spectrometer (Q-TOF Ultima™ Global, Micromass, Manchester, UK). The peptide mixtures (Section 11.4) were dissolved in 0.1% formic acid and 5 μL was injected onto a 0.35 x 5 mm $C_{18}$ micro pre-column cartridge (Water Symmetry 300 ™ $C_{18}$, 5 μm OPTI-PAK™ Trap Column, Waters, Millford, MA, USA). The peptide mixtures were retained while the solvent components were washed to waste. The trap then was brought online with a 75 μm x 100 mm $C_{18}$ NanoEase™ column (Waters, Millford, MA, USA). The peptides were sequentially eluted from the CapLC column with a gradient of 10-80% buffer B (acetonitrile/water/formic acid, 95:4.9:0.1, v/v/v) in 35 min. The flow rate was set at 30 μL/min during the loading and washing stages and was reduced to 4 μL/min before peptide elution from the column which due to a flow split, resulted in a spray flow rate of 250 nL/min. The eluted peptides were sprayed directly from the tip of the capillary column to the mass spectrometer for MS/MS analysis. The mass spectrometer was set to operate in automatic MS/MS switching mode. Only doubly and triply

protonated ions in the mass-to-charge (m/z) range from 400 to 1600 were selected for collision-induced dissociation (CID) using collision energies at 22-55 V. Proteins were identified using the in-house Mascot search engine against MSDB database (Mascot, Matrix Science, UK).

## Results and Discussion

I. Sample preparation

A highly effective method of sample preparation is required to obtain good 2-DE results. In general, the aims of sample preparation for 2-DE are to: (i) obtain and solubilize as many proteins as possible; (ii) minimize the presence of interfering contaminants; (iii) keep the sample preparation strategy as simple as possible to avoid protein loss. Typically, sample preparation includes the following steps: (i) lysis (breakage) of the cell wall, (ii) inactivation or removal of interfering substances (*e.g.*, proteases, nucleic acids), and (iii) maximization of protein solubilization (*i.e.*, disruption of protein aggregates or complexes into soluble individual polypeptides) (Dunn and Corbett, 1996). Many sample preparation methods have been described in the literature, and often the optimal protocol depends on the nature of the sample, *i.e.* its source. However, there is no published precedent for 2-DE gel sample preparation from *T. cutaneum*. Since proteins were very poorly resolved on initial 2D gels (Fig. 4A) of extracts from this organism, much effort was devoted to devising an optimal sample preparation method.

1. Cell lysis

To fully analyze all cytoplasmic proteins, cells must be effectively disrupted. The choice of lysis method depends on the nature of the sample. Baker's yeast (*S. cerevisiae*) is effectively disrupted by using glass bead homogenization (Jazwinski, 1990) or sonication (Harder *et al*, 1999) to break the cell wall. When either lysis method was applied to *T. cutaneum*, around 2 mg of crude protein was liberated per gram (wet weight) of cell pellets. It is possible that *T. cutaneum* might have a tougher cell wall than *S. cerevisiae*,

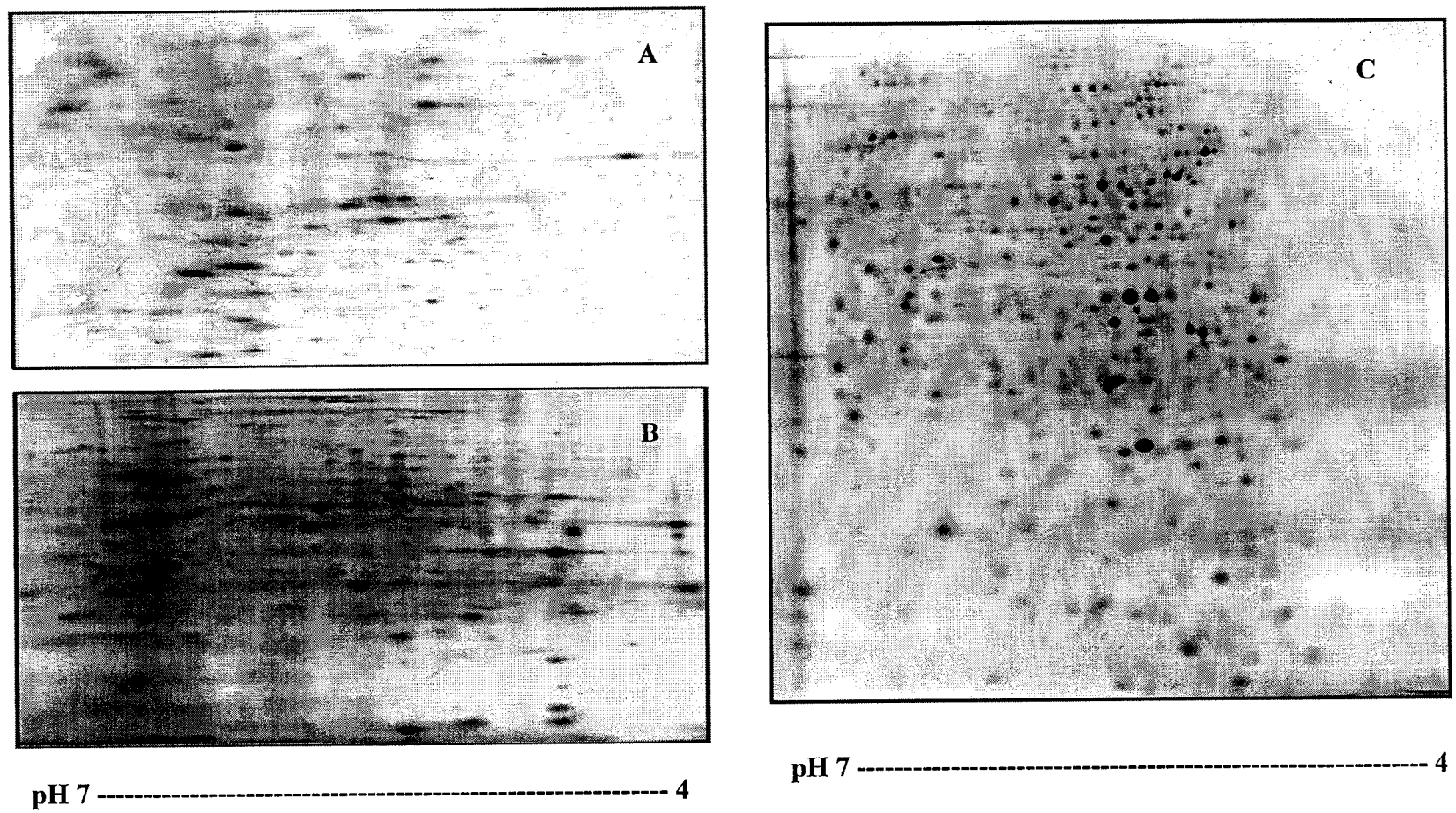


**Figure 4.** Comparison of sample preparation effects in 2D-gels of crude extract of *T. cutaneum*. (A) cells were sonicated then proteins were solubilized in the rehydration buffer containing 7M urea and 2M thiourea; (B) cells were sonicated with glass beads then proteins were solubilized in rehydration buffer containing 9M urea; (C) method (B) followed by DNase/RNase treatment and acetone precipitation.

necessitating more vigorous lysis methods. By combining the glass bead homogenization and sonication, the soluble protein extracted was 10-15 times higher than that obtained using either single lysis method alone. Since proteases may be activated during the vigorous cell lysis, a commercial protease inhibitor cocktail (Roche) was added directly to the lysis buffer to protect against proteolysis. A 2-DE gel of sample preparation using the modified lysis method is shown in Fig. 4B.

2. Removal of contaminants that affect the 2-DE results

Salts and nucleic acids (DNA, RNA) are the main non-protein impurities which can cause insufficient focusing in the first-dimension or result in background smears after silver staining (*e.g.* Fig. 4B). Therefore, two steps were added to the sample preparation procedure: (i) treating samples with a protease-free DNase/RNase mixture to reduce the nucleic acid contamination and (ii) precipitation of proteins by addition of ice-cold acetone to remove interfering compounds, such as salts, pigments and polyphenolics (Matsui *et al.,* 1997). To minimize the appearance of DNase and RNase on the 2-DE gel, the added amount of these enzymes was kept low (0.05 x volume of sample solution) and the incubation time was prolonged (2 hours). Finally, insoluble material was removed by centrifugation prior to first-dimension IEF. A 2-DE gel of a sample prepared using these modifications is shown in Fig. 4C.

3. Solubilization medium

Pretreatment of samples for IEF requires solubilization, denaturation, reduction, and alkylation to break the interactions between proteins. Commonly-used cocktails for

protein solubilization contain 2 M thiourea and/or 7 M urea, plus a high concentration of an efficient detergent (*e.g.*, 4% CHAPS), 1% DTT, and 2% carrier ampholytes (Rabilloud, *et al.*, 1997). Urea has been used for decades in isoelectric focusing to solubilize and unfold most proteins to their random conformation with all ionizable groups exposed to solution. Thiourea recently has been found to further improve solubilization, especially of membrane proteins (Musante *et al.*, 1998). Complete reductive cleavage of inter- and intra- chain disulfide bonds is commonly achieved with dithiothreitol (DTT), which is usually added to the solubilizing cocktails. CHAPS, a zwitterionic detergent helps prevent aggregation via hydrophobic interactions, and carrier ampholytes enhance protein solubility by minimizing protein aggregation due to charge-charge interactions (O'Farrell, 1975). A 2-DE gel of sample preparation using above modifications (DNase/RNase treatment and solubilization medium) is shown in Fig. 4C.

Initially, a method for sample preparation of soluble proteins from *S. cerevisiae* was tested on *T. cutaneum* extracts. A comparison of three different solubilization methods for *S. cerevisiae*, a yeast on which many proteomic studies have been done, yielded a preferred method of boiling the sample with SDS (1% w/v), followed by dilution with urea-thiourea buffer (7 M urea/2 M thiourea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v carrier ampholytes) (Harder *et al.*, 1999). Specific improvements over other methods tested were that protein degradation was minimized by boiling the sample in SDS immediately after sonication, so that protein disaggregation and solubilization of high molecular weight proteins were maximized. However, relatively few spots were observed on 2-DE gels of *T. cutaneum* extracts prepared in this way (compare Fig. 4A to

4B). The reason might be that the thiourea strongly quenches the alkylation process (as typically carried out with iodoacetamide, IAA) during equilibration after the first dimension (Galvani *et al.*, 2001). Minimizing the thiourea concentration in the urea-thiourea mixture (0.5 M thiourea, 8 M urea) previously was shown to give good resolution in 2-DE gels of soluble proteins (Musante *et al.*, 1998). Thiourea was eventually omitted from the sample buffer for 2-DE of *T. cutaneum* extracts, and the urea concentration was increased to 9 M. The modified sample buffer included 9.0 M urea, plus 100 mM DTT, 4% CHAPS, 0.2% IPG buffer (pH 4-7) for 2-DE of *T. cutaneum* proteins, and better resolution was obtained (Fig. 4C).

II. Experimental conditions for 2-D electrophoresis

1. Optimal pH range for immobilized pH gradient (IPG) gel

Immobilized pH gradient (IPG) IEF is the current method of choice for the first dimension of 2-DE for most proteomics applications (Görg *et al.*, 2000). A wide-range, linear pH 3-10 gradient is often useful for the initial analysis of a new type of sample. However, for many samples this may result in loss of resolution in the region pH 4-7, in which the isoelectric point (pI) values of many proteins occur (Görg *et al.* 1988). This problem can be overcome to some extent with the use of a pH 4-7 IPG IEF gel (Görg *et al.* 2000) or of a nonlinear pH 3-10 IPG IEF gel (Bjellqrist *et al.*, 1993), in which the pH 4-7 region has a much flatter gradient than in the pH 7-10 region.

For 2-DE of *T. cutaneum* extracts, a nonlinear pH 3-10 IPG gel strip (Pharmacia Biotech) was used first for an overview of total protein complex distribution (Fig. 5). Although

the resulting 2-DE separation was not satisfactory because of insufficient spaces and difficulty in visualizing low copy number proteins, it revealed that most of the proteins were distributed at the pH 4-8 range (Fig. 5). Therefore, pH 4-7 IPG gel strips were selected for further analysis of the *T. cutaneum* proteome.

2. Focusing time in first-dimension IEF

IEF is an electrophoretic method that separates proteins according to pI. In theory, the optimum focusing time required for the best quality and reproducibility is the time needed for the IEF separation pattern to achieve the steady state (Görg *et al.*, 1988; 1997). If the focusing time is too short, this will result in horizontal and vertical streaking, but over-focusing may produce horizontal smearing because of the electroendosmotic water and protein movement. After trial and error, the following optimum focusing procedure was established: (i) a reswelling step at 50 V to get rid of salts and buffer ions in the sample, (ii) gradual increase of the voltage (from 50 V to 8000 V over 10 h) in the focusing steps to optimize sample entry and separation, (iii) maximization of the final focusing time to give the best steady-state separation of proteins at their pI. The optimized focusing times used for protein samples from *T. cutaneum* are summarized in Table 3.

