Protein expression profiling of *Escherichia coli* wild type versus a Δ*ptsI* mutant and improvement of in-gel protein visualisation

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Andreas Lamanda

von Oberburg

Leiter der Arbeit:
Prof. Dr. Bernhard Erni
Departement für Chemie und Biochemie

Protein expression profiling of *Escherichia coli* wild type versus a Δ*ptsI* mutant and improvement of in-gel protein visualisation

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von
Andreas Lamanda
von Oberburg

Leiter der Arbeit:
Prof. Dr. Bernhard Erni
Departement für Chemie und Biochemie

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Der Dekan:

Bern, den 28. April 2004
Prof. Dr. G. Jäger

The proteomes of an *E. coli* K12 wildtype and an isogenic *ptsI* mutant were compared by protein expression profiling. Deletion of enzyme I of the phosphoenolpyruvate dependent phosphotransferase system did not cause a dramatic change in the expression profile. The deletion of enzyme I affected: (i) genes under catabolite repression (ii) glycolytic and tricarboxylic acid cycle enzymes, (iii) anaerobiosis-responsive enzymes, (iv) proteins of the PTS, (v) “general” stress proteins, (vi) proteins related to oxidative stress, (vii) proteins related to information transfer. Parallel to protein expression profiling a new fluorescent staining and destaining procedure for proteins in 1 and 2-D polyacrylamide gel matrices was developed. The new procedure was published in *Proteomics* 2004, *4,* 599–608 and was subject to US provisional patent application.

Content

Part I

Protein expression profiling of Escherichia coli wild type versus a Δ*ptsI* mutant

1. Introduction

1.1. The phosphotransferase system
1.1.1. Proteome analysis of *E.coli*
1.2. Proteomics and protein separation techniques
1.2.1. Proteomics
1.2.2. The Proteome
1.2.3. Protein expression profiling
1.2.4. Two dimensional gel electrophoresis
1.2.5. Chromatographic approaches
1.2.6. Protein identification
1.2.7. Methodical setup

2. Materials and Methods

2.1. Bacterial strains
2.2. Sample preparation, cell growth and harvesting
2.2.1. Cell growth and harvesting of MC4100HIC9
2.2.2. Sample preparation of cytoplasmic and membrane associated proteins
2.2.3. 2-D gel electrophoresis
2.2.4. Staining procedures
2.2.5. Optimisation of time-depending silver staining
2.2.6. Imaging
2.2.7. Gel conservation
2.2.8. Protein identification I
2.2.8.1. Edman degradation
2.2.8.2. MALDI-TOF MS
2.2.8.3. Nanospray ESI-TOF MS
2.2.8.4. Matching to SWISS-2DPAGE reference gel
2.2.9. Protein identification II
2.2.10. Western blotting
2.2.11. Image analysis and data processing
2.2.12. Statistical analysis
2.2.13. Selection criteria for protein identification
2.2.14. Data mining
2.2.15. Protein name and accession number
2.2.16. Regulation factor and map location
2.2.17. Calculation of theoretical and observed molecular weight M*r* and p*I*
2.2.18. Operon search

2.2.19. Regulon search
2.2.20. Pathway search
2.3. Molecular biology
2.3.1. Reporter plasmid construction
2.3.2. Isolation of genomic DNA
2.3.3. Isolation of plasmid DNA
2.3.4. β-galactosidase assays

3. Results and Discussion

3.1. Media selection for growth of *E. coli* MC4100 and MC4100Δ*ptsI*.
3.2. Growth of *E. coli* MC4100, MC4100Δpts and MC4100(pTSHIC9) and protein sample preparation
3.3. Quality control of protein sample and optimisation of silver staining procedure
3.4. Protein expression profiling
3.5. Construction of the expression profile
3.6. Statistical analysis
3.7. Protein identification
3.7.1. Protein identification with nanospray ESI-TOF
3.7.2. Protein identification by MALDI-TOF
3.8. Masses and isoelectric points
3.9. Validation of the protein identification
3.10. Validation of the regulation factors inferred from proteome analyses
3.11. Enzyme I related protein expression
3.11.1. Controlled expression of genes under catabolite repression
3.11.2. Controlled expression of genes which are not under catabolite repression
3.11.3. Constant expression of genes expected to be under catabolite repression
3.11.4. The effect of EI on glycolytic and tricarboxylic acid cycle enzymes
3.11.5. The effect of EI on anaerobiosis-responsive enzymes
3.11.6. The effect of EI on proteins of the PTS
3.11.7. The effect of EI on stress proteins
3.11.7.1. The effect of EI “general” stress proteins
3.11.7.2. The effect of EI proteins related to oxidative stress
3.11.8. The effect of EI proteins related to information transfer

4. Reference list

Part II

Improvement of in-gel protein visualisation

5. *Proteomics* 2004, 4, 599–608

Improved Ruthenium II tris (bathophenantroline disulfonate) staining and destaining protocol for a better signal-to-background ratio and improved baseline resolution

Part III

6. Appendices

6.1 Cover of US provisional patent application 60/454,979
6.2 Cover of *Proteomics* 4, 2004
6.3 Acknowledgements
6.4 Curriculum vitae

Part I

Protein expression profiling of *Escherichia coli* wild type versus a ∆*ptsI* mutant

1. Introduction

1.1. The Phosphotransferase System of E.coli.

Escherichia coli belong to the Enterobacteriaceae that can live almost everywhere in nature. *E. coli* is able to survive in a continuously changing surrounding and has evolved mechanisms to allow it to cope with conditions as diverse as the gut and the sewer. If a nutrient is exhausted the gram-negative bacteria can chemotactically move or adapt its metabolism to the use of other nutrients. To achieve this, *E.coli* has several sensory systems to monitor its surrounding [1]. *Escherichia coli* can utilize many different carbon sources whose metabolism can be turned on and off. This process is controlled by signals that are sensed on the outside of the cell, which are transformed into an intracellular response such as a change in protein expression or a change in enzyme activity.

The phosphoenolpyruvate (PEP):carbohydrate phosphotransferase system (PTS) is one of these sensory systems. It is widely distributed in bacteria but it does not occur in archebacteria, animals and plants [2]. The PTS of *E. coli* which is one of the larger PTS consists of 56 proteins. 28 are soluble proteins, 28 are membrane spanning carbohydrate transport systems (Figure 1).

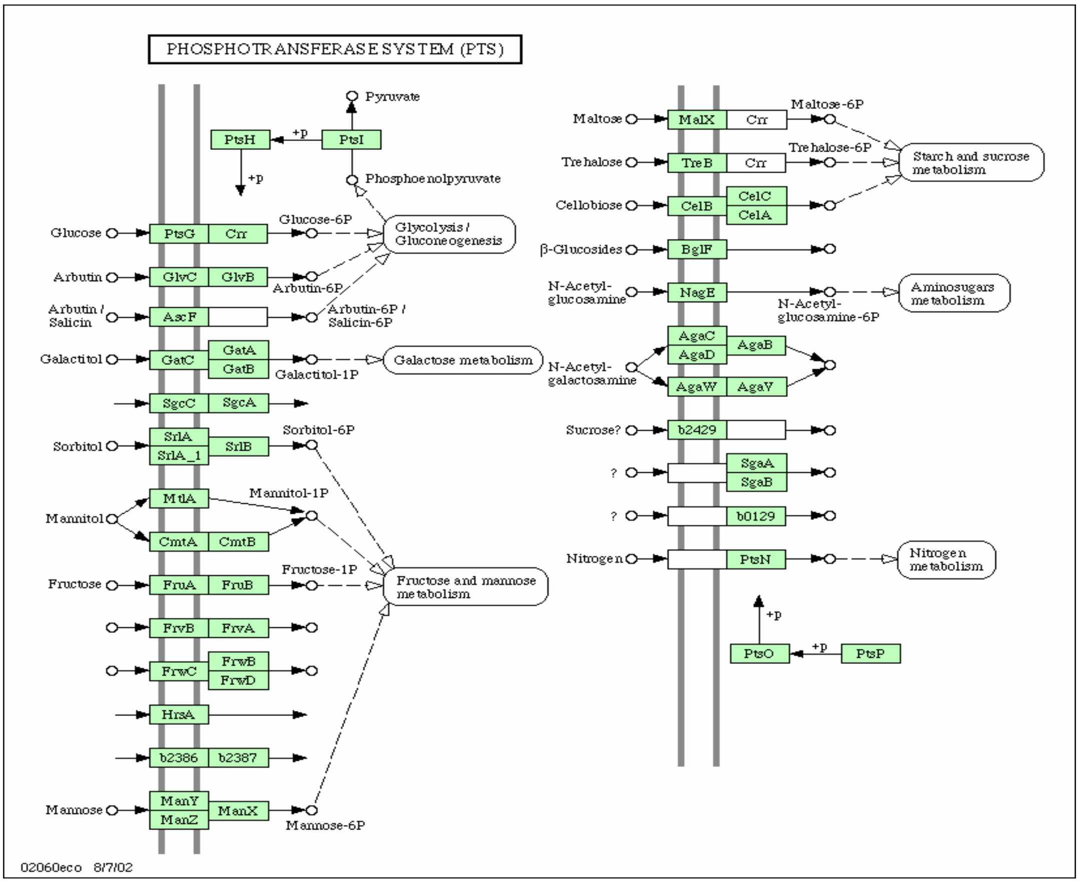
PHOSPHOTRANSFERASE SYSTEM (PTS)
Pyruvate
PtsH
+p
PtsI
+p
Phosphoenolpyruvate
Glucose
PtsG
Crr
Glucose-6P
Glycolysis / Gluconeogenesis
Arbutin
GlvC
GlvB
Arbutin-6P
Arbutin / Salicin
AscF
Arbutin-6P / Salicin-6P
Galactitol
GatC
GatA
GatB
Galactitol-1P
Galactose metabolism
SgcC
SgcA
Sorbitol
SrlA
SrlA_1
SrlB
Sorbitol-6P
Mannitol
MtlA
Mannitol-1P
CmtA
CmtB
Fructose
FruA
FruB
Fructose-1P
Fructose and mannose metabolism
FrvB
FrvA
FrwC
FrwB
FrwD
HrsA
b2386
b2387
Mannose
ManY
ManZ
ManX
Mannose-6P
Maltose
MalX
Crr
Maltose-6P
Trehalose
TreB
Crr
Trehalose-6P
Starch and sucrose metabolism
Cellobiose
CelB
CelC
CelA
β-Glucosides
BglF
N-Acetyl-glucosamine
NagE
N-Acetyl-glucosamine-6P
Aminosugars metabolism
N-Acetyl-galactosamine
AgaC
AgaD
AgaB
AgaW
AgaV
Sucrose?
b2429
?
SgaA
SgaB
?
b0129
Nitrogen
PtsN
Nitrogen metabolism
+p
PtsO
+p
PtsP
02060eco 8/7/02

Figure 1. Phosphotransferase system of *E. coli*. The figure is from www.Ecocyc.org

The PTS components can be organized in a matrix (Figure 2). The columns designate the biological functions that PTS-components can have the rows refer to position of a particular component along the phosphorylation cascade. The functions of these modules are carbohydrate transport and phosphorylation, antitermination of mRNA transcription, catabolite repression, inducer exclusion, and chemotactic signalling.

	Carbohydrate transport and phosphorylation	m-RNA antitermination	Catabolite repression	Inducer exclusion	Chemotaxis
Initiation of phosphoryl transport	**Enzyme I**	**Enzyme I**	**Enzyme I**	**Enzyme I**	**Enzyme I**
Joint	**Hpr**	**Hpr**	**Hpr**	**Hpr**	
Targetting subunits	**14 IIA domains**	**IIA Bgl**	**IIA Glc**	**IIAGlc**	**?**
Target loci	**14 IIBC IIBC Bgl IIBC Glc**	**BglG**	**Cya**	**Non PTS systems**	**Che protein switch**

Figure 2. Matrix like design of the PTS. The columns representing functional modules, the rows contain their protein constituents.

Enzyme I (EI) and the heat stable protein (HPr) are two protein constituents of the PTS. They are located in the central middle positions of the top two rows in the matrix which indicates their involvement in all functional PTS modules. *In vivo* they are located in the cytoplasm. In each module of the PTS phosphoryl transfer occurs in a cascade of reactions from enzyme I to an acceptor (carbohydrate) via Hpr and a class of protein kinases named enzymes IIA and carbohydrate specific transporters, enzymes IIB and IIC [1]. Alternatively the phosphoryl flow is directed away from the PTS to non PTS systems at the Hpr level.

Enzyme I initiates phosphoryl transfer from phosphoenolpyruvate (PEP) to HPr or alternatively to CheA a protein of the flagellar apparatus. Autophosphorylation occurs at the N3 position of a histidyl residue of enzyme I from which the phosphate is transferred to a histidyl residue of HPr. Hpr which is kind of a universal joint for the different modules of the PTS is phosphorylated on histidine 15, (position N1) and transfers the phosphorylgroup to a histidyl residue (N3) in any of the set of fourteen different enzyme IIA proteins/domains.

Each carbon source has its representative on the enzyme II level (row three, Figure 2). This level of the PTS has several functions. One of these functions is to transfer the phosphorylgroup to the sugar specific enzymes IIB and IIC of which IIC forms a channel for substrate transport across the membrane while IIB phosphorylates the transported substrate. This mechanism of transport coupled to chemical modification of the substrate is known as group translocation.

Another function at the enzyme II level is realized by the glucose specific enzyme IIA (IIA^{Glc}). Depending on the availability of glucose, the ratio of phospho-IIA^{Glc} to dephospho-IIA^{Glc} varies. The phosphorylation state of IIA^{Glc} controls the uptake of alternative carbon sources such as lactose, galactose or ribose [3]. In presence of glucose and other PTS substrates the steady state concentration of dephospho-IIA^{Glc} is increased and the one of P-IIA^{Glc} decreased. Dephospho-IIA^{Glc} acts as allosteric inhibitor of lactose permease (LacY, *lac* operon) and the ATP-binding protein subunit (MalK, *mal* operon) of the maltose permease and inhibits the uptake of lactose and maltose, which are inducers for LacY and MalK. This mechanism is known as *inducer exclusion* [3]. Thereby operons like the *lac, gal* or *mgl operons* which are under additional control of a second gene specific repressor, are only expressed at a basal level. Firstly because the inducer is not present in the cell and secondly because the catabolite activator protein remains inactive as long as the cAMP concentration is low. Finally inducer exclusion depends not only on the phosphorylation status of IIA^{Glc} but also on the concentration ratio of IIA^{Glc} and target molecules, the presence of substrate of the target molecules and the intracellular cAMP level [4].

When PTS substrates are present, the expression of proteins from other metabolic pathways is repressed. The balance is then on the side of phospho-IIA^{Glc} which does not stimulate adenylyl cyclase. Therefore the concentrations of cAMP and catabolite activating protein (CAP) are low. The cAMP-CAP complex, which is formed when higher cAMP levels are reached upon stimulation of adenylyl cyclase through phospho-IIA^{Glc}, represses the transcription of about 100 different operons. Glucose can be regarded as “catabolite” of lactose or maltose that is produced when the disaccharides are metabolized. Therefore, as long as the catabolite is present in amounts that can be metabolized, the genes that encode the machinery for lactose and maltose transport are repressed. This gave the name to the phenomenon: *catabolite repression* [3].

A third function on the enzyme II level is realized through EII^{Bgl}. EII^{Bgl} controls its own expression through phosphorylation of the regulatory protein BglG. In the absence of beta-glucosides, BglG is phosphorylated by EII^{Bgl} and is inactive in mRNA antitermination. Addition of inducer stimulates EII^{Bgl} to dephosphorylate BglG, allowing BglG to function as a positive regulator of operon expression.

In the bottom leftmost field of the PTS matrix, the carbohydrate specific transporters enzymes IIB (cytoplasmatic part) and IIC (trans-membrane part) are located. Enzymes IIB are phosphorylated by their specific IIA subunits on a cysteinyl residue. From there the phosporylgroup is transferred to the acceptor (carbohydrate) which is transported simultaneously across the inner membrane.

The PTS of *E. coli* has extensively been studied for specific carbohydrate transport, gene expression and its regulation during the last years. Table 1 shows a summary of operons encoding transport proteins for the different carbohydrate substrates.

Table 1. PTS transporters of *E. coli* and their substrates

transported substrate	**operon/gene**	**reference**
5-carbon sugars	*SgcABC*	[7, 8]
arbutin, salicin, cellobiose	*AscF*	[9]
fructose	*FrwCBD*	[10,11]
aromatic β-glucoside (Arbutin/Salicin)	*BeglFGB*	[2,5,12,13]
galacitol	*GatABC*	[14]
glucitol	*GutABE*	[15,16]
glucose	*PtsG/Crr*	[19-20]
glucose (presumably)	*MalX*	[21]
mannitol	*MtlA*	[22-25]
mannitol (cryptic)	*CmtAB*	[26]
mannose	*ManXYZ*	[27-31]
N,N'-diacetylchitobiose	*ChbA-ChbB-ChbC*	[32-36]
N-acetylgalactosamine (putative)	*AgaBCDVWX*	[37]
N-acetylglucosamine	*NagE, NagBACD*	[38-41]
trehalose	*TreB*	[42-45]
unknown specifity	*FrvAB*	[46-50]
unknown specifity	*GlvCB*	[51]
unknown specifity	*Frx*	[2]
unknown specifity	*ptsA*	[52]

The PTS is also involved in the metabolism so that it forms a cycle together with the glycolysis. For each sugar molecule that is transported by the PTS two PEP molecules are generated by glycolysis. One of these PEP molecules is used for the transport of the next sugar molecule [3]. Other pathways that are directly linked to the PTS are: the galactose metabolism, the fructose and mannose metabolism, starch and sucrose metabolism, amino-sugar metabolism and nitrogen metabolism.