3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) in the second-dimension

Two formats of SDS-PAGE gels were used for the second-dimension of 2-DE: (i) a single percentage gel containing 12.5% total acrylamide prepared for a vertical electrophoresis system and (ii) a ready-to-use gradient gel (12-14%) from Pharmacia

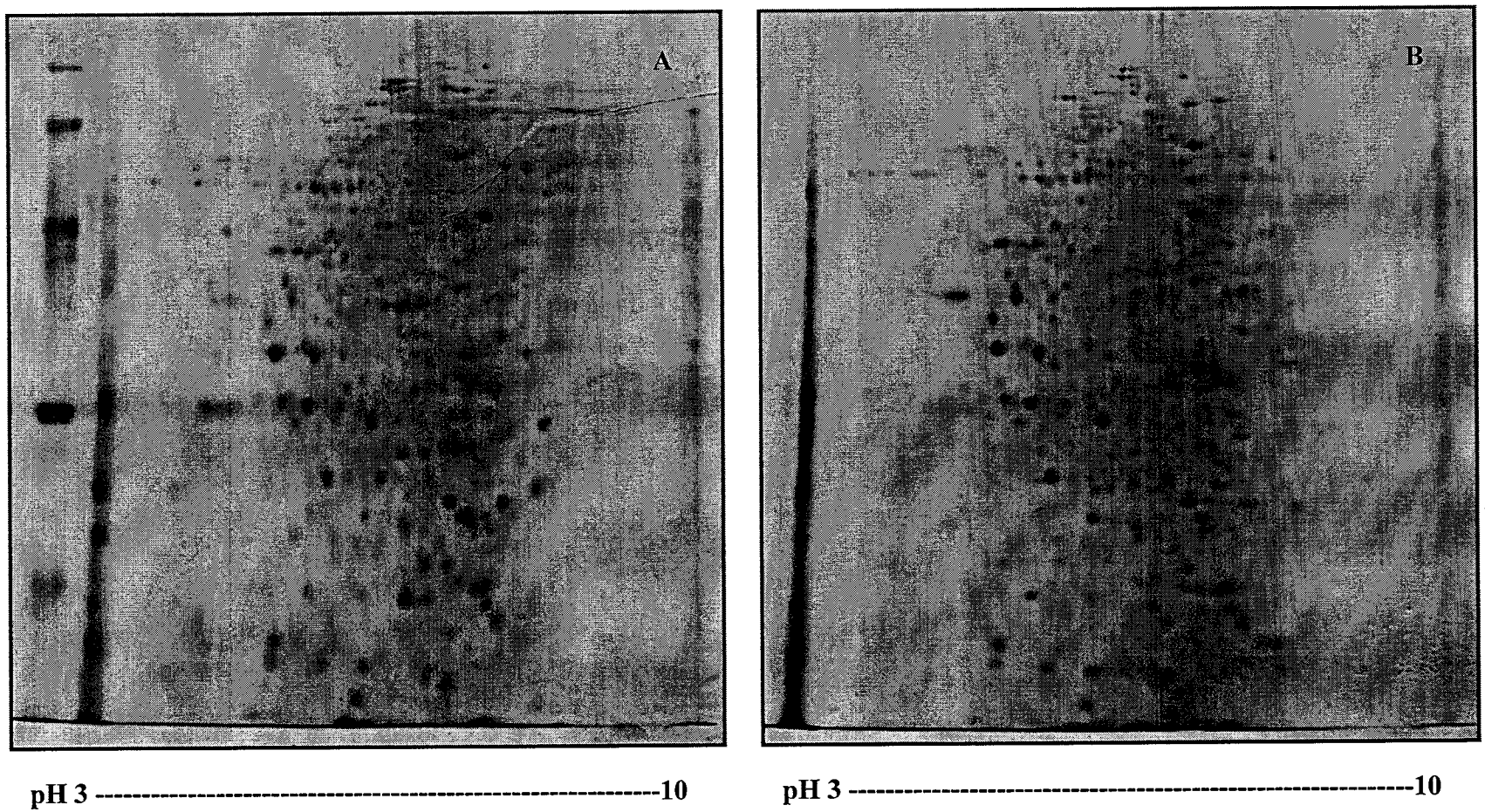


**Figure 5.** 2-DE gel maps of crude protein extracts from *T. cutaneum* grown with (A) phenol and (B) succinate, separated at pH 3-10, 11 cm IPG strips.

Biotech (ExcelGel™) used for a horizontal Multiphor electrophoresis system. The major advantage of the horizontal gel is that the gels are attached to a plastic support, thereby preventing breakage of the gel during the silver staining procedure. In addition, protein spot sharpness can be superior to that obtained in the most commonly used vertical system due to the decreased gel thickness (typically 0.5 mm in horizontal gel vs. 1.5 mm in vertical gel). During the course of this thesis work, about 12 ExcelGel™ gels were run on extracts from cells prepared from both growth conditions. At the same crude protein loading amount (500 μg), more spots were detected from the horizontal gel than from the vertical gel (about 800 spots vs. about 400 spots), indicating a better separation in the gradient resolving system than that in single-concentration resolving system (vertical gel) (Fig. 6). However, since the resolving-gel distance in the horizontal gel system is less than that in the vertical gel system (140 mm vs. 160 mm), good resolution was hard to achieve. This made the comparison of between-gel reproducibility fairly hard either in spot position or in the changes of spot intensities, especially in the high molecular weight region. In addition, ten spots with high intensity from one ExcelGel™ gel were excised for mass spectrometry analysis, but none of them yielded any mass spectral results. This indicated that the protein amount from a single spot separated in the thin format (0.5 mm thickness) gel was not sufficient to identify protein by mass spectrometry. Therefore, the vertical second-dimension SDS-PAGE system was used preferentially for 2-DE in this study in order to maximize the reproducibility of the gel pattern and the compatibility with mass spectrometry analysis. Application of these sample preparation and analysis procedures allowed the reproducible collection of high quality data such as those shown in Fig. 7.

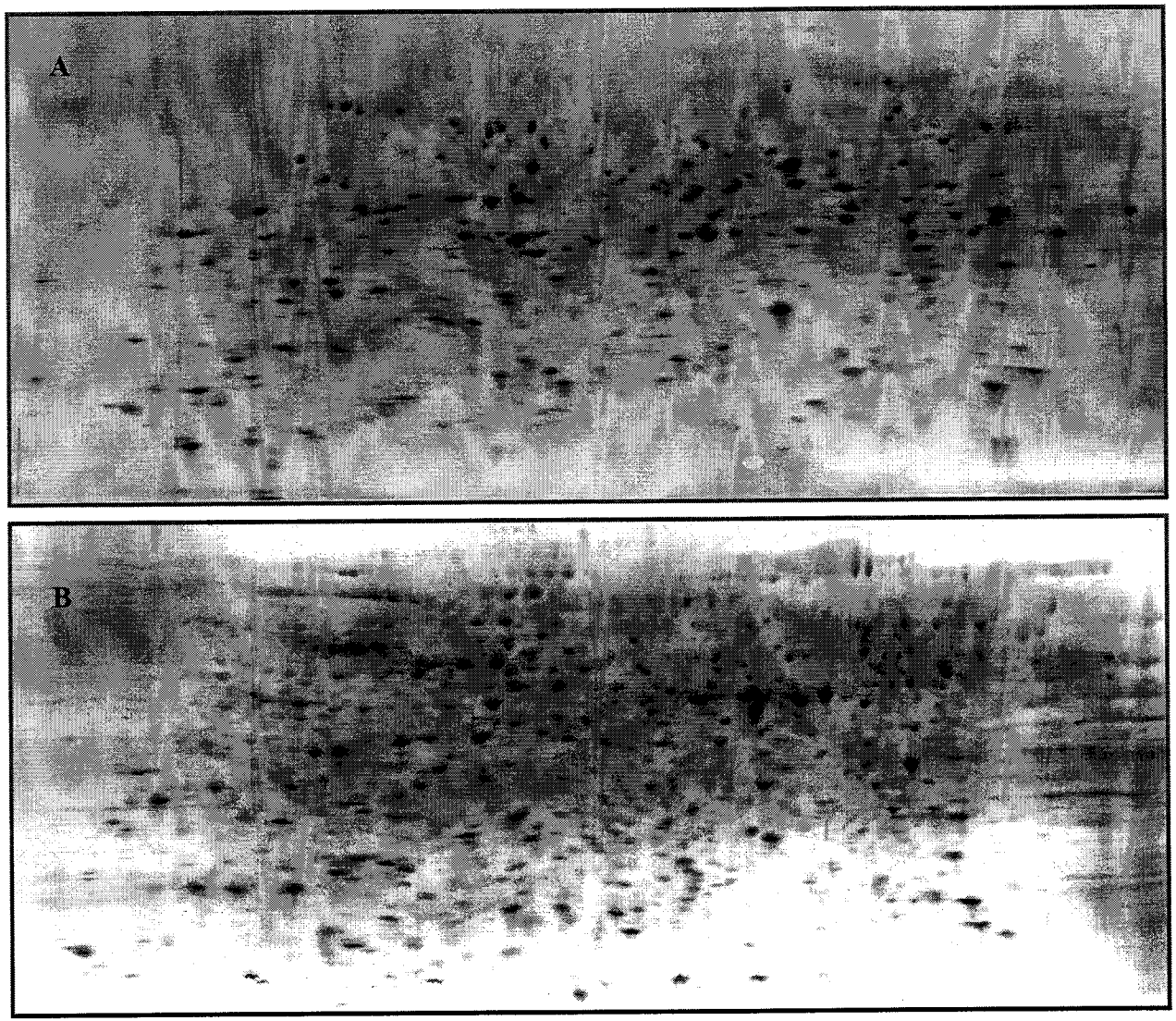


**Figure 6.** 2-DE gel maps of crude protein extracts from *T. cutaneum* grown with (A) phenol and (B) succinate. Samples were separated at pH 4-7, 17 cm IPG strip and horizontal gradient gel.

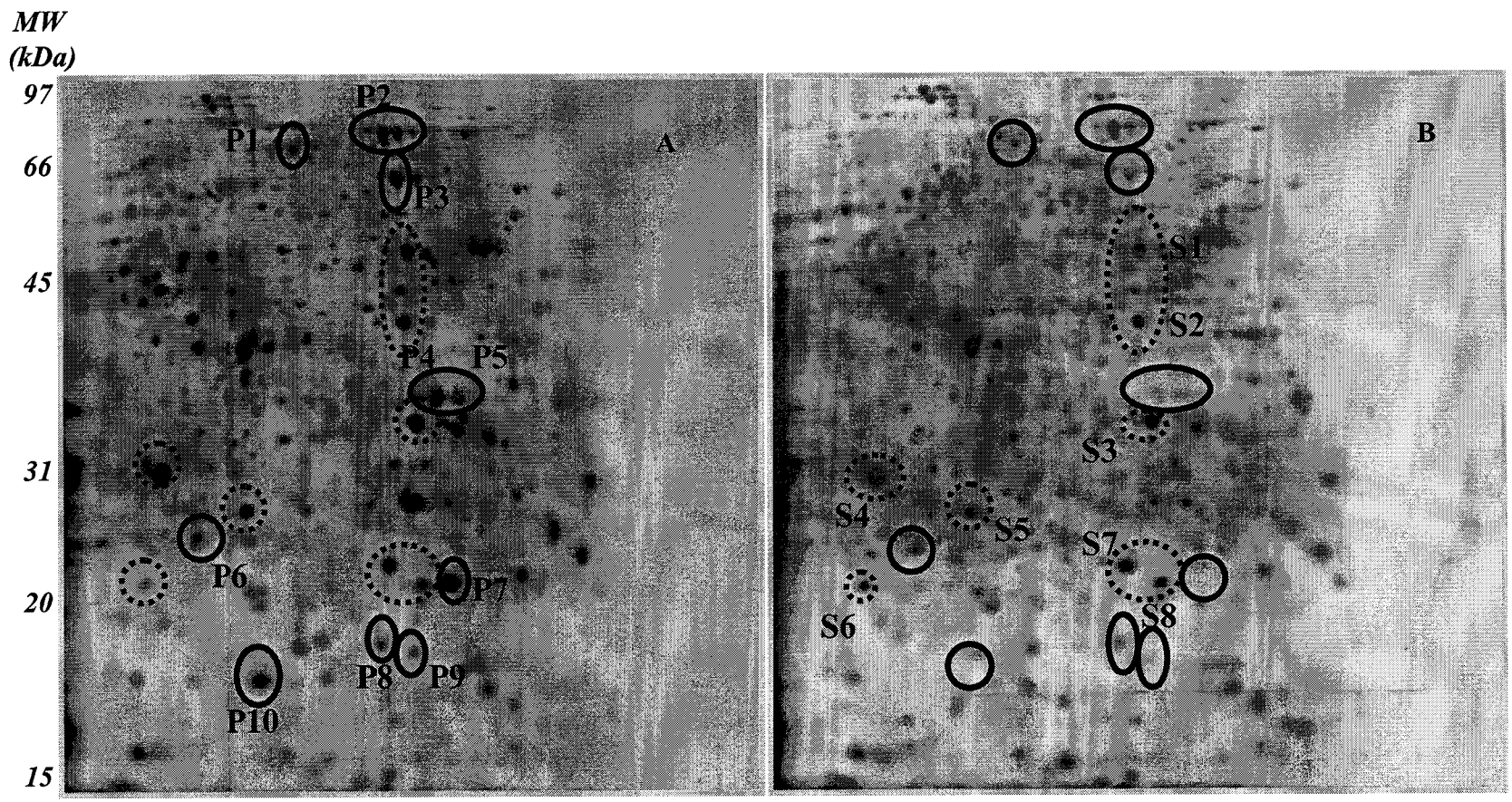


Figure 7. 2-DE gel maps of crude protein extracts from T. cutaneum rwon with (A) phenol and (B) succinate. Solid line Circles indicate over-expressed proteins in phenol-grown cells and dot line circles show proteins expressed equally under Both growth conditions.

III. Reproducibility of 2-DE gel patterns for phenol- and succinate-grown cells

A typical 2-DE gel can contain hundreds to thousands of protein spots. However, slight technical variability from gel to gel or batch to batch often yields variations in the number of spots. Some spots may appear in one gel but are not observed in a parallel gel of the same sample. Therefore, reproducibility analysis becomes necessary to ensure differentially expressed proteins between phenol-cultured and succinate-cultured conditions are real and not artifacts of experimental variables.

Four 2-DE gels were performed with the crude protein extracts from every batch of cells grown under both conditions. Two of the four gels were randomly selected for image analysis with ImageMaster™ 2D Elite software (Version 4.10). A unique function of this software is that it allows the user to create "averaged gels" comprising a representative set of spots generated automatically from a group of matched gels. In practice, two of the four gels described above were used to create an average gel, which represented the spots appeared on both gels and their averaged spot intensities. Another parallel averaged gel was created from a different batch of cells using the same method described above, to assess qualitative and quantitative reproducibility of the 2-DE gel pattern observed for both growth conditions.

"Qualitative variations" refer to the appearance or the disappearance of individual spots, so it is important to know whether or not spots shift position from gel to gel. In order to compare the exact reproducibility from batch to batch, the molecular weight and pI of each protein were assigned as parameters for comparing the positional reproducibility of

each spot in every batch. The positional variation of protein spots from phenol-grown cells and succinate-grown cells are displayed in Fig. 8 and Fig. 9, respectively. The high correlation coefficients obtained were no less than 0.9986, demonstrating an excellent reproducibility of positional pattern.

"Quantitative variations" refer to the optical density of individual spots on 2-DE gels. The intensity was calculated as spot volume or normalized volume. Since the intensity of each protein spot is related to the total intensity of all the proteins detected on the gel, the normalized volume is an effective means to normalize against the total protein volume. Thus the normalized volume was used to compare the batch-to-batch quantitative variation and the data are displayed in Fig. 10. The correlation factors for both growth conditions were close to the optimal diagonal line, which shows that a high level of batch- to-batch reproducibility was obtained. Thus, the changes in intensities of detected spots on the gels analyzed are meaningful.

In order to identify the proteins of *T. cutaneum* induced under phenol-cultured conditions, 2-DE and comparative analyses were performed using phenol- and succinate-grown cells, respectively. The crude extract proteins were prepared from cells grown on phenol and succinate in media according to the methods described in *Materials and Methods*. More than 400 protein spots were revealed on silver-stained 2-DE gels and the protein maps. Following reproducibility analysis, 466 protein spots were detected in phenol-induced cells and 450 protein spots were present in succinate-induced cells (Fig. 7). It thus appears that phenol induced the production of more proteins than succinate did. Image

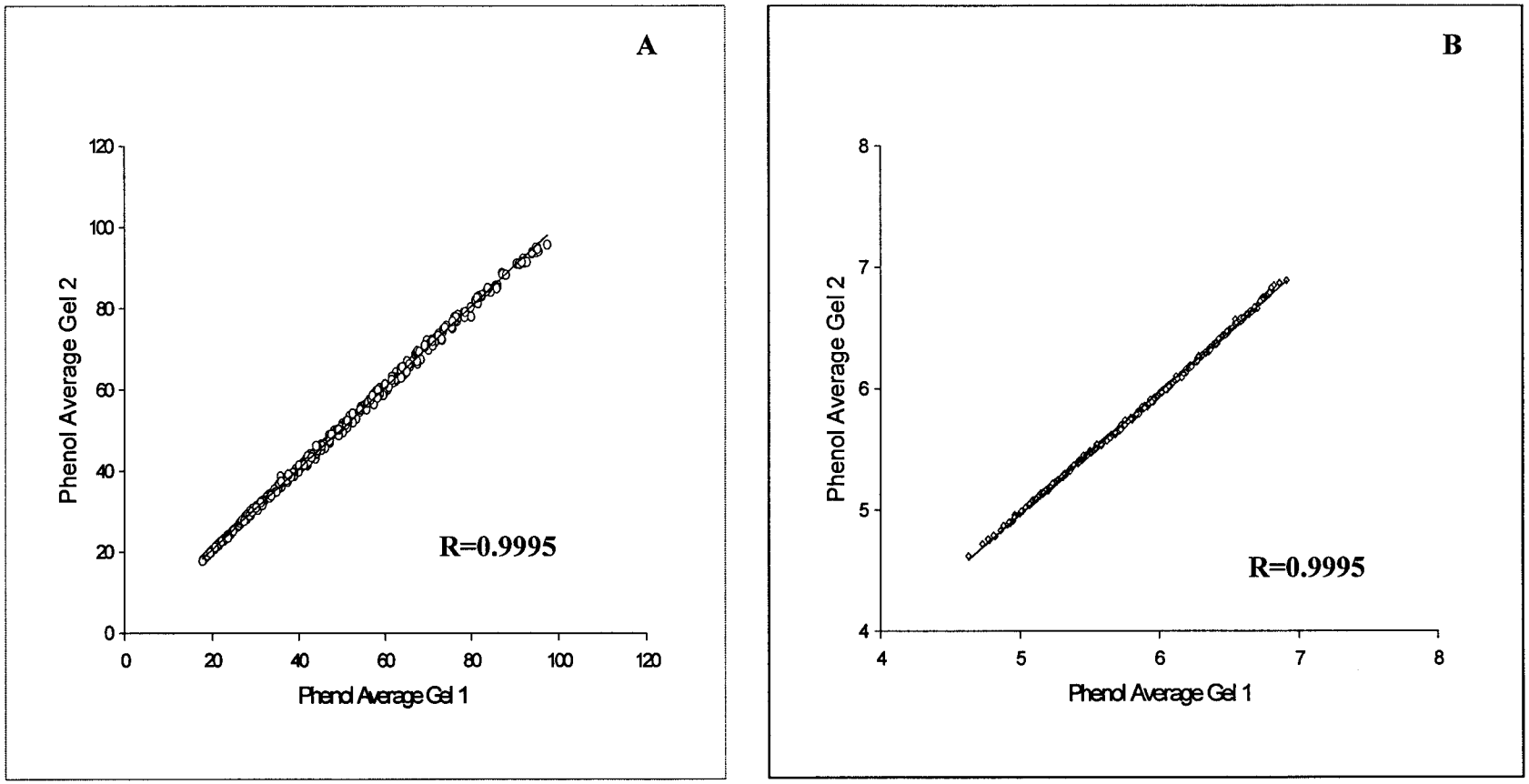


**Figure 8.** Positional reproducibility of 2-DE gels of phenol-grown cells based on (A) molecular weight (kDa) and (B) pI.

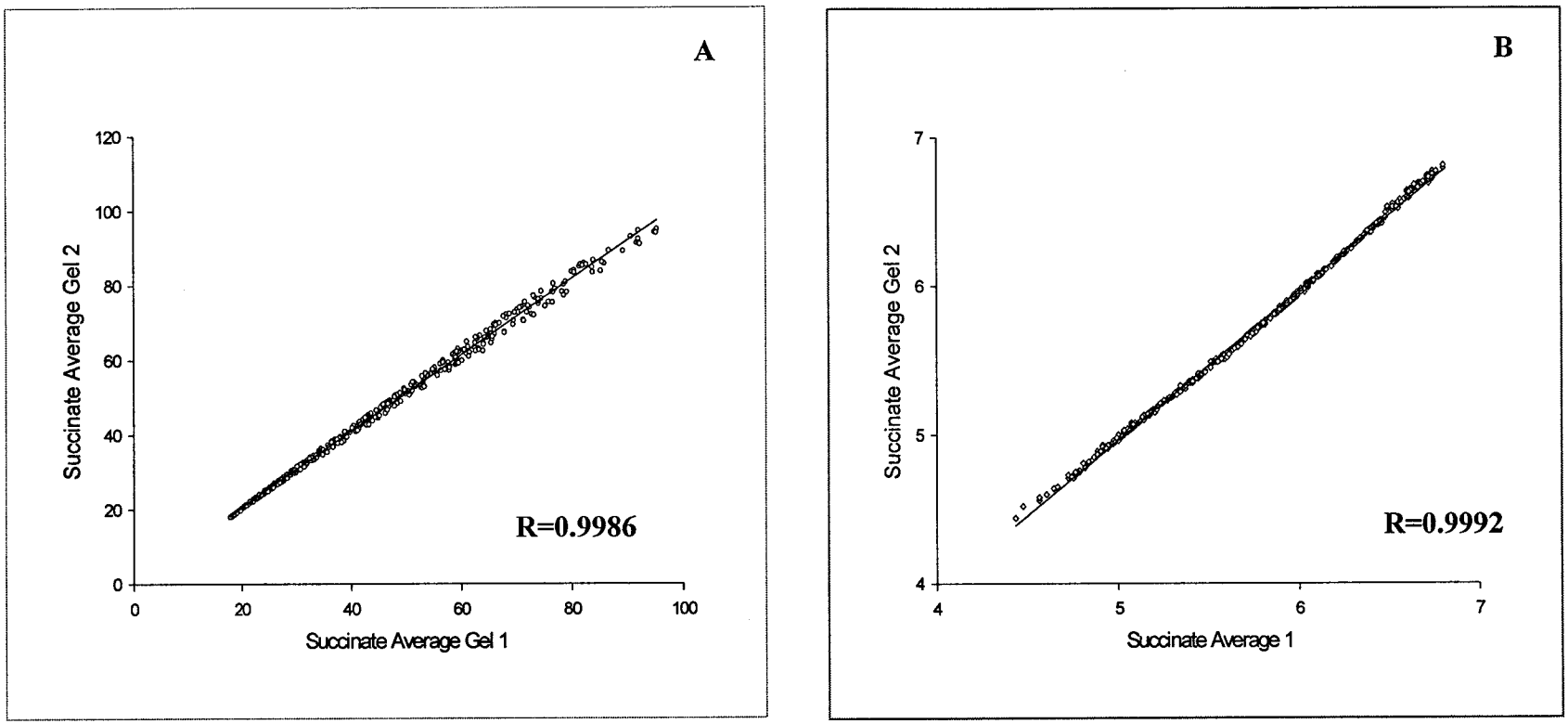


**Figure 9.** Positional reproducibility of 2-DE gels of succinate-grown cells based on (A) molecular weight (kDa) and (B) pI

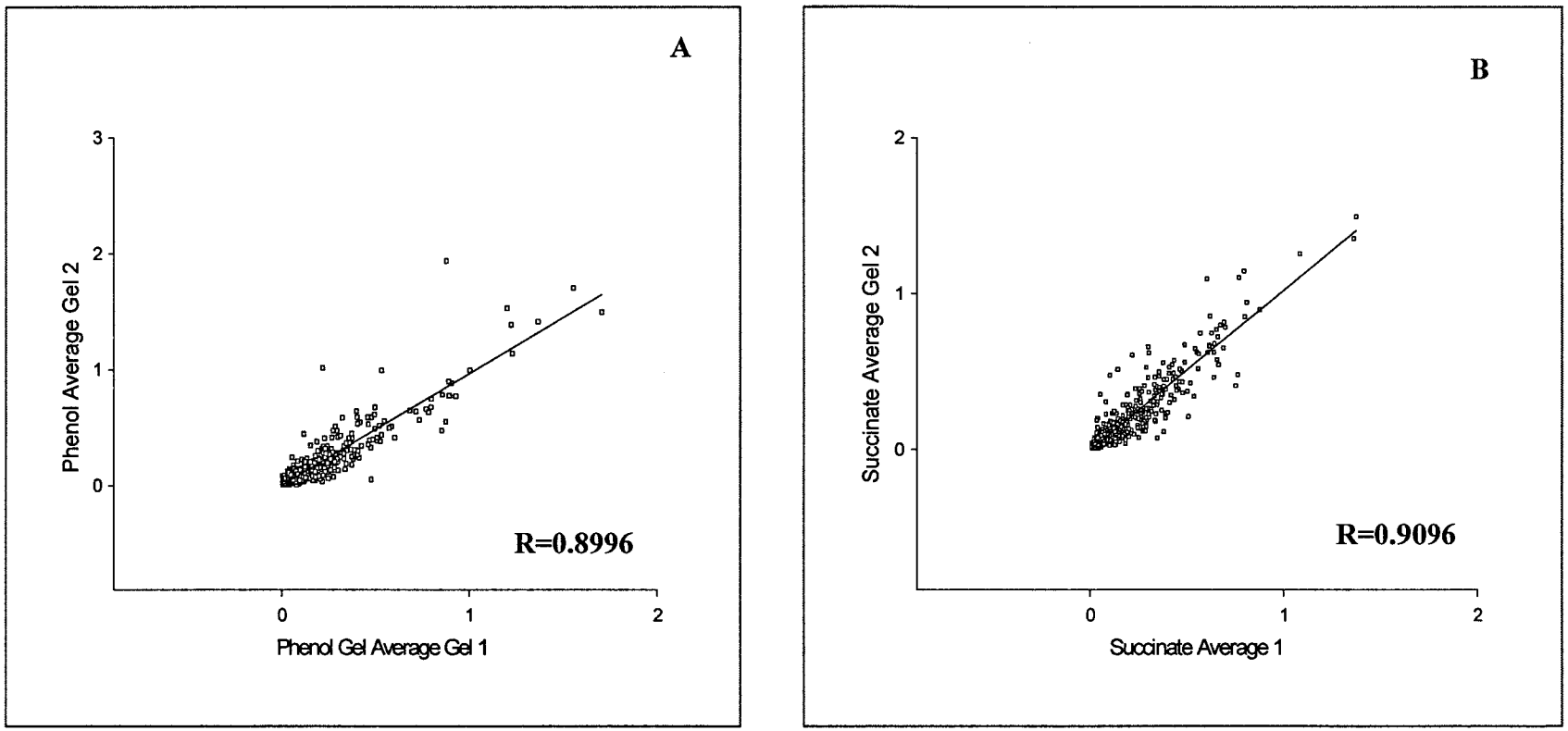


**Figure 10.** Normalized spot volume reproducibility in (A) phenol-grown cells and (B) succinate-grown cells

analysis showed that 45 protein spots were up-regulated (more intense) in gels of extracts from phenol-grown cells (data not shown). The 10 protein spots with the greatest intensity differences are respectively numbered 1-10 in Fig. 7A: the normalized spot volumes of the ten protein spots can be compared in Table 4. In addition, eight intense protein spots appear with equal intensity in cells obtained under both growth conditions were selected for mass spectrometry analysis in order to learn about proteins involved in normal cellular functions, which still remains poorly understood in *T. cutaneum* (Fig. 7B).