In 1992 Reizer and Saier speculated about a PTS-catalyzed protein phosphorylation that would provide a regulatory link between carbon and nitrogen assimilation [5, 6] but so far there is no experimental evidence for this theory.

1.1.1. Proteome analysis of E. coli. In recent years, it has become clear that, in addition to the regulation of the expression of specific genes or operons, there are global regulatory systems that control the simultaneous expression of a large number of genes in response to a variety of environmental stimuli. The proteomic approach has become very popular to detect these changes. Several studies concerning the *E. coli* proteome can be found in the literature (Table 2).

Table 2. Twin proteome studies of *E. coli* under different conditions.

Proteome analysis of E.*coli*	**Number of detected protein spots**	**expr. changed**	**up reg.**	**down reg.**	**Ref.**
Acetate and formate stress	800	84	37	17	[100]
Global metabolic regulation in response of different carbon sources	1000	52	-	-	[98]
Comparison of parent vs. L-Threonin overproducing strain	800	54	19	35	[103]
pH-dependant expression of periplasmic proteins	800	45	23	22	[99]
Proteome analysis of metabolic engineered *E.coli* strain that overproduces PHB	not known	13	-	-	[104]
Proteomic pattern of gel entrapped vs. free floating *E.coli*	1000	94	35	59	[105]
Acid and base induced proteins	300	18	-	-	[106]

Knowing the facts about the involvement of the PTS in regulatory functions we aimed for the following targets by applying proteomics:

1. Displaying the proteomes of an *E. coli* wild type strain (MC4100) and strain MC4100Δ*ptsI* lacking the *ptsI* gene and therefore enzyme I on 2-D electropherograms.
2. Comparison of the spot patterns, localisation and quantification of proteins with altered expression (protein expression profiling).
3. Identification of proteins with altered expression, especially such with unknown function.
4. Application of an orthogonal method to evaluate the results from 2-D Page.
5. Comparison and if possible integration of the resulting data into the existing model or suggestion of new regulatory functions of enzyme I.

1.2. Proteomics and protein separation techniques

1.2.1. Proteomics. The concept of “Proteomics” was proposed in September 1994 at the second Siena 2-D electrophoresis meeting to define protein-based gene expression analysis [60]. Ten years later, proteomics is defined as the search for quantitative changes of expression levels by simultaneous analysis of complex protein mixtures like cell lysates or tissue extracts. [61]. The term proteomics or proteome analysis usually also includes the serial coupling of three steps: sample preparation, protein separation and protein identification.

1.2.2. The proteome. Originally, the *proteome* was defined as the proteins expressed by a genome under specified conditions [62].The proteome may contain hundreds up to several thousand protein species depending on the origin of the cell type. Not all of these proteins are expressed at the same time and with the same abundance. Changes in protein expression occur continuously during the cell cycle. For this reason the proteome is not a fixed feature in an organism compared to its genome. Instead, it is a dynamic entity, changing upon development of the cell, the environmental conditions and mutations. As consequence proteomes are much more complex than their corresponding genomes. Several other phenomena additionally contribute to the complexity of the proteome. For example: gene splicing leading to different gene products, post-translational modifications such as proteolytic cleavage, phosphorylation or glycosylation. Due to this fact it is not possible to display a final or “the proteome map” of a cell, tissue or organism. To solve this problem an approach named *protein expression profiling* can be applied.

1.2.3. Protein expression profiling. This approach makes it possible to compare the proteomes of two or more “discrete states” of a cell. The expression level of each protein of the proteome is quantified at a defined time point of cell growth in a massive parallel approach. The resulting time independent expression profile is then statistically compared to its “Twin” which represents the counterpart derived under controlled altered conditions. Such a comparative approach on unchanged and altered proteome is therefore also called “Twin proteomic approach” [63]. A scheme of the approached used in this study is shown in Figure 12.

During the last few years a vast array of separation techniques appeared for performing quantitative proteome analysis. These methods can be divided into two major classes: conventional two-dimensional map analysis which couples orthogonally a charge-based step (isoelectric focusing) to a size-based separation (SDS-electrophoresis), and two-dimensional procedures [64].

1.2.4. Two dimensional gel electrophoresis. If 2-D electrophoresis is applied for proteome analysis, two stages prior to protein identification are added. In-gel visualisation of proteins that is followed by image and statistical analysis. A detailed discussion of traditional and new protein in-gel visualisation methods is given in sections B and C.

The roots for the first dimension of 2-D electrophoresis can be found in the year 1912 when the first isoelectric focussing fractionating was described by Ikeda and Suzuki [65]. The ancestor of SDS Page, which is used as second dimension, was described by Tiselius in 1937 [66]. The first 2-D electrophoretic separation according to completely independent physico-chemical parameters (isoelectric point and size) of proteins was done in 1970 by Kenrick and Margolis [67]. They combined native IEF with pore gradient SDS gel electrophoresis. O'Farrell, Klose and Scheele replaced the pore gradient SDS gel electrophoresis by the Lämmli [68] system in 1975 [69-71]. In 1982 Bjellqvist replaced the denaturing IEF in the first dimension by immobilised pH gradient (IPG) gels [72] and Görg brought the technique in 1988 close to the recent standard [73].

2-D electrophoresis is by far the most popular method in proteomics nowadays. Applying 2-D-PAGE enables for differential display of paired samples and statistical analysis performed on sets of gels via powerful software packages, such as MELANIE, Carol, PDQuest, Z3, 2-D Advance, Image Master, Progenesis and Phoretix evolution. Due to their heterogeneity and significant differences in abundance, the detection and identification of all proteins expressed in cells and tissues is a major challenge in proteome analysis [74]. 2-D electrophoresis in combination with mass spectrometry is the actual working horse in proteome analysis. During the last years, 2-D electrophoresis with immobilized pH gradients (IPGs) has constantly been refined [75]. The development of IPGs between pH 2.5 and pH 12 has enhanced the resolution of the proteome and, in particular, facilitated the analysis of very acidic and very alkaline proteins, whereas the introduction of overlapping narrow-range IPGs permits higher resolution [75].

Today 2-D electrophoretic separation can resolve as much as 1500 proteins per gel. [76]. No other method is capable to provide this resolving power on a reproducible basis. In the ideal case 95% of these proteins can be identified by mass spectrometric methods [77, 78]. But still there remain limitations to the current technology to separate the broad diversity of proteins present in the cell. Dynamic ranges for some protein species, for example albumine vs. cytokines in human blood plasma, are undoubtly outside the range of 2-D electrophoresis [79]. Hydrophobic proteins are usually underrepresented, since these are poorly soluble in the standard sample buffer for 2-D electrophoresis.

There are several different approaches to reach higher resolution and to detect low abundant and hydrophobic proteins. Hartinger, MacFarlane and also Langen and Ueffing aimed for separation of hydrophobic proteins by combining a discontinuous electrophoresis system that uses the cationic detergent 16- BAC in the first dimension with conventional discontinuous SDS-PAGE [81-84]. Although this was not a real two dimensional separating setup, they were able to separate protein off diagonal, due to the different binding properties of the proteins to the detergents 16-BAC (cationic) and SDS (anionic).

A technically challenging attempt to increase the resolving power of 2-D electrophoresis was undertaken by Inagaki in 2002 by increasing the gel size. Fourteen large 2-DE gels (twelve 24 cm x 70 cm gels and two 18 cm x 70 cm gels) were assembled into a 93 cm x 103 cm virtual gel. Data derived from this experiment suggested that the large virtual gel can display more than 11,000 protein spots expressed in a 1-100,000 dynamic range in cells [85].

In 2003 C. Eckerskorn reported the expansion of the two dimensional separation by adding a third electrophoretic dimension, a free flow electrophoresis step (FFE), prior to 2-DE [86, 87].

Another recent attempt to add a third separating dimension to 2-D electrophoresis was reported by Lee in 2003. The Method is called 3-D SDS-PACGE (sodium dodecyl sulphate polyacrylamide cube gel electrophoresis). IEF is combined with a 12% SDS-PAGE with 2-[N-morpholineo]ethanesulfonic acid (Mes) running buffer that separates low molecular weight proteins. The third dimension is a 7.5% SDS-PAGE cube gel with tris-glycine running buffer to separate the high molecular weight proteins [88].

1.2.5. Chromatographic approaches. When column chromatographic approaches are used for proteome analysis visualisation, image and statistical analysis, three complex steps, are eliminated. Chromatographic approaches mostly rely on analysis of tryptic digests. More often one sample is parallel digested with trypsin, Glu-C and subtilisin to create overlapping peptides. Micropore HPLC or other devices used for peptide separation are coupled directly to the mass spectrometer for continuous peptide identification.

Relative Differences between two samples are measured by labelling techniques. The most prominent technologies are the ICAT, MCAT, and MudPIT. In the so called isotope coded affinity tag (ICAT) approach cysteine residues are labelled with the ICAT reagent which comprises a thiolspecific reactive group, a linker (heavy with deuterium or light with hydrogen) and biotin. Relative protein levels can be measured by MALDI-TOF. One drawback is that protein without cysteines are not accessible with this method. Another drawback is the huge amount of data that is generated during continuous MS-analysis [89].