IV. Cross-species protein identification by mass spectrometry

Since the *T. cutaneum* genome is not sequenced and no public genome-sequencing effort is underway, the biggest problem after having solved the sample preparation issues was identifying proteins from *T. cutaneum* with the limited amount of information available in protein databases. The number of proteins that have been characterized in the Swiss-Prot and TrEMBL protein databases is only four from the genus *T. cutaneum* (as of May 2004). Nanospray MS/MS and LC-MS/MS have been used for generating peptide sequences which can be employed in identifying unknown proteins by cross-species identification (Wilkins and Williams, 1997; Shevchenko *et al.*, 2000). Therefore to enhance the number of proteins that could be identified from *T. cutaneum,* cross-species identification against all species of fungi was used to conduct protein database searches. This allowed searches against species such as the yeast *S. cerevisiae*, the dimorphic yeast *C. albicans*, the filamentous fungus *N. crassa* where complete genome sequence exist, and other filamentous fungi where genome sequencing is underway, such as *Aspergillus sp.* The generated amino acid sequences were identified by MASCOT (Perkins *et al.*,

Table 4. Comparison of normalized volumes of selected protein spots from 2-DE gel of crude extracts of *T. cutaneum* under phenol- and succinate- growth conditions

| Spot | phenol-induced protein | succinate-induced protein |
|---|---|---|
| 1 | 0.1180 | 0.0064 |
| 2 | 0.1760 | 0.0658 |
| 3 | 0.2601 | 0.0074 |
| 4 | 0.2417 | - |
| 5 | 0.2398 | - |
| 6 | 0.1538 | 0.1047 |
| 7 | 0.3335 | - |
| 8 | 0.0668 | - |
| 9 | 0.0591 | - |
| 10 | 0.2466 | - |

1999) or by manual *de novo* sequencing, then were aligned to homologues in the database using MS BLAST (Shevchenko *et al.*, 2000; Shevchenko *et al.*, 2002).

1. Analysis of spots from 2-DE gels of crude protein extracts

Seven of ten proteins highly expressed in phenol-grown cells (Fig. 7) were successfully analyzed by *de novo* sequencing, and five proteins were identified by sequence similarity using the MS/MS sequences and *de novo* interpretation (Table 5). Spots P1 and P2 were identified as heat-shock proteins (HSP70), which are very important in protein folding and degradation of damaged proteins: such activities are important in allowing organisms to survive during periods of environmental stress (Hartl and Hayer-Hartl, 2002). HSP70 proteins comprise a large family of highly related proteins with chaperone activity that are used for diverse cellular functions including assisting with the post-translational unfolding and translocation of nuclear-encoded proteins through the lipid bilayers of organelles (Bukau and Horwich, 1993). An HSP70 family protein also was identified in spot P8. These proteins probably play a role in facilitating the assembly of multimeric protein complexes inside the ER (van der Heide *et al.*, 2000). The expression of three different HSP70 proteins during growth on phenol, suggests that they may play a critical role in providing tolerance against aromatic stress.

Spot P10 was surprisingly identified as phenol hydroxylase, one of the few proteins fully sequenced from *T. cutaneum*, but the experimental molecular mass (19.5 kDa) is far from its actual size (75.2 kDa). This observation indicates that it is not sufficient to include

Table 5. nanoESI-MS/MS sequences and MS-BLAST similarity searches* of proteins upregulated in phenol-grown cells

| Spot No. | $Mr_{exp}$ (kDa) | $pI_{exp}$ | m/z | Sequence | Homologous protein | Organism | Accession No. |
|---|---|---|---|---|---|---|---|
| P-1 | 76.2 | 6.0 | 894.4 (2+) | IINEPTAAAIAYGLDKK | HSP70 | *P. graminis* | Q01877 |
| | | | 790.8 (2+) | AVITVPAYFNDSQR | | | |
| P-2 | 80.4 | 5.62 | 783.4 (2+) | LLGNFQLTGIPPAPK | HSP70 | *S. pombe* (Fission yeast) | P22774 |
| | | | 614.3 (2+) | DAGKIAGLDVLR | | | |
| P-3 | 66.7 | 5.56 | 791.4 (2+) | AAVEEGIVAGGGVALIR | | | Unidentified protein |
| P-5 | 31.4 | 5.40 | 546.8 (2+) | TLVTQLFDR | | | Unidentified protein |
| | | | 575.3 (2+) | GTLVTQLFDR or TLVTQLFDR with +57 modification | | | |
| P-7 | 21.9 | 5.34 | 620.8 (2+) | VDYEPNFDLK | STB2 protein | *S. cerevisiae* | P46679 |
| P-8 | 20.2 | 5.60 | 607.8 (2+) | VSGTIAGLEVLR | HSP70 | *P. angusta* (yeast) | Q9HG01 |
| | | | | | | *C. curvatus* | P87036 |
| P-10 | 19.5 | 6.09 | 534.8 (2+) | IPDTLPGISR | Phenol Hydroxylase | *T. cutaneum* | P15245 |
| | | | 728.3 (2+) | HILDSIAEVSDTR | | | |
| | | | 458.9 (3+) | HILDSIAEVSDTR substitution (V-I) | | | |
| | | | 432.8 (2+) | VLSEYVR | | | |
| | | | 935.9 (2+) | VLIVGAGPA | | | |
| P-11# | 32 | 8.5 | 632.3 (2+) | QDVVLLPQVPR | Malate dehydrogenase | *T. emersonii* | Q8X1C8 |
| | | | | | | | Q8TG27 |

| | | | | |
|---|---|---|---|---|
| 610.3 (2+) | LFGVTTLDVVR | Hypothetical protein | *N. crassa* | Q7S1D6 |

$Mr_{exp}$. Observed molecular weight from 2-DE gel.

$pI_{exp}$. Observed isoelectric point from 2-DE gel.

# This spot was obtained from IPG 3-10 gel (gel image not shown).

* Above MS-BLAST similarity searches are against Swiss-Prot and Tremble databases. The identified sequences shown the table correspond to the protein with greater sequence coverage (above 80%). The organism shown in the table correspond with the highest score provided by MS-BLAST. "L" (Leucine) and "I" (Isoleucine) in the peptide sequences were not differentiated by low-energy ESI-MS/MS. The "L" or "I" shown in the table correspond to the protein sequences with the closest hit provided by PROWL (Rockefeller University).

protease inhibitors during cell lysis and protein extraction, but they also may need to be included in the IEF steps, *e.g.* reswelling buffer (Finnie and Svensson, 2002). It also suggests that the proteases of *T. cutaneum* are very robust and are able to withstand the strong denaturing conditions of sample preparation: this may relate to this organism's survival ability under harsh soil-based conditions. A yeast SIN3 binding protein (STB2) was found in spot P7. As a transcriptional repressor, SIN3 may play an important role in regulating mitochondrial respiratory activity (Pile *et al.*, 2003). Moreover, the sequence in spot P11 (observed on a gel with first dimension pH 3-10, data not shown) was identified as a hypothetical protein as well as malate dehydrogenase which is involved in a central metabolic process. Since one of the products of phenol degradation in *T. cutaneum* is succinyl-CoA, which feeds into the Krebs cycle and is converted to malate, malate dehydrogenase may be elevated for that reason. The presence of these two proteins implies that phenol degradation results in a change in regulation of energy metabolism although the exact relationship is less clear.

Spots P3 and P5 could not be identified based on the sequences obtained from the Fungi database. However, spot P3 was similar to a 60 kDa ATP-binding chaperon in from bacteria (Accession No. P94798). It suggests that some of the proteins from *T. cutaneum* are more like bacteria than fungi. Five of eight proteins expressed equally under both growth conditions were analyzed by MS/MS but none of these proteins could be identified by similarity searches (Table 6). One obvious reason for this is that the protein database for *T. cutaneum* is extremely small, and even though the similarity search was

Table 6. nanoESI-MS/MS sequences and MS-BLAST similarity searches* of proteins expressed in both phenol- and succinate-grown cells

| Spot No. | $Mr_{exp.}$ (kDa) | $pI_{exp.}$ | m/z | Sequence | Homologous protein | Organism | Accession No. |
|---|---|---|---|---|---|---|---|
| S-2 | 40.4 | 5.42 | 564.4 (2+) | GLYEL | | | Unidentified protein |
| S-3 | 30.6 | 5.37 | 453.8 (2+) | RPYEKGR | | | Unidentified protein |
| S-4 | 26.9 | 6.55 | 584.3 (2+) | VGVVDPYS and NCKY | | | Unidentified protein |
| S-6 | 22.0 | 6.60 | 458.0 (2+) | DAGKIAAGLD | | | Unidentified protein |
| S-7 | 22.7 | 5.48 | 577.3 (2+) | SSGTSYPDVLK | | | Unidentified protein |

$Mr_{exp.}$ Observed molecular weight from 2-DE gel.

$pI_{exp.}$ Observed isoelectric point from 2-DE gel.

* Above MS-BLAST similarity searches are against Swiss-Prot and Tremble databases. The identified sequences shown the table correspond to the protein with greater sequence coverage (above 80%). The organism shown in the table correspond with the highest score provided by MS-BLAST. "L" (Leucine) and "I" (Isoleucine) in the peptide sequences were not differentiated by low-energy ESI-MS/MS. The "L" or "I" shown in the table correspond to the protein sequences with the closest hit provided by PROWL (Rockefeller University).

employed against all known proteins from Fungi, only highly conserved proteins such as HSPs were identified.

2. Analysis of spots from 2-DE gels after pre-fractionation with anion exchange chromatography

Even though phenol hydroxylase, a major enzyme in the phenol degradation pathway was detected by analysis of crude extracts, another previously sequenced pathway enzyme, *cis, cis*-muconate cyclase, was not. A possible explanation for this is that the abundance of this enzyme was not high enough to be detected in the whole cell extract, although its theoretical pI/MW (5.02/41kD) place it within the 2-DE separation range. Alternatively, this protein might overlap with a more abundant one on the 2D gel. To improve the possibility of identifying additional proteins that are less abundant, crude extracts from phenol-induced cell were applied to a DEAE anion exchange column to reduce the complexity of the total protein mixture. Many of the enzymes involved in phenol degradation in this organism elute within a gradient of 0-0.3 M NaCl (Powlowski and Dagley, 1985). As described in *Material and Methods*, protein fractions were collected by eluting at 0.3 M NaCl and were shown to contain phenol hydroxylase activity. Fig. 11 shows protein expression profiles by 2-DE of these DEAE fractions of proteins from *T. cutaneum* cultured in succinate and phenol containing media. About 170 spots were detected for each growth condition, and about 30 spots were enriched after growth on phenol. Five spots were successfully analyzed by using nano-ESI MS/MS (Fig. 12), and identified by Mascot and BLAST searches (Table 7). Spot PF2 was similar to a hypothetical protein and two central metabolic enzymes: 3-isopropylmalate

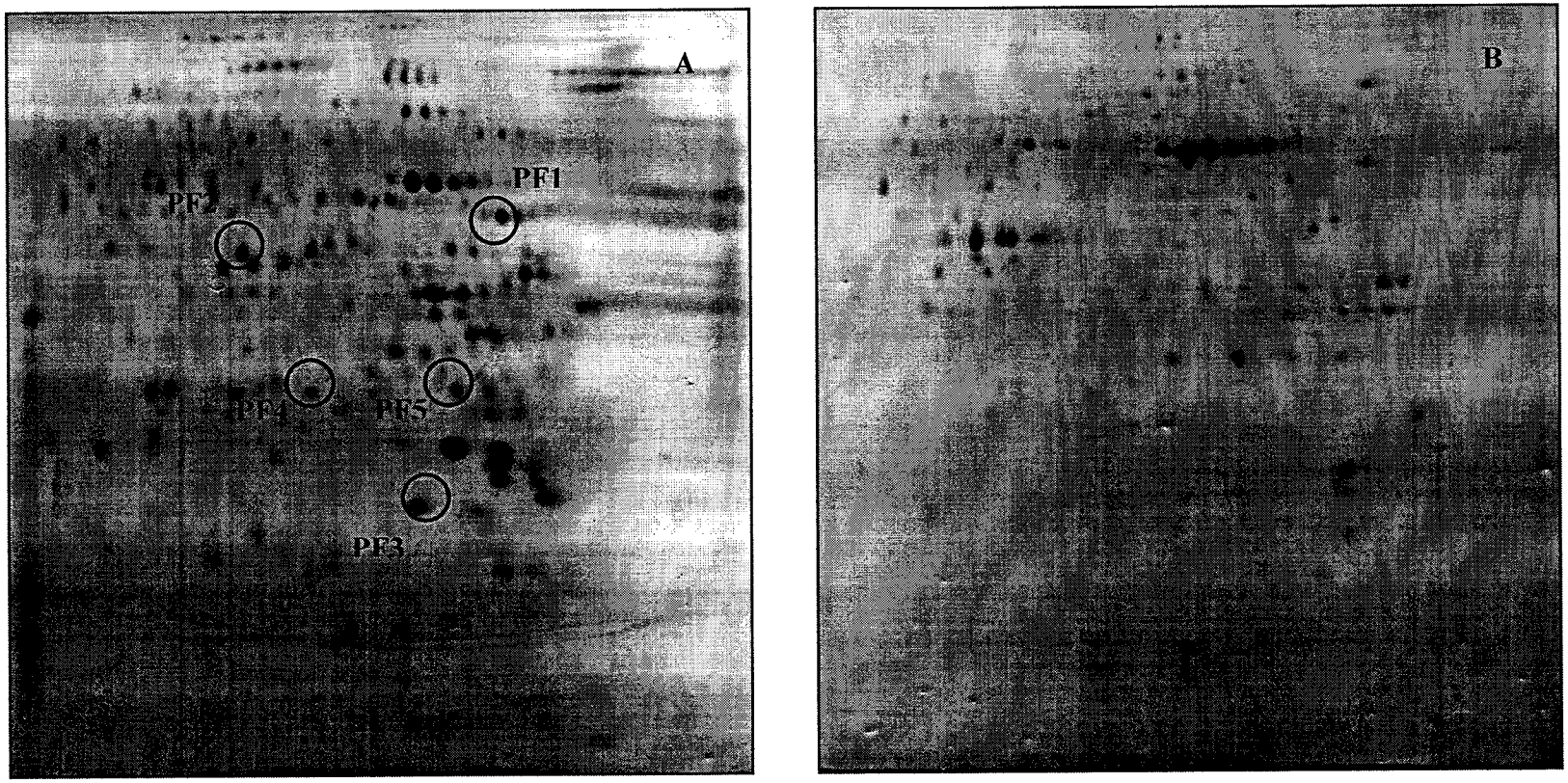


**Figure 11.** 2-DE gels showing protein maps of DEAE fractions from *T. cutaneum* grown with (A) phenol and (B) succinate. Solid line circles indicate enriched proteins in extracts from phenol-grown cells. The DEAE fractions from cells grown on succinate were used as control.

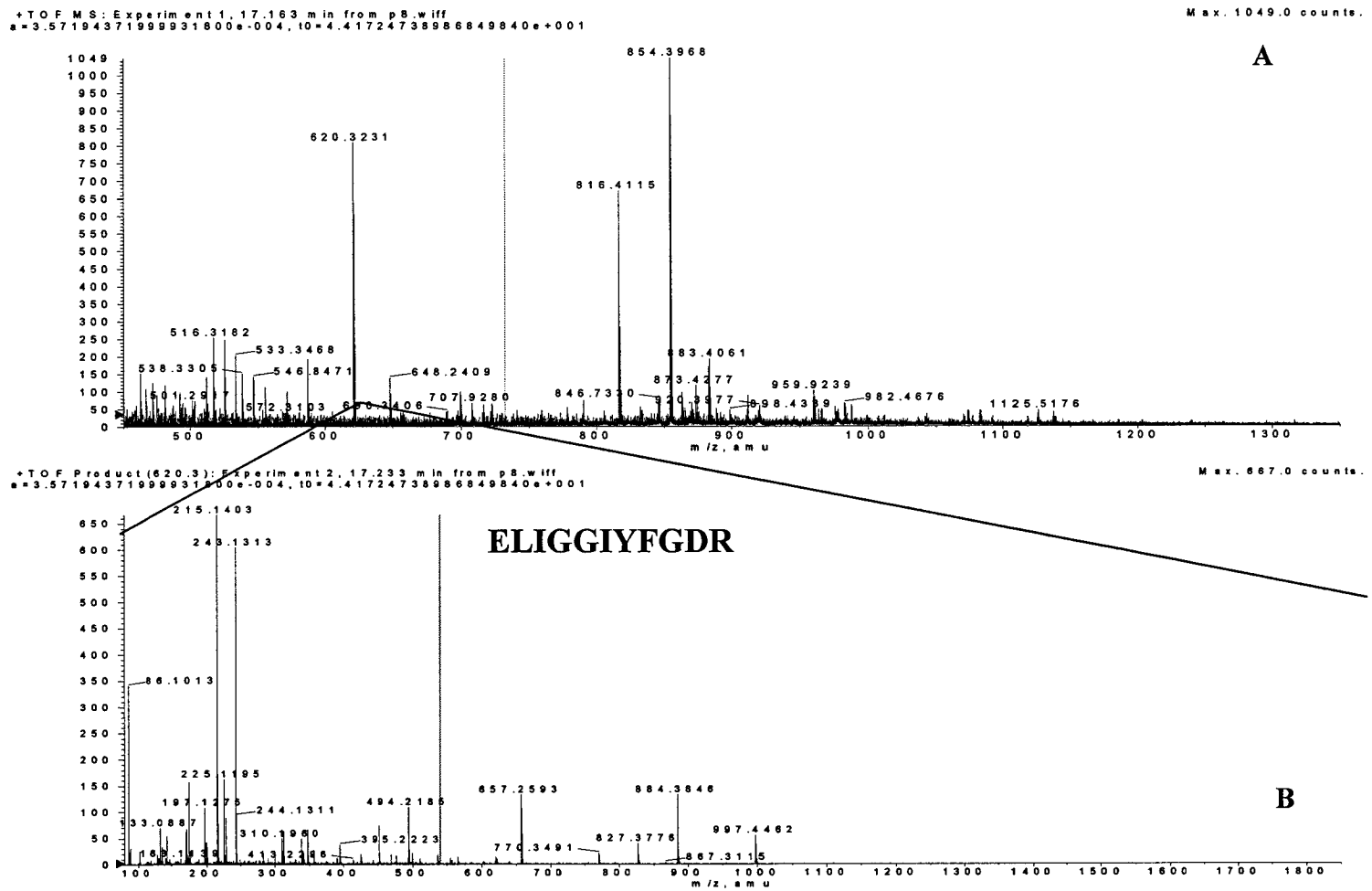


Figure 12. nanoESI-MS/MS spectrum of spot PF-1. The parent ions (A) and one of the daughter ion (m/z=620.30, 2+) (B) shown the generated peptide sequence.

Table 7. nano-ESI-MS/MS sequences and MS-BLAST similarity searches* of proteins from DEAE fractions of phenol-grown cells

| Spot No. | $Mr_{exp.}$ (kDa) | $pI_{exp.}$ | m/z | Sequence | Homologous protein | Organism | Accession No. |
|---|---|---|---|---|---|---|---|
| PF-1 | 52.1 | 5.0 | 620.30 (2+) | ELIGGIYFGDR | 3-isopropylmalate dehydrogenase | *A. niger* | P87256 |
| | | | 625.30 (2+) | HRCAGY(IL)ENK | | | |
| | | | 701.30 (2+) | V(IL)DSSADGGYDFMR | | | |
| PF-2 | 47.9 | 6.06 | 610.40 (2+) | LFGVTTLDVVR | Hypothetical protein | *N. crassa* | Q7S1D6 |
| | | | 931.50 (2+) | PSVTEL | | | |
| | | | 989.0 (2+) | PVNSTV | | | |
| PF-3 | 24.0 | 5.35 | 512.27 (2+) | NEFQLSASK | Ferri reductase transmembrane component | *S. cerevisiae* | P63746 |
| | | | 555.25 (2+) | APLLSAD | | | |
| PF-4 | 32.7 | 5.78 | 789.40 (2+) | GAFTGEISP | Triosephosphate isomerase | *A. oryzae* | Q9HGY8 |
| | | | 712.30 (2+) | V(IL)ACIGETFEER | | | |
| PF-5 | 32.7 | 5.18 | 688.37 (2+) | NSTINDLMLLK | SIP1 protein | *S. cerevisiae* | P32578 |
| | | | 696.40 (2+) | Met oxidization of 688.37 | | | |
| | | | 656.30 (2+) | DDPDVV | | | |

$Mr_{exp.}$ Observed molecular weight from 2-DE gel.
$pI_{exp.}$ Observed isoelectric point from 2-DE gel.
* Above MS-BLAST similarity searches are against Swiss-Prot and Tremble databases. The identified sequences shown the table correspond to the protein with greater sequence coverage (above 80%). The organism shown in the table correspond with the highest score provided by MS-BLAST. "L" (Leucine) and "I" (Isoleucine) in the peptide sequences were not differentiated by low-energy ESI-MS/MS. The "L" or "I" shown in the table correspond to the protein sequences with the closest hit provided by PROWL (Rockefeller University).

dehydrogenase A (spot PF-1) and triosephosphate isomerase (spot PF-4). A protein kinase substrate, SIP1 protein (spot PF-5) and an integral membrane protein, ferric reductase transmembrane component 4 (spot PF-3) were also identified. However, no phenol degradation pathway enzymes have been identified from these samples.

An additional set of samples prepared from spots excised from 2-DE gels prepared either from crude extracts or from DEAE fractions was obtained from Dr. L. Sahlman. A total of 16 protein spots were excised from the gels and first analyzed by MALDI-QTOF-MS. Only two spots could be identified by peptide mass fingerprinting, and both were phenol hydroxylase (spot PCE-8 and PE-2) (Fig. 13 A and 13 B). This confirmed that the peptide mass profiles generated from MALDI-QTOF-MS were not an efficient tool for identifying proteins with no known sequences in the databases. When the same set of samples were analyzed by CapLC-MS/MS, identities of 12 spots were obtained (Table 8). In this sample set, the pathway enzymes, phenol hydroxylase and *cis, cis*-muconate cyclase I, were detected in more than one of the analyzed spots (Table 8). Four additional proteins were identified as isocitrate dehydrogenase (spot PE-1), inorganic pyrophosphatase (spot P-4), enolase (spot SEE) and malate dehydrogenase (spot SCE-3). Among these spots, spot SEE produced good matches to enolase, mRNA polymerase II subunit and coronin-like protein. Matches to three unrelated proteins indicated that this spot might represent the comigration of three proteins. Similar occurrences of comigrating proteins were reported in a recent evaluation of 2-DE gel based techniques with *S. cerevisiae* and *E. coli* (Corthals *et al.*, 2000). A transcriptional regulation-related

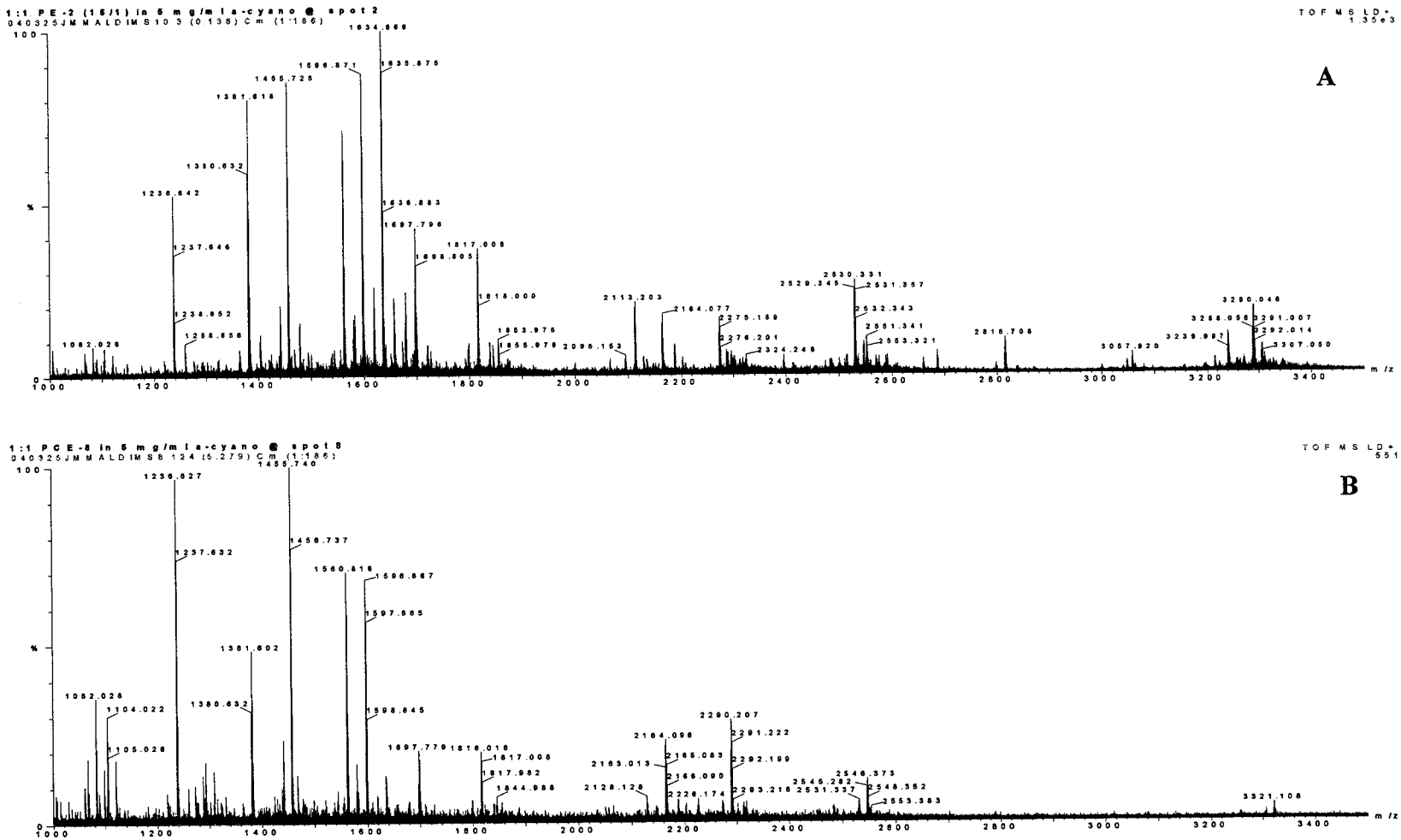


Figure 13. MALDI-QTOF-MS spectrum shows phenol hydroxylase at spot PE-2 (A) and PCE-8 (B).