Mass coded abundance tagging (MCAT) is based on differential guanidination of lysine residues with *O*-methylisourea. From full-scan MS spectra the relative abundance of sister peptide species can be determined. [90].

In the multidimensional protein identification technology (MudPIT) introduced by Link in 1999, peptides are separated with a biphasic micro capillary column packed with strong cation exchange material followed by a reversed phase material. Peptides are directly eluted into a tandem mass spectrometer [91]. This approach is also called shot gun proteomics [92].

1.2.6. Protein identification. The final step in proteome analysis is protein identification. During the last ten years amino-terminal sequencing, which requires blotting on a membrane, has been replaced by tryptic in-gel digest and mass spectrometric analysis of the resulting peptides. For Edmann degradation [93], protein amounts in the range hundreds of nano mols were necessary. This amount has drastically dropped to a few pico mols which are necessary for mass spectrometry today [94].

Matrix assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF MS [95]) in combination with peptide mass fingerprinting (PMF) [96] is the common choice of method for protein identification. Combined with an algorithm for reliable evaluation of mass spectrometric data and with automatic data base search, MALDI-TOF MS is a fast and reliable method [97-99].

1.2.7. Methodical setup. During the methodical set up of the two dimensional gel electrophoresis system different staining procedures for polyacrylamide gels were evaluated. This led finally to the development of a new staining and destaining procedure which was published in March 2004 in "*Proteomics*" (see Part II). The method in Part II was commercialised by the University of Berne together with Fluka AG and brought to US provisional patent application. The application recipe is shown in part III.

2. Materials and Methods

2.1. Bacterial strains.

E.coli strain MC4100 is F-araD139 Δ(araF-lac)U169 rpsL150 relA1 thi fib5301 deoC1 ptsF25 rbsR [100] and MC4100Δ*ptsI* is MC4100ptsI::Ω [101]. MC4100(pTSHIC9) carries a plasmid pTSHIC9 encoding HPr, EI and IIA^{Glc}.

2.2. Sample preparation, Cell growth and harvesting.

Overnight starter cultures were inoculated with a single colony from an LB plate and grown overnight at 37° on an orbital shaker. To adapt *MC4100, MC4100Δpts* to amino acid medium (AA-medium) 500 µl of the overnight culture were transferred into 50 ml of AA-medium and again grown overnight. On the following day this starter cultures was diluted 1:100 into 800 ml AA-medium in a 5 l Erlenmeyer flask. AA-medium contains per 1 litre (0.5 g alanine, 0.48 g arginine, 0.4 g aspartic acid, 0.037 g cysteine, 0.33 g glutamine, 0.54 g glycine, 0.06 g histidine, 0.23 g isoleuzine, 0.42 g lysine, 0.09 g methionine, 0.13 g phenylalanine, 0.1 g proline, 2.1 g serine, 0.23 g threonine, 0.17 g tyrosine, 0.23 g valine, 100 ml of a sterile filtered (0.2 µm) solution of (7 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl), 20 mg biotin (sterile filtered 0.2 µm), 20 mg thiamine (sterile filtered 0.2 µm), 0.12 g $MgSO_4$, 0.012 g $CaCl_2$). Batch cultures were grown in AA-medium under vigorous shaking to OD$_{600}$=0.8 and harvested by centrifugation at 2600 x g (Sorvall GSA, 4000 rpm) at 4°C. The resulting pellet was washed four times by re-suspension in 20 ml of ice-cold low salt washing buffer (3 mM KCl, 1.5 mM KH_2PO_4, 68 mM NaCl, 9 mM NaH_2PO_4), and stored frozen at -20°C.

Cell growth and harvesting of MC4100HIC9. A fresh overnight culture of *E. coli MC4100HIC9* was diluted 1:100 in 800 ml of AA-medium. Batch cultures containing 50 µg/ml ampicilline were grown under vigorous shaking. At a density of OD$_{600}$=0.15, 0.5 mM IPTG was added. The culture was grown until OD$_{600}$=1.0 and processed as described above.

2.2.1. Sample preparation of cytoplasmic and membrane associated proteins. On the next day the cells were thawed on ice, re-suspended in 600 µl of 8 M urea and broken by five times sonification for 45 seconds with a Labsonic

1510 sonifier (Bender & Hobein, Zürich). Between the sonification steps the lysate was allowed to cool. The resulting cell lysate was centrifuged for ten minutes at maximum speed in an Eppendorf centrifuge to remove the cell debris. The supernatant was centrifuged for 30 min at 352900 x g at 4°C to remove the bacterial membranes. The resulting supernatant containing urea soluble proteins was split into 1, 3 and 10 μl aliquots and frozen in liquid nitrogen and stored at -80°C.

2.2.3. 2-D gel electrophoresis. (i) *Analytical gels*. For the first dimension IPG strips (pH 3-10, Amersham Pharmacia) were rehydrated for 10 h with 300 μl sample buffer (8 M urea, 2% CHAPS, 18 mM DTE, 0.5% IPG buffer 3-10 and traces of bromophenole blue) containing 80 μg *E. coli* proteins. Isoelectric focussing was carried out in an IPGphor (Amersham Pharmacia Biotech) with the following settings: 20°C, 200 μA per strip, 150 Vh (1 h, 150 V, step-n-hold), 300 Vh (1 h, 300 V, step-n-hold) 17500 Vh (5 h, 3500 V, step-n-hold) 63250Vh (gradient, constant volt-hour-area).

(ii) *Preparative gels*. IPG strips were rehydrated with 300 μl sample buffer containing 800 μg *E. coli* proteins for 10 h on the IPGphor under 30 V. Focussing was carried out with the following settings: 20°C, 200 μA per strip, 150 Vh (1 h, 150 V, step-n-hold), 300 Vh (1 h, 300 V, step-n-hold) 17500 Vh (5 h, 3500 V, step-n-hold) 276000 Vh (gradient), 80000 Vh (10 h, 8000 V, step-n-hold).

Equilibration (for i and ii): The IPG strips were equilibrated for 12 min with 5 ml/strip solution I (50 mM Tris-HCl, pH 7.0, 6 M urea, 30% (v/v) glycerol, 2.5% (w/v) DTE) and 12 min with 5 ml/strip solution II (50 mM Tris-HCl, pH 7.0, 6 M urea, 30% (v/v) glycerol, 2.5% (w/v) iodoacetamide).

Second dimension (for i and ii). Six 12.5% homogeneous polyacrylamide gels (30.8% T, 2.6% C, 15 cm *16 cm *1.5 mm) were prepared in a Hoefer SE 600 Series Multicaster. 83 ml deionized water, 64 ml 1.5 M Tris-HCl, pH 8.8, 105 ml acrylamide/piperazine diacrylamide (PDA) solution (0.8 g PDA (Bio-Rad) per 100 ml 30% (w/v) acrylamide (Rotiphorese GelA, Roth), sterile filtered (0.2 μm)), 0.2 g sodium thiosulfate and 5 ml 10% APS were mixed and polymerisation was started with 0.15 ml TEMED. The freshly poured gels were covered with water saturated butanol, allowed to polymerise for 2 h, taken out of the casting chamber, wetted with water, wrapped in Saran foil, stored at RT for 24 h to allow completion of polymerisation and finally stored at 4°C until use.

1.5 cm of the basic and acidic end from the IPG strip were cut off, the strip was inserted on top of the polyacrylamide gel and sealed with 0.4% agarose in running buffer, containing traces of bromophenol blue on top of the second dimension. The second dimension was run at 2°C for 5 h at 50 mA per gel (500 V, non limiting) in Lämmli buffer (25 mM Tris, 198 mM glycine, 0.1% SDS (w/v)) [68] in a Hoefer Dalt 600 chamber. In cases where proteins were blotted on PVDF membranes for spot identifiaction by Edman sequencing, the Lämmli buffer system was replaced with the Schagger and Jagow buffer system (25 mM Tris, 198 mM tricine, 01% (w/v) SDS).

2.2.4. Staining procedures. Volumes of 200 ml were used for each step of the staining procedure. Analytical gels were stained according to the time-optimized and up-scaled silver nitrate procedure described by Swain and Ross (with an additional step which combined stop and storage (3 g Tris base, 10 ml acetic acid 99% in 490 ml water) [102]. Preparative gels were stained either with Coomassie Blue (saturation staining 48 h) or the colloidal Blue staining kit (Novex 46-7015) for 24 h and destained to the desired spot intensity. SYPRO Ruby was purchased from Bio-Rad (170-3125) and used according to Berggren [103].