Table 8. Cap-LC-MS/MS sequences and MS-BLAST similarity searches* of proteins from samples obtained from Dr. L. Sahlman

| Spot No. | Mr (kDa) | pI | M/Z | Sequence | Homologous protein | Organism | Accession No. |
|---|---|---|---|---|---|---|---|
| PE-2 | 19.5 | 6.0 | 534.76 (2+) | IPDTLPGISR | phenol hydroxylase | *T. cutaneum* | P15245 |
| | | | 618.78 (2+) | YHQVVLHQGR | | | |
| | | | 701.35 (2+) | TLESLKNLGLADK | | | |
| | | | 566.22 (3+) | STKVYNGQADGLQCR | | | |
| PE-4 | 32 | 6.4 | 594.82 (2+) | FTPEVVLANAK | phenol hydroxylase | *T. cutaneum* | P15245 |
| | | | 636.78 (2+) | EAGEIETVHCK | | | |
| | | | 698.85 (2+) | FYVQLQVDRTK | | | |
| | | | 709.38 (2+) | TKFTPEVVIANAK | | | |
| | | | 770.35 (2+) | VFIAGDACHTHSPK | | | |
| | | | 513.90 (3+) | VFIAGDACHTHSPK | | | |
| PE-6 | 34 | 5.2 | 690.76 (2+) | VYNGQADGLQCR | phenol hydroxylase | *T. cutaneum* | P15245 |
| | | | | | | *T. beigelii* | Q7M4T8 |
| PCE-8 | 19 | 6.0 | 534.79 (2+) | IPDTLPGISR | phenol hydroxylase | *T. cutaneum* | P15245 |
| | | | 690.79 (2+) | VYNGQADGLQCR | | | |
| PE-1 | 23 | 5.1 | 690.31 (2+) | LSPMADVVFGSTR | muconate cycloisomerase I | *T. cutaneum* | P46057 |
| | | | 497.20 (2+) | HAFGDQYR | isocitrate dehydrogenase | *A. niger* | P79089 |
| | | | | | | *S. cerevisiae* | P53982 |
| | | | | | | *A. nidulans* | Q96UN7 |
| | | | | | | *A. nidulans* | Q96UN6 |
| | | | | | hypothetical protein | *N. crassa* | Q7RW77 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PEE | 48 | 5.25 | 560.76 (2+)<br>690.31 (2+)<br>698.31 (2+) | ANGTLDFGGLR<br>LSPMADVVFGSTR<br>LSPMADVVFGSTR + oxidation | muconate cycloisomerase I | *T. cutaneum* | P46057 |
| PCE-6 | 29.5 | 5.4 | 690.33 (2+) | LSPMADVVFGSTR | muconate cycloisomerase I | *T. cutaneum* | P46057 |
| PE-3 | 28 | 5.9 | 530.27 (2+) | LADKLDSAVK | mating-type protein beta 1 | *C. cinereus* | Q9Y7A5<br>Q9Y7A6 |
| PE-7 | 78 | 5.2 | 614.83 (2+) | ADLVNNLGTIAK | HSP90 | *C. albicans*<br>*C. tropicalis*<br>*C. neoformans* | P46598<br>Q9P8I3<br>Q71QT8 |
| P-4 | 35 | 5.2 | 551.74 (2+) | LNDIEDVER | inorganic pyrophosphatase | *S. pombe*<br>*C. pelliculosa* | P19117<br>Q875I2 |
| PCE-3 | 23 | 5.1 | 656.85 (2+) | VHLVAIDIFTGK | eukaryotic translation initiation factor 5A | *N. crassa*<br>*C. albicans*<br>*S. cerevisiae*<br>*E. gossypii* | P38672<br>O94083<br>P23301<br>Q753F8 |
| SEE | 44 | 5.3 | 451.73 (2+)<br>768.34 (2+)<br>717.33 (2+)<br>669.31 (2+) | ACNALLLK<br>IEEELGENAIYAGK<br>QIGIWDAFNIEK<br>HIKWLNQGWR | enolase<br><br>coronin-like protein<br>RNA polymerase II subunit RPB2 | *P. citrinum*<br><br>*S. cerevisiae*<br>*B. Montana* | Q96X46<br><br>Q06440<br>Q873V9 |
| SCE-3 | 23 | 5.2 | 546.28 (2+) | IQFGGDEVVK | malate dehydrogenase | *S. cerevisiae*<br>*S. cerevisiae*<br>*S. pombe*<br>*P. brasiliensis*<br>*P. orbiculare* | P17505<br>Q6Q5N4<br>Q9Y7R8<br>Q7ZA65<br>Q9Y750 |

$Mr_{exp}$. Observed molecular weight from 2-DE gel.

$pI_{exp}$. Observed isoelectric point from 2-DE gel.

* Above MS-BLAST similarity searches are against Swiss-Prot and Tremble databases. The identified sequences shown the table correspond to the protein with greater sequence coverage (above 80%). The organism shown in the table correspond with the highest score provided by MS-BLAST. "L" (Leucine) and "I" (Isoleucine) in the peptide sequences were not differentiated by low-energy ESI-MS/MS. The "L" or "I" shown in the table correspond to the protein sequences with the closest hit provided by PROWL (Rockefeller University).

protein, mating-type protein beta 1 (spot PE-3), a translation factor (spot PCE-3) and an HSP90 protein (PE-7) were also highly expressed during growth on phenol. The potential significance of HSP-90 expression was discussed earlier, but the roles of the other proteins in phenol degradation are not obvious. Apart from the proteins with classified functions, a protein with unknown function was found in spot SCE-6, which was more highly expressed during growth on succinate.

Overall, prefractionation by DEAE chromatography allowed the detection of some proteins that were difficult to detect in crude extracts. Cross-species identification by two tandem mass spectrometry methods was successfully applied in discovering some proteins involved in phenol degradation, as well as in the other fundamental processes of protein synthesis, protein folding, energy generation and electron transport despite the absence of genome sequence data for *T. cutaneum*. More genome sequence data will undoubtedly be required for further in-depth analysis of the proteomic study of aromatic degradation in *T. cutaneum*.

V. Other preliminary work and future considerations

1. Investigation of protein expression during growth with *p*-cresol vs. phenol

*T. cutaneum* is capable of using various aromatic compounds as sole carbon and energy sources. In addition to phenol, its methylated derivative, *p*-cresol, has been well studied in terms of catabolic pathway and enzymology (Powlowski and Dagley, 1985a; 1985b). As no proteomic studies have been done to compare overall protein expression during

growth on these two similar carbon sources, a proteomic analysis of *p*-cresol degradation by *T. cutaneum* was initiated.

After growth of *T. cutaneum* on *p*-cresol as the sole carbon and energy source, cells were harvested and extracts were prepared for 2-DE using the same procedures as for crude extracts from phenol-induced cells. From one batch of *p*-cresol-grown cells, about 250 spots were detected on a 2-DE gel (Fig. 14). Sixty spots were expressed differently compared to phenol-grown cells: 16 spots were up-regulated in *p*-cresol-grown cells from which 8 spots appeared to be present in *p*-cresol-grown cells whereas 44 spots were more highly expressed in phenol-grown cells and 12 spots were only detected in exacts from phenol-grown cells. However, no mass spectrometry results have yet been obtained for these proteins. These are preliminary data since the reproducibility of patterns from different batches of cells has not been examined. However, it is clear from the results presented here that there are major differences in protein expression related to the presence of the methyl group on the ring. Although it has been shown that the enzymology of the degradation of *p*-cresol is similar to that of phenol degradation (Powlowski and Dagley, 1985a; 1985b), cells grow much more slowly in the presence of *p*-cresol compared to phenol (Fig. 15), and only tolerate 0.02% *p*-cresol as compared to 0.03% phenol). These observations suggest that *p*-cresol is more toxic than phenol, and it may be that the differences in protein expression are a response to the different toxicities of the two growth substrates or differences in growth rate.

2. The membrane proteome of *T. cutaneum* during growth on phenol

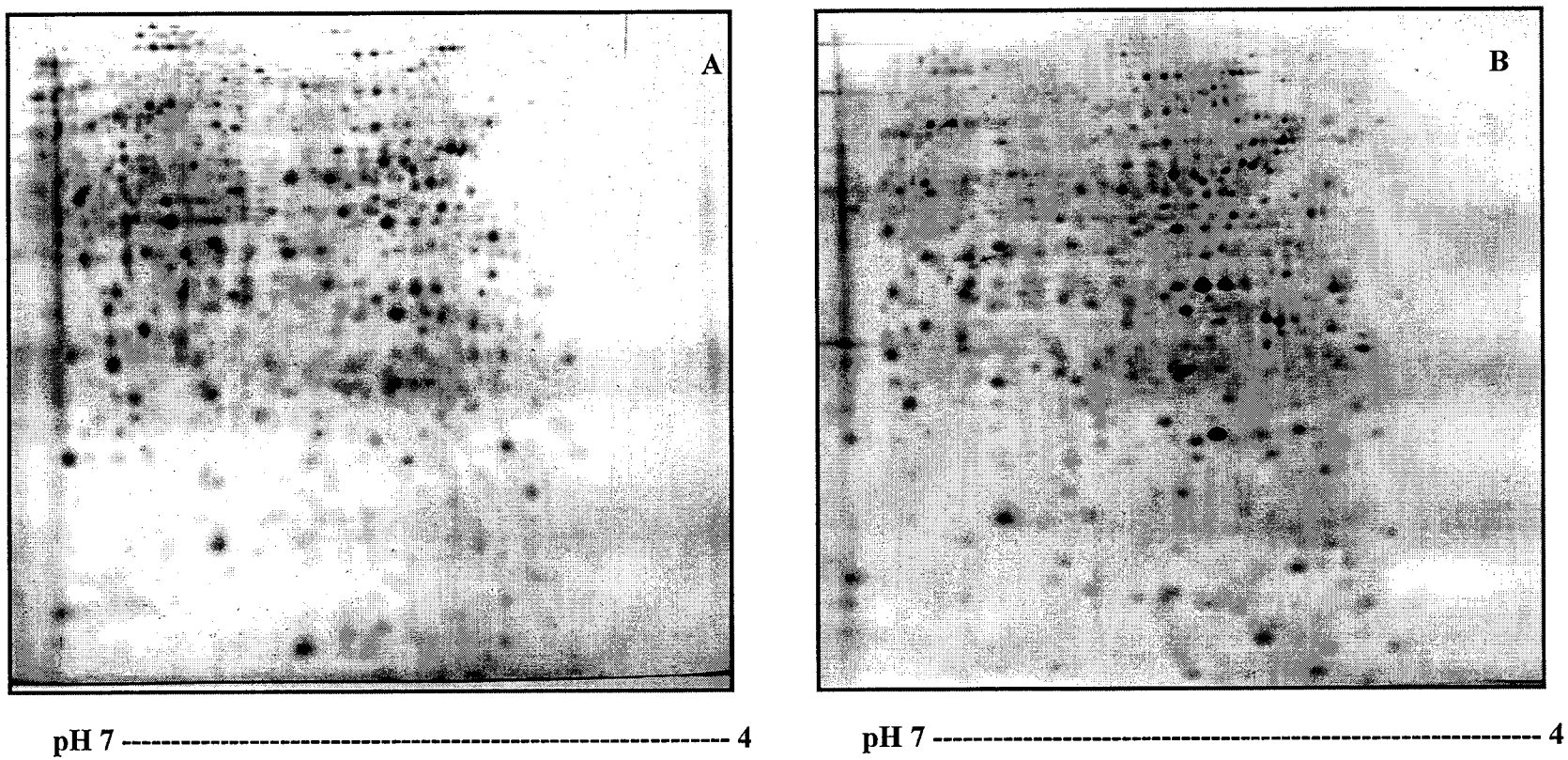


**Figure 14.** 2-DE gels of crude protein extracts from *T. cutaneum* grown with (A) *p*-cresol and (B) phenol, separated at pH 4-7, 11 cm IPG strips.

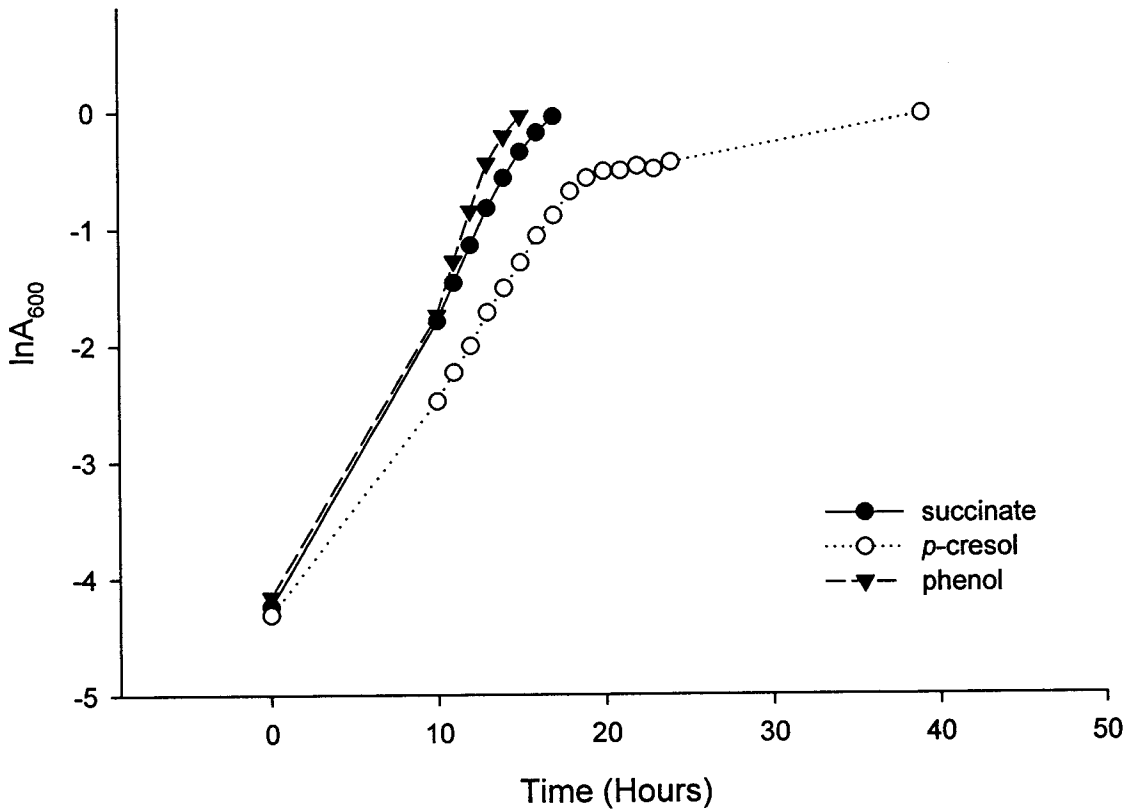


**Figure15.** Growth curve of *T. cutaneum* grown with succinate, phenol and *p*-cresol.

Membrane proteins play important roles in various cellular functions including signal transduction, toxin resistance and metabolism and transport processes. Although proteomics technologies have made rapid progress in the analysis of soluble proteins in recent years, proteomics analysis of membrane proteins has lagged behind. Only a few reports on membrane or organellar proteomes (e.g., mitochondria and plasma membrane) in fungi and yeast (Grinyer *et al.*, 2004; Navarre *et al.*, 2002) have been published, but none of them examine an entire complement of membrane proteins present in a cell under a specific condition. In particular, no publications have been reported for membrane proteomes during microbial growth on aromatic compounds. It was, therefore, of interest to study the membrane proteins expressed in response to aromatic degradation in *T. cutaneum* as they may be involved in transporting aromatic compounds into cell, or resistance to toxic effects of phenol.

A preliminary study involved extraction of membrane proteins from *T. cutaneum* grown on phenol or succinate, followed by 2-DE separation (Fig. 16). The result was not very satisfactory since many soluble protein spots were still present in the membrane extracts, and some horizontal streaking was visible on the gel, especially for the sample from phenol-grown cells. These problems indicated that the sample preparation method must be modified to enrich for membrane proteins, and the problem of incomplete isoelectric focusing must be addressed. Since there is no report of a global membrane protein extraction method for yeast and fungi, the method used for the preparation of membrane proteins for 2-DE was based on some principles referred to in Molloy's comprehensive review (Molloy, 2000). For future work, sub-fractionation will be required to enrich for

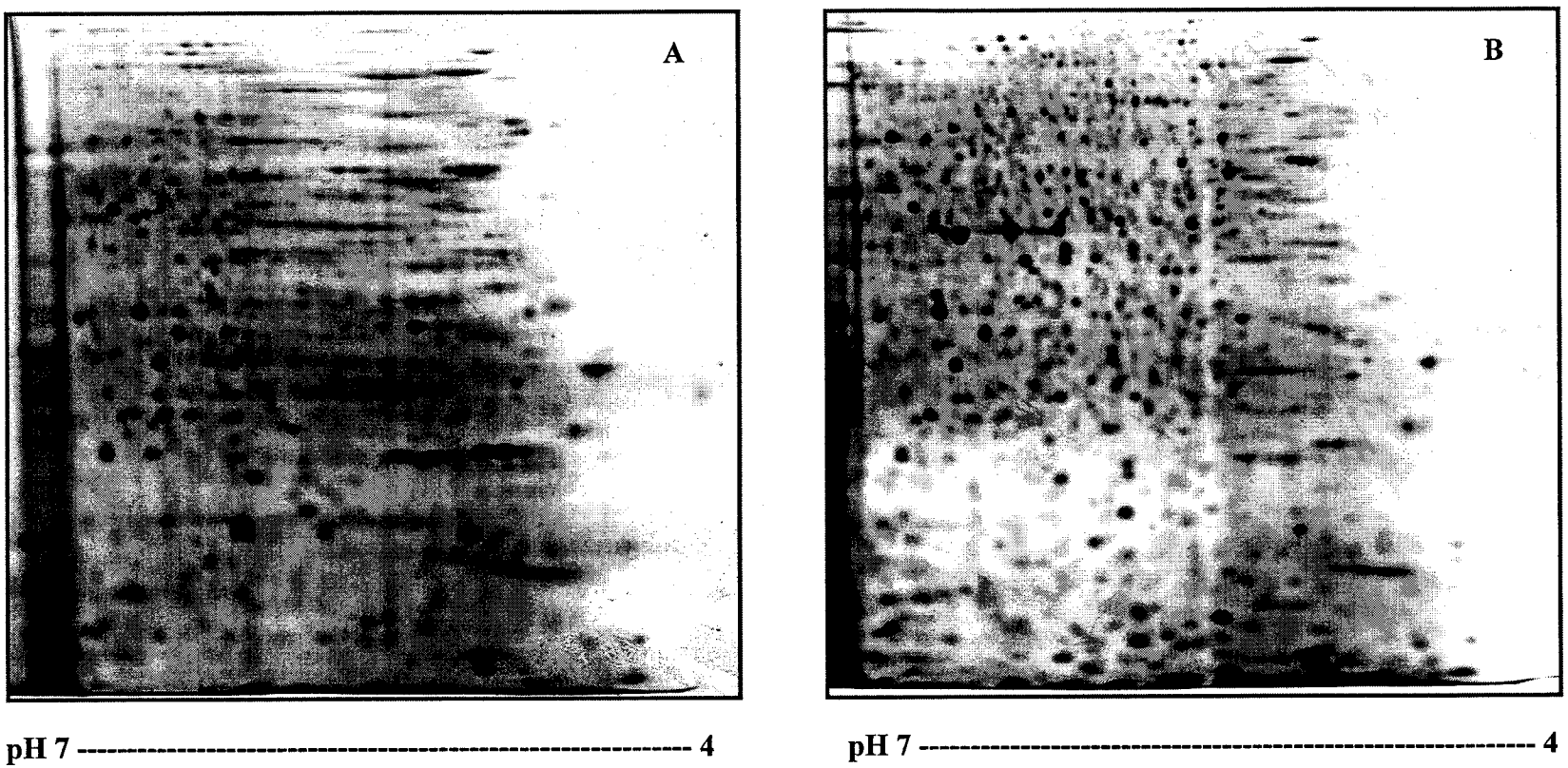


**Figure 16.** 2-DE gels of membrane proteins from *T. cutaneum* grown with (A) phenol and (B) succinate, separated at pH 4-7, 11 cm IPG strips.

membrane proteins. A recent publication reported a method used to prepare membrane proteins for proteomic analysis by using high pH membrane fractionation then followed by digestion with proteinase K. This procedure is optimized specifically for the global analysis of both membrane proteins and soluble proteins from membrane-containing samples (Wu *et al.*, 2003). A second problem associated with 2-DE of membrane proteins is that these proteins become increasingly insoluble as they near their isoelectric points (pI). An alternative separation strategy is the use of 2-D blue native gel electrophoresis (Brooks et al., 2002). In this method membrane protein complexes remain intact during the first dimension, which is performed under native conditions, and are subsequently resolved during second-dimension SDS-PAGE. Technical modifications such as these may allow the study of the changes in the membrane proteome of *T. cutaneum* in response to growth on aromatic compounds.

## VI. Concluding remarks

In summary, proteomic analysis combining 2-DE and MS/MS has been successfully applied in studying proteins of *T. cutaneum* differentially expressed in response to growth on phenols. The optimized sample preparation methods and experimental conditions provide a great advantage in quantitative analysis and potential high-throughout applications. For the first time, the global effects on the *T.cutaneum* proteome caused by phenol have been examined. Cross-species identification provides a powerful tool in protein identification with an unknown genome. Except for phenol hydroxylase and *cis, cis* - muconate cyclase, the other 24 proteins identified from *T. cutaneum* were not previously known. The sequences obtained from these proteins can be used for isolating

full-length genes to confirm the identifications made based on the mass spectrometry results, and to more fully characterize these genes. Once the genes are identified, the encoded proteins can be expressed and purified to study what effects over-expression have on, for example, sensitivity of the cells to phenol. Also, it might be possible to knock out some of these genes to examine what effect that has on phenol degradation. In most cases, we are still for from being able to connect the identified protein with an actual function in phenol degradation. However, the work presented in this thesis reveals an initial whole-cell protein reference map for the model fungus *T. cutaneum* and will aid further proteomic-based studies of this organism, and other fungi in, better understanding the mechanisms involved in aromatic degradation.

## Reference


Aebersold, RH., Leavitt, J., Saavedra, RA., Hood, LE. And Kent, SB. (1987) *Proc. Natl. Acad. Sci. USA.* **84(20):** 6970-4

Anderson, JJ. and Dagley, S. (1980) *J.Bacteriol.* **141(2):** 534-543

Anderson, JJ. and Dagley, S. (1981) *J.Bacterio.***146(1):** 291-7

Anderson, L. and Seilhamer, J. (1997) *Electrophoresis* **18:** 533-537

Babnigg, G. and Giometti, CS. (2004) *Nucleic Acids Res.* **32:** D582-585

Barber, M., Jolles, P., Vikas, E. and Lederer, E. (1965) *Biochem. Biophys. Res. Comm.* **18:** 469-473

Behrend, G. (1890) *Trichomycosis nodosa. Ber Klin Wochemmschr* **21:** 464-467

Bier, M. (1998) *Electrophoresis* **19:** 1057-1063

Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, JC and Hochstrasser, D. (1993) *Electrophoresis* **14(12):** 1357-65

Blomberg, A., Blomberg, L., Norbeck, J. (1995) *Electrophoresis* **16:** 1935-1945

Bradford, MM. (1976) *Anal. Biochem.* **72:** 248-254

Brooks, PS., Pinner, A., Ramachandran, A., Coward, L., Barnes, S., Kim, H. and Darley-Usmar, VM. (2002) *Proteomics* **2:** 969-977

Bruneau, JM., Magnin, T., Tagat, E., Legrand, R., Bernard, M., Diaquin, M., Fudali, C. and Latge, JP. (2001) *Electrophoresis* **22:** 2812-2823

Bukau, B. and Horwich, AL. (1993) *Cell* **92:** 351-366

Burgoyne, L., Case, ME. and Giles, NH. (1969) *Biochim. Biophys. Acta.* **191(2):** 452-62

Corthals, GL., Wasinger, VC., Hochstrasser, DF. and Sanchez, JC. (2000) *Electrophoresis* **21:** 1104-15

Dagley, S. (1971) *Adv. Microb. Physiol.* **6(0):** 1-46

Dagley, S. (1985) *In* Moo-Young, M., Bull, AT., and Dalton, H. (ed.), Comprehensive biotechnology, vol. 1. Pergamon Press Ltd., Oxford. p. 483-505

Dagley, S. (1986) *In* Sokatch, JR. (ed.), *The bacteria*, 10: 527-555

Dancik, V., Addona, TA., Clauser, KR., Vath, JE. And Pevzner, PA. (1999) *J. Comput. Biol.* **6(3-4):** 327-42

Detmer, K.M. and Massey, V. (1985) *J. Biol. Chem.* **260:** 5998-6005

Duffner, FM., Kirchner, U., Bauer, MP., and Muller, R. (2000) *Gene* **256(1-2):** 215-21

Dunn, MJ. and Corbett, JM. (1996) *Methods Enzymol.* **271:** 177-203

Ehrt, S., Schirmer, F. and Hillen, W. (1995) *Mol. Microbiol.* **18(1):** 13-20.

Enroth, C., Huang, W., Waters, S., Neujahr, H., Lindqvist, Y. and Schneider, G. (1994) *J. Mol. Biol.* **238(1):** 128-30

Enroth, C., Neujahr, H., Schneider, G. and Lindqvist, Y. (1998) *Structure* **6(5):** 605-17

Enroth, C. (2003) *Acta. Crystallogr D. Biol. Crystallogr.* **59 (Pt 9):** 1597-1602

Fenn, JB., Mann, M., Meng, CK., Wong, SF. and Whitehouse, CM. (1989) *Science* **246:** 64-71

Finnie, C. and Svensson, B. (2002) *Anal. Biochem.* **11(2):** 182-6

Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL, Witney AA, Wolters D, Wu Y, Gardner MJ, Holder AA, Sinden RE, Yates JR, Carucci DJ. (2002) *Nature* **19(6906):** 520-6

Fountoulakis, M., Langen, H., Evers, S., Gray, C. and Takeacs, B. (1997) *Electrophoresis* **18:** 1193-1202

Fountoulakis, M., Takeacs, B. and Langen, H. (1998a) *Electrophoresis* **19:** 761-6

Fountoulakis, M., Juranville, JF., Roder, D., Evers, S., Berndt, P. and Langen, H. (1998b) *Electrophoresis* **19:** 1819-1827

Fountoulakis, M., Takeacs, B., Berndt, P. and Langen, H. (1999) *Electrophoresis* **20:** 2181-2195

Gaal, A. and Neujahr, HY. (1980) *Biochem. J.* **191(1):**37-43

Gaal, A. and Neujahr, HY. (1981) *Arch. Microbiol.* **130:** 54-8

Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H,

Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. (2002) *Nature* **415(6868):** 141-7

Giuffrida MG, Pessione E, Mazzoli R, Dellavalle G, Barello C, Conti A, Giunta C. (2001) *Electrophoresis* **22(9):** 1705-11

Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S., Mische, S.M. (1999) *Electrophoresis* **20:** 601-605

Görg, A., Postel, W., Günther, S. (1988) *Electrophoresis* **9:** 531-546

Görg, A., Obermaier, C., Boguth, G., Csordas, A., Diaz, JJ. And Madjar, JJ. (1997) *Electrophoresis* **18:** 328-337

Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R. and Weiss, W. (2000) *Electrophoresis* **21:** 1037-1053

Grinyer, J., McKay, M., Herbert, B. and Nevalainen, H. (2004) *Current Genetics* **45(3):** 170-175

Guy, G.R., Philip, R., Tan, Y.H. (1994) *Electrophoresis* **15:** 417-440

Gygi, SP., Rochon, Y., Franza, BR. and Aebersold, R. (1999) *Mol. Cell. Biol.* **19(3):** 1720-1730

Gygi, SP., Rist, B., Corthals, G. L., Zhang, Y., Rochon, Y. and Aebersold, R. (2000) *Proc. Nat. Acad. Sci. USA* **97:**9390-9395

Hara, SH., Murakami, H., and Oba, T. (1971) *J. Ferment. Technol.* **49:** 330-337

Harder, A., Wildgruder, R., Nawroki, A., Fey, S.J., Larsen, P.M., and Görg, A. (1999) *Electrophoresis* **20:** 826-829

Hartl, FU. & Hayer-Hartl, M. (2002) *Science* **295:** 1852-8

Hecker, M. (2003) *Adv. Biochem. Eng. Biotechnol.* **83:** 57-92

Henzel, WJ., Billeci, TM., Stults, JT., Wang, SC., Grimley, C. and Watanabe, C. (1993) *Proc. Natl Acad. Sci. USA* **90:** 5011-5015