2.2.5. Optimisation of time-depending silver staining. Ten samples each of three concentrations of purified IIAB^{Man} (2300 ng, 230 ng, or 23 ng per band) were run on a 17.5 % gel, prepared as described [68] in a Protean 3 Mini cell system (BioRad). During silver staining bands of each concentration were cut out with a scalpel, removed from the Silver nitrate staining solution and immediately transferred into stop solution. The first samples were removed after 160 s and further samples every 60 s until 900 s. The ten samples of each concentration were arranged in a time dependant grey scale ladder and scanned as described in section Imaging. To detect time dependant staining saturation, the profiles of stained protein bands were drawn with the in-built profiling tool of Phoretix 2D Advance V 5.0.

2.2.6. Imaging. Coomassie Blue and Silverstained gels and Western blots were scanned on a flatbed scanner (HP Deskscan, DeskScanII V2.3) with the following scanning parameters: 300*300 dpi, 8 bit black and white picture (256 grey shades, two times sharpened), contrast 125, and brightness 125. SYPRO Ruby stained gels were scanned as described in Materials and Methods section B.

2.2.7. Gel conservation. For permanent storage, gels were placed on a piece of Whatman filter paper, covered with a sheet of saran wrap (Migros Art.7045.444), air bubbles removed, and the sandwich was dried in the gel dryer (Bio-Rad Model 543) at 80°C under vacuum for 2 hours. The dried gels were sealed in laminate at Copy Trend Längassstrasse.

2.2.8. Protein identification I. Proteins were identified by Edman degradation, MALDI-TOF MS and PMF, nanospray ESI-TOF-MS or matching to the SwissProt *E. coli* reference gel from SWISS-2-D PAGE [11].

2.2.8.1. Edman degradation. An unstained 2-D gel run in Shagger Jagow buffer was incubated in blotting buffer (50 mM boric acid, pH 9, 0.1% (w/v) SDS, 20% (v/v) methanol) for 30 minutes, and the PVDF membrane was wetted for 30 minutes in methanol. Proteins were then wet blotted during 90 minutes at 500 mA. The PVDF membrane was rinsed in Coomassie Blue staining solution for 5 s and destained in a mix of 50% methanol, 10% acetic acid overnight. Protein spots were cut out with a scalpel and subjected to Edman degradation by PD Dr. J. Schaller, Department of Chemistry and Biochemistry, Bern.

2.2.8.2. MALDI-TOF MS. Coomassie or Colloidal Blue stained protein spots were cut out of the gel. The gel pieces were destained with 100 mM ammonium bicarbonate in 30% acetonitrile. Proteins were digested with trypsin [97] and peptide masses identified by MALDI TOF as described [98]. The probability of a false positive match of an observed MS-spectrum was determined for each analysis [99]. Analysis was done in the Lab of PD Dr. H. Langen at F. Hoffmann-La Roche, Ltd., Basel.

2.2.8.3. Nanospray ESI-TOF MS. Coomassie blue stained protein spots were cut out of the gel, destained and digested as described above. A ZipTip column was cleaned three times with each 10 µl of a solution of 50% acetonitlile (HPLC grade) 0.1% TFA. The column was activated by three times rinsing with 10 µl 0.1% TFA. 10 µl peptides digest supernatant were pipeted in the ZipTip column. The unbound material was washed out with two times with 0.1% TFA and the peptides were then eluted with 10 µl of 50% acetonitlile 0.1% TFA solution. The eluate was taken up with a nanospray glass tip that was loaded on the Q-Star Pulsar MS and analysed by Johannes Hewel, Department of Chemistry and Biochemistry, Bern. Neutral monoisotopic tryptic peptide masses were transferred to the software Mascot or PepIdent and subjected to peptide mass fingerprint against the Swiss-Prot, NCBI, TrEMBL and newTeEMBL databases. Database search was restricted to *E. coli.* One missed trypsin cleavage was allowed and mass tolerance set to 0.1 Dalton for reliable protein identification. Oxidized methionine and carbamidomethyl modifications were taken into account. Identifications were accepted when the protein score reached significance according to [104].

2.2.8.4. Matching to SWISS-2-D PAGE reference gel. The reference gel published on the SWISS-2-D PAGE website [105] was downloaded and corresponding spots in SWISS-2-D PAGE reference gel were manually matched to the "master gel" of the experiment.

2.2.9. Protein identification II. The exact position on a 2-D gel of three reference proteins of the phosphoenolpyruvate carbohydrate phosphotransferase system (PTS) was identified as follows. Firstly by doping a cell extract with each of 1 μl (6 μg) purified E1, HPr and IIA^{Glc}. Secondly by inspection of the 2-D gel of an extract prepared from *E. coli* over expressing E1, HPr and IIA^{Glc} from a plasmid (pTSHIC9), and thirdly by western blotting.

2.2.10. Western blotting. An unstained 2-D gel was incubated in blotting buffer (50 mM boric acid, pH 9, 0.1% (w/v) SDS, 20% (v/v) methanol) for 30 minutes before blotting 45 minutes at 500 mA on a nitrocellulose membrane. The membrane was washed twice with TBS (10 mM Tris/HCl pH 8.0, 150 mM NaCl) and saturated with 1% milk powder in TBS for 10 minutes. After rinsing twice with TBS the membrane was incubated overnight with the specific antibody (TBS, 0.5% (w/v) BSA, AB: 1:5000 anti IIA^{Glc} (rabbit)). After washing the membrane two times for 3 minutes, the membrane was incubated with a rabbit-specific and peroxidase coupled IgG (TBS, 0.5% (w/v) BSA, AB: 1:5000 (v/v)) for 2 h. The membrane was washed two times for three minutes with TBS and then treated with peroxidase staining solution (TBS, 6% (v/v) chloro-1-naphthol 0.3% (w/v) in MeOH, 0.002% (v/v) H_2O_2 30%). After 2 to 5 min the antibody bound spots became visible and the reaction was stopped by washing the membrane with TBS.

2.2.11. Image analysis and data processing. Samples of two cultures of *E. coli* MC4100 and MC4100Δ*ptsI* were each resolved on six 2-D gels forming one experimental set. Scanned gel images were processed with the Phoretix 2D Advance V 5.0, 5.1 or V.6.01 gel analysis software running either on an DELL DimensionTM 8100 Multimedia PC (Intel pentium 4 processor 1.8 GHz, 256 MB RAM memory) running on Microsoft Windows 2000 or DELL Dimension 4550 PC (Intel pentium 4 processor 2.8 GHz, 1024 MB RAM memory) running on Microsoft Windows XP. Spot intensities were determined for all spots in all gels of the experiment with the following parameters: Sensitivity 9850, Operator size 27, Noise factor 5, Background 1, Split factor 8-9, Minimum spot area was set to 16 pixels. Detected spots with volumes below 2500 AU and/or circularity below 0.3 were filtered out and deleted. After background subtraction in "mode of non-spot" the spot volumes were normalized as fractions of the total spot volume per gel. Conversion to ppm was done by multiplying each normalized spot volume by 10^6. The analyzed gels were assembled to sets of "slave gels", "assistant gels" (Figure 12) and a "master gel" (reference gel) (Figure 13). The qualitatively best gel was chosen as master gel. Corresponding spots in the master gel and in the slave and assistant gels were matched by the built-in procedures of Phoretix 2D advance.

Correspondence criteria for protein spots of gels from MC4100. Spots were considered as "matched" if:

$M(WT)P_z(x_iy_j) = S_1P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = S_2P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = S_3P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = S_4P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = S_5P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = S_6P_z(x_iy_j)$, then

$S_1P_z(x_iy_j) = S_2P_z(x_iy_j)$ and

$S_1P_z(x_iy_j) = S_3P_z(x_iy_j)$ and

$S_1P_z(x_iy_j) = S_4P_z(x_iy_j)$ ... and all permutations.

Correspondence criteria for protein spots of MC4100Δ*ptsI*:

$M(WT)P_z(x_iy_j) = T_1P_z(x_iy_j)$ and

$M(WT)P_z(x_iy_j) = T_2P_z(x_iy_j)$ and so on.

The matching criteria between protein spots of MC4100 and MC4100Δ*ptsI* is then:

$S_1P_z(x_iy_j) = T_1P_z(x_iy_j)$ and $S_2P_z(x_iy_j) = T_1P_z(x_iy_j)$ and all permutations

M(WT/Mu)= master gel created from a "wild type" gel. $P_z(x_iy_j)$ = Spot with number z and x-coordinate I and y-coordinate j. S_1 = Slave gel (MC4100 series). T_1 = Slave gel (MC4100Δ*ptsI* series).

The following matching parameters were set: Search box size 64 and Vectorbox size 8. Matching was manually supported. Enzyme I, Hpr and IIA^{Glc} were localized from the assistant gels. The matched spots were checked by eye and edited if necessary. Spot volumes of matched spots that were present in at least three of six slave gels were averaged. Raw and averaged data sets were exported for further processing and statistical analysis with Microsoft Excel 2002 SP-2.

2.2.12. Statistical analysis. (i) First spots were removed from the data sets if they were detected in only one or two of the six repeats (x,y-direction outlayers).
(ii) Standard deviations were calculated.
(iii) Assuming that the spot intensity values were normally distributed in the data sets the averaged spot intensities were subjected to a two tailed, unpaired heteroskedatic *t*-significance test. The *t*-Test was done with Excel according the formula: =TTEST(($Array_1$);($Array_2$);2;3), $Array_1$= spot set of MC4100, $Array_2$= corresponding spot set of MC4100Δ*ptsI,* 2= two tailed, 3= heteroskedatic. The test value T was calculated according to: $T= (X-Y)/sqrt((var_1/m)+(var_2/n))$, X=average of spot intensities in MC4100, average of spot intensities in MC4100, Y=average of spot intensities in MC4100Δ*ptsI*, m=number of samples in the MC4100 spot set, n=number of samples in MC4100Δ*ptsI* spot set, var1=variance of the MC4100 spot set, var2=variance of the MC4100Δ*ptsI* spot set (with variance as square of the standard deviation) [108-110]. Excel calculated the degrees of freedom according to the Welch-Satterwaite-Approximation: $d.f.=((var_1/m)+(var_2/n))^2/((var1/m)^2)/(m-1)+(var_2/n)^2)/(n-1))$, d.f.= degrees of freedom, var_1=variance of the MC4100 spot set, var_2=variance of the MC4100Δ*ptsI* spot set, m=number of samples in the MC4100 spot set, n=number of samples in MC4100Δ*ptsI* spot set. The p-value was calculated by integration of

the area below the t-distribution curve for the calculated degrees of freedom above test value T and compared to the p-value of t critical.

(iv) The null hypothesis (no difference between MC4100 and MC4100Δ*ptsI*) was judged from the p-value output from Excel. It was rejected if the combined data sets were not *t*-distributed hence the calculated p-value was lower than the p-value for t-critical (α=0.05) and hence the two compared means belonged to two different data sets [111].

(v) The resulting p-values of the *t*-significance tests were transformed into s = significant ($p<0.05$) or ns = not significant ($p>0.05$). (vi) In spot sets that contained more than 4 spots single outlayers in z-direction were identified according to the formula ($X_n+1> x + ks$) and removed. (X_n=value of the outlayer, X=arithmetical average calculated without the outlayer, s=standard deviation calculated without the outlayer, k=coefficient value (for n=4 ->k=7, for n=5->k=6) [106, 107]). For spot sets that remained in the data set spot averages standard deviations and p-values were new calculated.

2.2.13. Selection criteria for protein identification. Protein spots were selected for protein identification from preparative gels applying the following five serial criteria. Firstly, protein spots must be reproduced in a least three out of six gels of MC4100 or MC4100Δ*ptsI* respectively. If a spot appeared in just on gel in a set of gels, it was regarded as absent. Secondly there must be a factor of 2 between two corresponding (averaged spot volumes). Thirdly the *t*-significance level ($p < 0.05$) must be reached and finally all selected spots had to pass a visual control as fourth criteria. The last criterion to fulfil was the reproducibility of a spot in preparative 2-D electrophoresis. For proteins with unchanged expression that were chosen for landmark identification, spots had to meet criteria one and four and a change protein expression below factor 2.

2.2.14. Data mining. Each identification was annotated with the Swiss-Prot entry protein name, SWISSPROT accession number, molecular weight and isoelectric point (theoretical vs. observed), location in 2-D master map, catalytic function (if available), regulation factor between *E. coli* strains MC4100 and MC4100Δ*ptsI*, operon structure and function, presence of upstream regulatory sites (involvement in a regulon) and pathway involvement.

2.2.15. Protein name and accession number. Protein Identifications were first linked to SWISSPROT entry protein name and accession number by searching each identification (gene name) on us.expasy.org with the full text search tool.

2.2.16. Regulation factor and map location. Regulation factors were taken from protein spots that fulfilled the selection criteria for protein analysis and map locations from the 2-D master gel.

2.2.17. Calculation of theoretical and observed molecular weight Mr and pI. Theoretical M*r* and p*I* were calculated at the website of [113]. Observed M*r* and p*I* were calculated by measuring all x- and y-coordinates of each spot. The measured distances (x) were entered in the formula $pI=3.5+(x*5.5/16.8)$ for the first dimension and (y) in $Mr=e^{((y+45.173)/5.0956))}$ for the second dimension.

2.2.18. Operon search. Each identified gene was compared to the DNA Data Bank from the Center for Information Biology of Japan [114]. Additionally all genes that remained unmatched were searched in the E*. coli* genome at the National Center for Biotechnology Information [115].

2.2.19. Regulon search. Upstream regulatory sites were searched in the EcoCyc and Rockefeller regulon databases for *E. coli* K-12 [116, 117].

2.2.20. Pathway search. Metabolic pathway involvement, identified gene products were searched in the EcoCyc database [116].

2.3. Molecular Biology.

2.3.1. Reporter plasmid construction. The promoter regions of five selected genes were each amplified from purified genomic DNA of *E. coli* W3110 with primer pairs introducing Hind III and Nsi I endonuclease restriction sites (details of primers and reaction conditions see tables 3-8). Purified PCR fragments and reporter plasmid pBRPdhaR-lacZ (gift of Christoph Bächler) were digested with Hind III and Nsi I, PCR product and vector fragment isolated by agarose gel electrophoresis (0.8% agarosegel, 200 mA, 80 V, limiting) and ligated [118] to afford the five reporter constructs: pBRGapA-lacZ, pBRYfiD-lacZ, pBRFrwC-lacZ, pBPflC-lacZ and pBRYeaD-lacZ. Success of ligation was checked by growth on

LB amp plates and restriction analysis [118]. Each construct was sequenced by Microsynt.

2.3.2. Isolation of genomic DNA. Genomic DNA from *E. coli* for PCR was purified by the DNeasy tissue kit from Promega following the instructions of the manufacturer.

2.3.3. Isolation of plasmid DNA**.** Purification of Plasmid DNA was done with the Wizard Plus SV Mini- or Midiprep DNA purification Kit and the QIAquick PCR purification kit as described in [118] or following the manufacturer's manual.
DNA restriction, restriction analysis and *DNA Ligation and transformation of* $CaCl_2$ *competent cells* were performed as described in [118].

Table 3. PCR primers.

Primer 1 (Nsi I, promoter pflC2, 303 Bp)

(5’*TATGCAGAACCA*ATGCATCCTGGATCTCCTTCGACA3’)

Primer 2 (Hind III, promoter pflC2)

5’*CCC*AAGCTTAAGGGCCAACAGCGGTACTGAAGTCAG3’

Primer 3 (Hind III, promoter pflC (frwC) 3077 Bp)

5’*CCC*AAGCTTTCACAAATAAATTCCACAATCAGGGCCA3’

Primer 4 (Nsi I, promoter yfiD 346 Bp)

5’*AATGCAGAACCA*ATGCATCCGGAAAAATATCCGCAG3’

Primer 5 (Hind III, promoter yfiD)

5’*CCC*AAGCTTTGGGTAGTTGGCGTCAGCGTTTTGCGT3’

Primer 6 (Nsi I, promoter yeaD 318 Bp)

5’*AATGCAGAACCA*ATGCATTTTTACAGGTAAAAAAAA3’

Primer 7 (Hind III, promoter yeaD)

5’*CCC*AAGCTTAGCTGCAACTTACGAGCAGATCAAAGC3’

Primer 8 (Nsi I, promoter gapA 368 Bp)

5’*AATGCAGAACCA*ATGCATATATTCCACCAGCTATTT3’

Primer 9 (Hind III, promoter gapA)

5’*CCC*AAGCTTTGCCGAAGGTTTATTAGCCATTTGCTC3’

Underlined: Restriction sites for Nsi I and Hind III, *Italic:* base composition immediately after restriction site for optimal restriction, recommended by NEB [119].

Table 4. Reaction conditions for PCR I

PCR I: pflC2	**Amount [µl]**	**Concentration**
Reaction buffer	5	100 mM Tris HCl pH 8.85, 250 mM KCl, 50 mM $(NH4)_2SO_4$, 20 mM $MgSO_4$, (Roche)
dNTP	3.75	4 mM dNTP Mix (dATP, dTTP, dGTP,dCTP)
Template (genomic DNA)	2	
Primer 1	1	100 µM
Primer 2	1	100 µM
Pwo Polymerase	1	Roche (5 u/µl)
Water	37	

Programme Thermocycler Perkin Elmer Thermo Cycler

Cycles	**Time, Temperature**
1	2', 94 °C
30	1', 94 °C / 1' 67 °C / 35'', 72°C
1	10', 72 °C

Table 5. Reaction conditions for PCR II

PCR II: pflC(frwC)	**Amount [µl]**	**Concentration**
Reaction buffer	5	100 mM Tris HCl pH 8.85, 250 mM KCl, 50 mM $(NH4)_2SO_4$, 20 mM $MgSO_4$, (Roche)
dNTP	3.75	4 mM dNTP Mix (dATP, dTTP, dGTP,dCTP)
Template (genomic DNA)	2	
Primer 1	1	100 µM
Primer 3	1	100 µM
Pwo Polymerase	1	Roche (5 u/µl)
Water	37	