Hochstrasser, D., Patchornik, A., and Merril, C. (1988) *Anal. Biochem.* **173:** 412-423

Hochstrasser, AC., James, RW., Pometta, D. and Hochstrasser, D. (1991) *Applied. Theor. Electrophor.* **1:** 333-337

Hodges, PE., McKee, AH., Davis, BP., Payne, WE. and Garrels, JI. (1999) *Nucleic Acids Res.* **27(1):** 69-73

Hoogland, C., Sanches, JC., Tonella, L., Binz, PA., Bairoch, A., Hochstrasser, DF. and Appel, RD. (2000) *Nucleic Acids Res.* **28:** 286-288

James, P., Quadroni, M., Carafoli, E. and Gonnet, G. (1993) *Biochem. Biophys. Res. Commun.* **195:** 58-64

Jazwinski, SM. (1990) *Methods Enzymol.* **182:** 154-174

Jiménez, C.R., Huang, L., Qiu, Y., and Burlingame, A.L. (1998) *Current Protocols in Protein Science* 16.4.1-16.4.5

Jones, KH., Trudgill, PW. and Hopper, DJ. (1995) *Arch. Microbiol.* **163(3):** 176-81

Kahng HY, Cho K, Song SY, Kim SJ, Leem SH, Kim SI. (2002) *Biochem. Biophys. Res. Comm..* **295(4):** 903-9

Kälin, M., Neujahr, HY., Weissmahr, R., Sejlitz, T., Jöhl, R., Fiechter, A., Käppeli, O., and Fiechter, A. (1981) *Biotechnol. Lett.* **3:** 541-6

Kanapin, A., Batalov, S., Davis, M.J., Gough, J., Grimmond, S., Kawaji, H., Magrane, M., Matsuda, H., Schonbach, C., Teasdale, RD., Yuan, Z; RIKEN GER Group; GSL Members. (2003) *Genome Res.* **13(6B):** 1335-44

Kamath, AV., Dasgupta, D. and Vaidyanathan, CS. (1987) *Biochem. Biophys. Res. Commun.* **145(1):** 586-95

Kamath, AV., Rao, NA. and Vaidyanathan, CS. (1989) *Biochem. Biophys. Res. Commun.* **165(1):** 20-6

Käppeli, O. and Fiechter, A. (1982) *Biotechnol. Bioeng.* **24:** 2519-2526

Karas, M. and Hillenkamp, F. (1988) *Anal. Chem.* **60:** 2299-2301

Kim SI, Song SY, Kim KW, Ho EM, Oh KH. (2003) *Res Microbiol.* **154(10):** 697-703

Klein, J., Hackel, U. and Wagner, F. (1979) *ACS Symp. Ser.* **106:** 101-118

Klose, J. (1975) *Human Genetik* **26:** 231-243

Klose, J. and Kobalz, U. (1995) *Electrophoresis* **16:** 1034-1059

Klose, J. (1999) *Electrophoresis* **20:** 643-652

Krug, M., Ziegler, H. and Straube, G. (1985) *J. Basic Microbiol.* **25(2):** 103-10
Krug, M. and Straube, G. (1986) *J. Basic Microbiol.* **26(5):** 271-81

Laaser, G., Möller, E., Jahnke, KD., Bahnweg, G., Prillinger, H., and Prell, HH. (1989) *Syst. Appl. Microbiol.* **11**: 170-5

Laemmli, UK. (1970) *Nature* **22:** 680-5

Lampkin, SL. $4^{H}$, Cole, KW., Vitto, A. and Gaertner, FH. (1976) *Arch. Biochem. Biophys.* **177(2):** 561-5

Langen, H., Gray, C., Roder, D., Juranville, JF., Takeacs, B. and Fountoulakis, M. (1997) *Electrophoresis* **18:** 1184-1192

Link, AJ., Eng, J., Schieltz, DM., Carmack, E., Mize, GJ., Morris, DR., Garvik, BM. and Yates, JR. III. (1999) *Nature Biotechnol.* **17:** 676-682

Lopez, MF., Berggren, K., Chernokalskaya, E., Lazarev, A., Robinson, M. and Patton, WF. (2000) *Electrophoresis* **21:** 3673-3683

Mann, M., Hojrup, P. and Roepstorff, P. (1993) *Biol. Mass Spectom.* **22:** 338-345

Mann, M., Hendrickson, RC. and Pandey, A. (2001) *Anu. Rev. Biochem* **70:** 434-73

Matsui, NM., Smit, DM., Clauser, KR., Fichmann, J., Andrews, LE., Sullivan, CM., Burlingame, AL. and Epstein, LB. (1997) *Electrophoresis* **18:** 409-417

Mazur, P., Pieken, WA., udihas, SR., Williams SE., Wong, S. and Kozarich JW. (1994) *Biochemisty* **33(7):** 1961-70

Milstein, OA., vered, Y., Sharma, A., Gressel, J. and Flowers, HM. (1983) *Appl. Environ. Microbiol.* **4:** 55-61

Milstein, OA., Trojanowski, J., Huttermann, A. and Gressel, J. (1988) *Microbios.* **55(222):** 7-16

Molloy, MP. (2000) *Anal. Biochem.* **280(1):** 1-10

Mörtberg, M. and Neujahr, HY. (1985) *J. Bacteriol* **161(2):** 615-9

Mörtberg, M. and Neujahr, HY. (1986) *J. Bacteriol* **168(2):** 734-8

Mortz, E., Krogh, T.N., Vorum, H., and Görg, A. (2001) *Proteomics* **1:** 1359-1363
Neidhardt, FC., Appleby DB., Sankar, P., Hutton, ME. and Phillips, TA. (1989) *Electrophoresis* **10(2):** 116-22

Musante, L., Candiano, G. and Ghiggeri, GM. (1998) *J. Chromatogr. B.* **705(2):** 351-6

Navarre, C., Degand, H., Bennett, KL., Crawford, JS., Mørtz, E. and Boutry, M. (2002) *Proteomics* **2:** 1706-1714

Neujahr, HY. and Varga, JM. (1970) *Eur. J. Biochem.* **13(1):** 37-44

Neujahr, HY. and Gaal, A. (1973) *Eur J Biochem.* **35(2):** 386-400.

Neujahr, HY. (1990) *Bioprocess Technol.* **5:** 321-348

Nordlund I, Powlowski J, Shingler V. (1990) *J Bacteriol.* **172(12):** 6826-33.

O'Farrell, PH. (1975) *J. Biol. Chem.* **250:** 4007-4021

Paca, J., Komarkova, E., Prell, A., Stiborova, M. and Sobotka, M. (2002) *Folia Microbiol (Praha)* **47(6):** 701-7

Pappin, DJC., Hojrup, P. and Bleasby, AJ. (1993) *Curr. Biol.* **3:** 327-332

Patterson, SD and Aebersold, RH. (2003) *Nat. Genet.* **33(suppl):** 311-23

Patton, WF. (2000) *Electrophoresis* **21(6):** 1123-44

Pedersen, SK., Harry, JL., Sebastian, L., Baker, J., Traini, MD., McCarthy, J., Manoharan, A., Wilkins, MR., Gooley, AA., Righetti, PG., Packer NH., Williams, KL. and Herbert, BR. (2003) *J. Proteome Res.* **2(3):** 303-311

Perkins, DN., Pappin, DJ., Creasy, DM. and Cottrell, JS. (1999) *Electrophoresis* **20:** 3551-3567

Pile, LA., Spellman, PT., Katzenberger, RJ. and Wassarman, DA. (2003) *J. Biol. Chem.* **278:** 37840-37848

Pitarch, A., Sanchez, M., Nombela, C. and Gil, C. (2003a) *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **787(1):** 101-128

Pitarch, A., Sanchez, M., Nombela, C. and Gil, C. (2003b) *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **787(1):** 129-148

Powlowski, J., Ingebrand, J. and Dagley, S. (1985) *J.Bacteriol.* **163:** 1136-1141

Powlowski, J. and Dagley, S. (1985) *J.Bacteriol.* **163:** 1126-1135

Powlowski, J. and Shingler, V. (1990) *J. Bacteriol.* **172(12)**: 6824-40

Powlowski, J., Sahlman, L. and Shingler, V. (1993) *J.Bacteriol.* **175(2):** 377-85

Powlowski, J. and Shingler, V. (1994) *Biodegradation* **5(3-4):** 219-36

Quadroni, M. and James, P. (1999) *Electrophoresis* **20:** 664-677

Rabilloud, T. (1996) *Electrophoresis* **17:** 813-829

Rabilloud, T., Adessi, C., Giraudel, A.and Lunardi, J. (1997) *Electrophoresis* **18:** 307-316

Reardon KF, Kim KH. (2002) *Electrophoresis* **23(14):** 2233-2241

Reiser, J. (1992) *J. Bacterial* **174:** 7112-7120

Reiser, J., Glumoff, V., Kälin, M., Ochsner, U.A., and Fiechter, A. (1994) *J Bacteriol.* **176(10):** 3021-32

Reiser, J., Ochsner, UA., Kälin, M., Glumoff, V. and Fiechter A. (1996) Non conventional Yeast in Biotechnology. A Handbook. Klaus Wolf editor, Springer.

Rigaut , G., Shevechenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) *Nature Biotechnol.* **17:** 1030-2

Righetti, PG., Wenisch, E. and Faupel, M. (1989) *J. Chromatogr.* **475:** 293-309

Righetti, PG., Wenisch, E., Jungbauer, A., Katinger, H. and Faupel, M. (1990) *J. Chromatogr.* **500:** 681-696

Righetti, PG., Castagna, A.and Herbert, B. (2001) *Anal. Chem.* **73:** 320A-326A
Santha, R., Savithri, HS., Rao, NA., and Vaidyanathan, CS. (1995) *Eur. J. Biochem.* **230(1):**104-10

Santoni, V., Molly, MP. and Rabilloud, T. (2000) *Electrophoresis* **21:** 1054-1070

Shemyakin, MM., Ovchinnikov, YA., Kiryushkin, AA., Vinogradova, EI., Miroshnikov, AI., Alakhov, YB., Lipkin, VM., Shvetsov, YB., Wulfson, NS., Rosinov, BV., Bochkarev, VN. and Burikov, VM. (1966) *Nature* **211:** 361-366

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) *Anal. Chem.* **68:** 850-858

Shevchenko, A., Sunyaev, S., Lobota, A.and Bork, P. (2000) *Anal. Chem.* **73:** 1917-1926

Shevchenko, A., Sunyaev, S., Liska, A. and Bork, P. (2002) *Methods. Mol. Biol.* **211:** 221-234

Shingler V, Bartilson M, Moore T. (1993) *J Bacteriol.* **175(6):** 1596-604.

Shoda, M. and Udaka, S. (1980) *Appl. Environ. Microbiol.* **39(6)**: 1129-33

Sez, IS. and Dagley, S. (1984) *J. Bacteriol.* **59(1):** 353-9

Taylor, JA. and Johnson, RS. (2001) *Anal. Chem.* **73(11):** 2594-604

Taylor, MG. and Massey, V. (1990) *J. Biol. Chem.* **265 (23):** 13687-13694

Taylor, SW., Fahy, E. and Ghosh, SS. (2003) *Trends Biotechnol.* **21(2):** 82-8

Thatcher, DR. and Cain RB. (1970) *Biochem J.* **120(4):** 28P-29P

Thatcher, DR. and Cain RB. (1974a) *Eur. J. Biochem.* **48(2):** 549-56

Thatcher, DR. and Cain RB. (1974b) *Eur. J. Biochem.* **48(2):** 557-62

Thatcher, DR. and Cain RB. (1975) *Eur. J. Biochem.* **56(1):** 193-204

Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I. and Davison, M. (2001) *Proteomics* **1:** 377-396

van der Heide, M., van der Klei, IJ., Hollenberg, CP. and Veenhuis, M. (2000)
Varga, JM. and Neujahr, HY. (1970) *Eur. J. Biochem.* **12(3):** 427-34

Varga, JM. and Neujahr, HY. (1972) *Acta. Chem. Scand.* **26(2)**: 509-12

Voss, T. and Haberl, P. (2000) *Electrophoresis* **21:** 3345-3350

Wall, DB., Kachman, MT, Gong, S., Hinderer, R., Parus, S., Misek, DE., Hanash, SM.and Lubman, DM (2000) *Anal. Chem.* **72:** 1099-1111

Washburn, MP., Wolters, D. and Yates, JR 3rd. (2001) *Nat. Biotechnol.* **19(3):** 242-7
Waters, S. and Neujahr, HY. (1994) *Prot. Express Purif.* **5:** 534-540

Welch, GR. And Gaertner, FH. (1975) *Proc. Natl. Acad. Sci. USA.* **72(11):** 4218-22

West, M., Emerson, GW. And Sullivan, PA (1990) *J. Gen. Microbiol.* **136(pt8):** 1483-1490

Wilkins, MR., Sanchez, JC., Gooley, AA., Appel, RD., Humphery-Smith, I., Hochstrasser, DF. and Williams, KL. (1995) *Biotechnol. Gene. Eng. Rev.* **13:** 19-50.

Wilkins, MR. and Williams, KM. (1997) *J. Theor. Biol.* **186(1):** 7-15

Wilm, M. and Mann, M. (1996a) *Anal. Chem.* **68:** 1-8

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T. and Mann, M. (1996b) *Nature* **379:** 466-469.

Wu, CC., MacCoss, MJ., Howell, KE. and Yates III, JR. (2003) *Nat. Biotech.* **21:** 532-538

Yan, J.X., Wait, R., Berkelman, T., Harry, RA., Westbrook, JA., Wheeler, CH. And Dunn, MJ. (2000) *Electrophoresis* **21:** 3666-72

Yates, JR. III, Speicher, S., Griffin, PR. and Hunkapiller, T. (1993) *Anal. Biochem.* **214:** 397-408

Zhang, L., Takaya, N., Kitazume, T., Kondo, T. and Shoun, H. (2001) *Eur. J. Biochem.* **268(11):** 3198-204