Programme Thermocycler Perkin Elmer Thermo Cycler

Cycles	**Time, Temperature**
1	2', 94 °C
30	1', 94 °C / 1' 67 °C / 4'20'', 72°C
1	10', 72 °C

Table 6. Reaction conditions for PCR III

PCR III: yfiD	**Amount [µl]**	**Concentration**
Reaction buffer	5	100 mM Tris HCl pH 8.85, 250 mM KCl, 50 mM $(NH_4)_2SO_4$, 20 mM $MgSO_4$, (Roche)
dNTP	3.75	4 mM dNTP Mix (dATP, dTTP, dGTP,dCTP)
Template (genomic DNA)	2	
Primer 4	1	100 µM
Primer 5	1	100 µM
Pwo Polymerase	1	Roche (5 u/µl)
Water	37	

Programme Thermocycler Perkin Elmer Thermo Cycler

Cycles	**Time, Temperature**
1	2', 94 °C
30	1', 94 °C / 1' 67 °C / 35", 72°C
1	10', 72 °C

Table 7. Reaction conditions for PCR IV

PCR IV: yeaD	**Amount [µl]**	**Concentration**
Reaktion buffer without $MgSO_4$	5	100 mM Tris HCl pH 8.85, 250 mM KCl, 50 mM $(NH_4)_2SO_4$, (Roche)
$MgSO_4$ solution	1	$MgSO_4$ 25mM, (Roche)
dNTP	3.75	4 mM dNTP Mix (dATP,dTTP,dGTP,dCTP)
Template (genomic DNA)	2	
Primer 6	1	100 µM
Primer 7	1	100 µM
Pwo Polymerase	1	Roche (5 u/µl)
Water	35	

Programme Thermocycler Perkin Elmer Thermo Cycler

Cycles	**Time, Temperature**
1	2', 94 °C
30	1', 94 °C / 1' 55 °C / 40", 72°C
30	1', 94 °C / 1' 60 °C / 40", 72°C
30	1', 94 °C / 1' 65 °C / 40", 72°C
30	1', 94 °C / 1' 67 °C / 40", 72°C
30	1', 94 °C / 1' 69 °C / 40", 72°C
1	10', 72 °C

Table 8. Reaction conditions for PCR V

PCR V: gapA	**Amount [µl]**	**Concentration**
Reaktion buffer without $MgSO_4$	5	100 mM Tris HCl pH 8.85, 250 mM KCl, 50 mM $(NH_4)_2SO_4$, (Roche)
$MgSO_4$ solution	6	$MgSO_4$ 25mM, Roche
dNTP	3.75	4 mM dNTP Mix (dATP, dTTP, dGTP,dCTP)
Template (genomic DNA)	2	
Primer 6	1	100 µM
Primer 7	1	100 µM
Pwo Polymerase	1	Roche (5 u/µl)
Water	24	

Programme Thermocycler Perkin Elmer Thermo Cycler

Cycles	**Time, Temperature**
1	2', 94 °C
30	1', 94 °C / 1' 60 °C / 40'', 72°C
1	10', 72 °C

2.3.4. β-galactosidase assays. A modified procedure of Miller [120] was used. Cells were grown in batch cultures in AA-medium or M9AA to an OD_{550} of 0.5-0.8. To 0.5 ml of cell culture, 0.5 ml of buffer Z (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol, pH 7) was added. Cells were lysed by the addition of two drops of chloroform, one drop of 1% SDS solution and vortexing for 30 seconds. 110 μl of lysate were added to the wells of a 96-well microtiter plate and 1:1 serially diluted with buffer Z. The plate was preheated for 5 min at 28°C and the reaction was then started by adding 30 μl of ONPG (o-nitrophenyl-β-D-galactopyranoside, 4 mg/ml in buffer Z). The increase of OD_{420} was measured for 20 minutes in a Spectramax 250 plate reader controlled by SoftmaxPro 3.1.1. To calculate the β-galactosidase activity in Miller units, the empirical formula was applied: (Miller units = V_0 x 3.37 x OD_{550}-1 x dilution [121]).

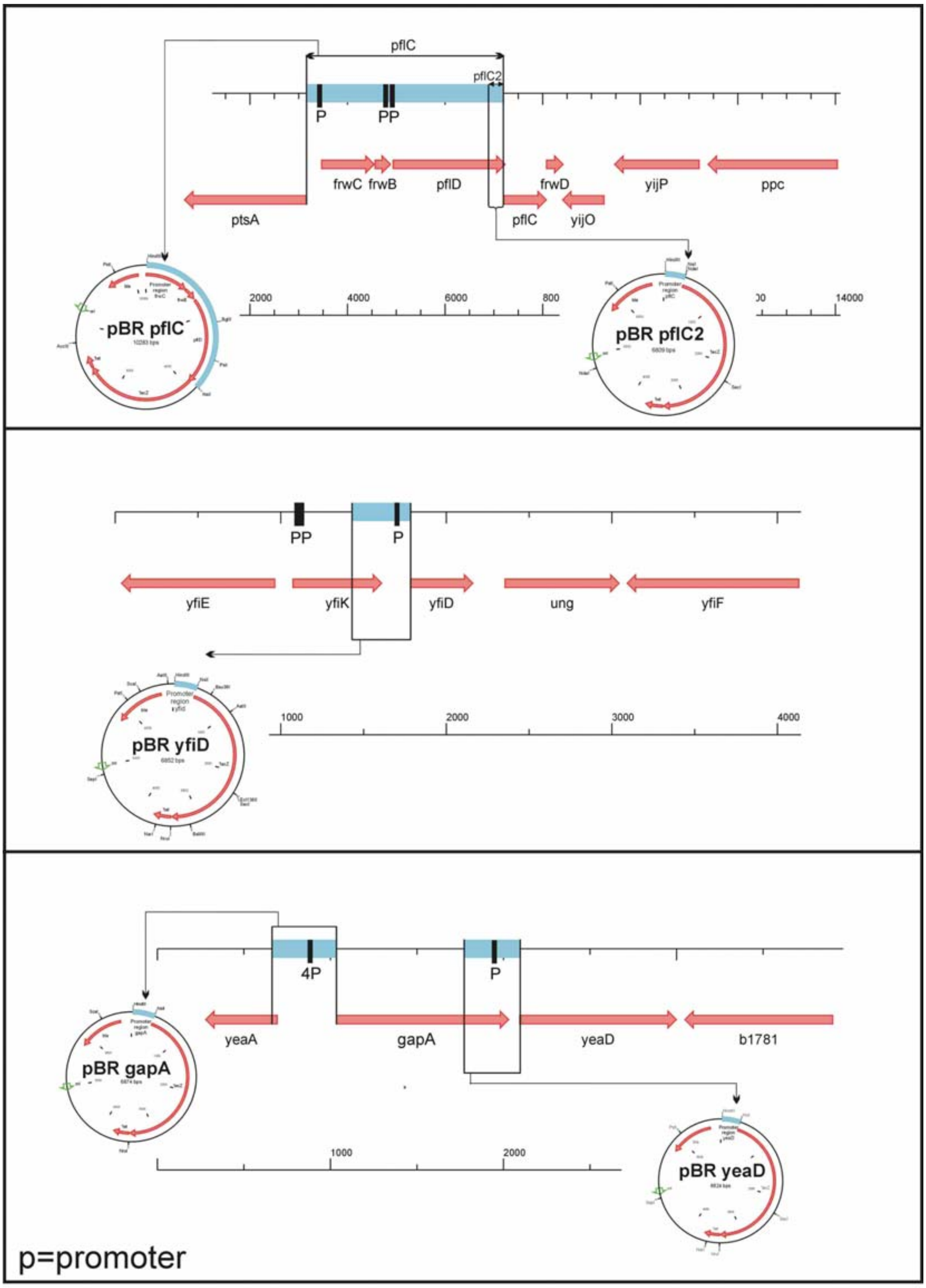
pflC
pflC2
P
PP
frwC
frwB
pflD
frwD
yijP
ppc
ptsA
pflC
yijO
2000
4000
6000
800
14000
pBR pflC
Promoter region frwC
pBR pflC2
Promoter region pflC
PP
P
yfiE
yfiK
yfiD
ung
yfiF
pBR yfiD
Promoter region yfiD
1000
2000
3000
4000
4P
P
yeaA
gapA
yeaD
b1781
pBR gapA
Promoter region gapA
pBR yeaD
Promoter region yeaD
1000
2000
p=promoter

Figure 3a

Figure 3a and 3b. Plasmid construction (3a) and reporter plasmids (3b)

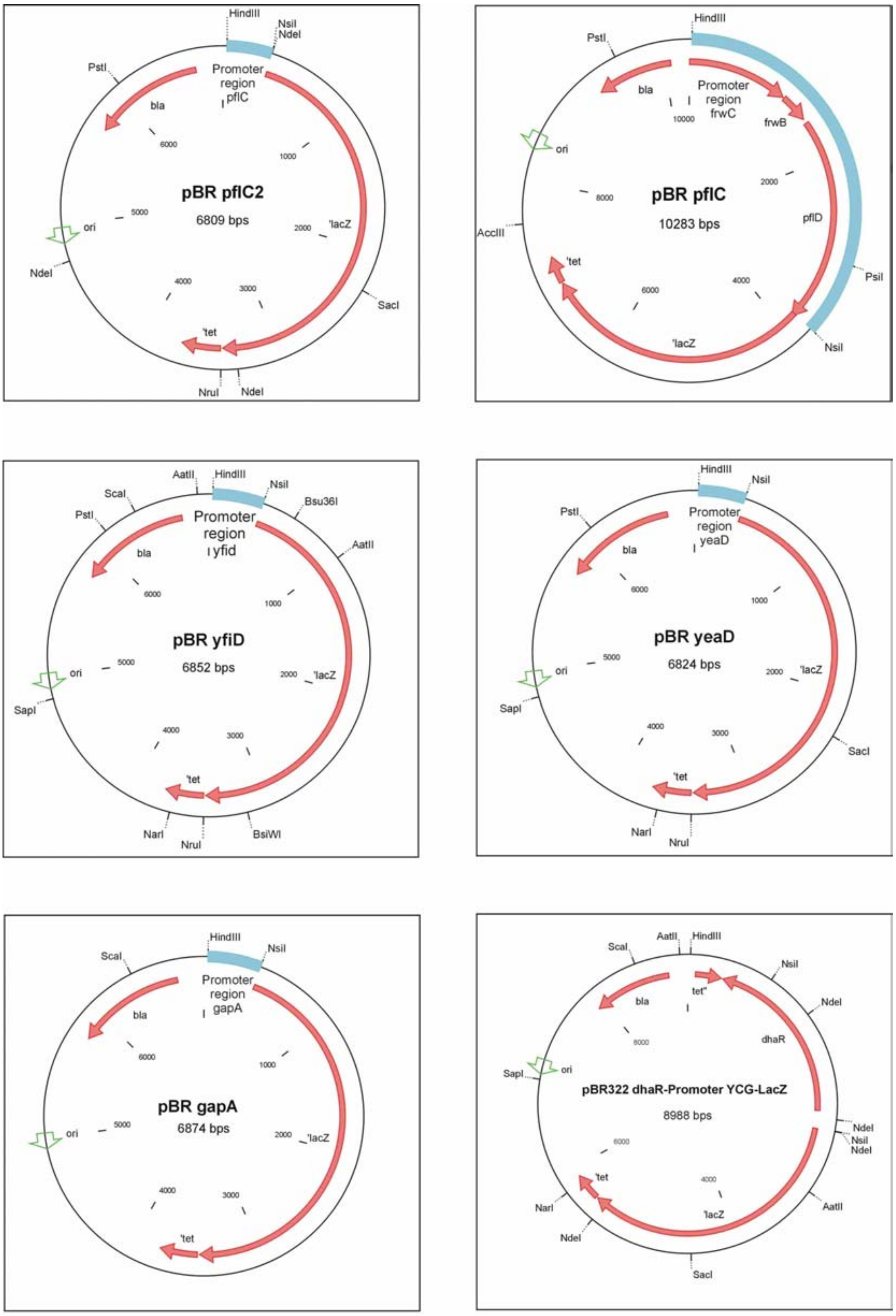
HindIII
NsiI
NdeI
PstI
Promoter region pflC
bla
6000
1000
pBR pflC2
6809 bps
5000
ori
2000
'lacZ
NdeI
4000
3000
SacI
'tet
NruI
NdeI
HindIII
PstI
bla
Promoter region frwC
10000
frwB
ori
2000
8000
pBR pflC
10283 bps
pflD
AccIII
'tet
6000
4000
PsiI
'lacZ
NsiI
AatII
HindIII
NsiI
ScaI
Bsu36I
PstI
Promoter region yfid
AatII
bla
6000
1000
pBR yfiD
6852 bps
5000
ori
2000
'lacZ
SapI
4000
3000
'tet
NarI
NruI
BsiWI
HindIII
NsiI
PstI
Promoter region yeaD
bla
6000
1000
pBR yeaD
6824 bps
5000
ori
2000
'lacZ
SapI
4000
3000
SacI
'tet
NarI
NruI
HindIII
NsiI
ScaI
Promoter region gapA
bla
6000
1000
pBR gapA
6874 bps
5000
ori
2000
'lacZ
4000
3000
'tet
ScaI
AatII
HindIII
NsiI
bla
tet"
NdeI
8000
dhaR
SapI
ori
pBR322 dhaR-Promoter YCG-LacZ
8988 bps
NdeI
NsiI
NdeI
6000
'tet
4000
AatII
NarI
'lacZ
NdeI
SacI

Figure 3b

3. Results and discussion

3.1. Media selection for growth of E. coli MC4100 and MC4100Δptsl.

The medium containing defined concentrations of amino acids as carbon source was chosen for two reasons. (i) It offers the possibility of *in vivo* labelling by substitution of methionine with radioactive ^{35}S methionine while optimized growth conditions can be maintained in future pulse labelling experiments. (ii) Amino acid metabolism does not directly depend on PTS proteins, unlike the metabolism of carbohydrates, which are transported by components of the PTS and of which the uptake is strongly influenced by the PTS.

Cells were grown in AA-medium and in LB-medium (for comparison) and the optical density of the cultures was measured (Figure 4). The generation times were calculated form the slope of the growth curves and the maximum optical density in the stationary phase was determined. They are listed in Table 9.

Table 9. Growth parameters of *E. coli* strains MC4100 and MC4100Δ*ptsI.*

	AA-medium		**LB-medium**	
	$t_{1/2}$ (min)	$A_{600\ nm}$ max.	$t_{1/2}$ (min)	$A_{600\ nm}$ max.
MC4100	28	3.6	24	4.4
MC4100Δ*ptsI*	44	2.6	34	4.4

Although the generation times of MC4100 and MC4100Δ*ptsI* were not exactly identical, they appeared similar enough. The difference should not unduely distort the outcome of the proteome comparison. The knockout phenotype of MC4100Δ*ptsI* was reconfirmed on McConkey glucose indicator plates. The MC4100 wild-type formed red colonies, indicating that glucose was transported and metabolised whereas the *ptsI* mutant formed yellow colonies, as expected of a strain unable to transport glucose.

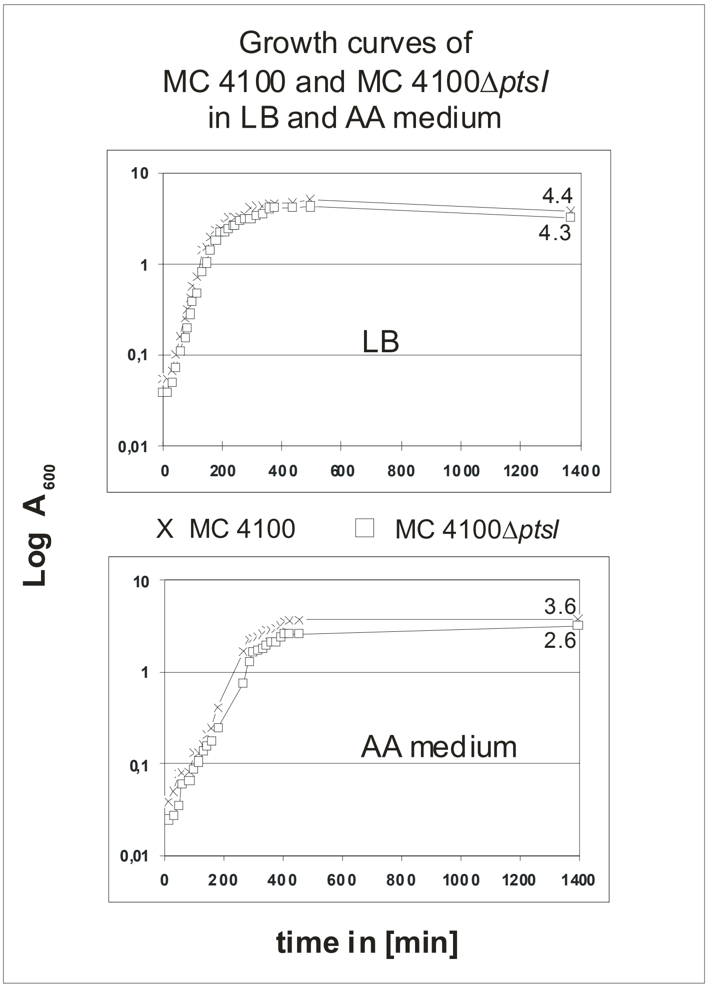

Figure 4. Growth curves of MC4100 and MC4100Δ*ptsI*. Growth was monitored in LB (top) vs. AA-Medium (bottom). In AA-medium both strains grew slower than in LB and grew to lower final cell density than in LB.

3.2. *Growth of E. coli MC4100, MC4100Δpts and MC4100(pTSHIC9) and protein sample preparation*.

Batch cultures of both strains were grown on a shaker to mid log phase (OD_{600}=0.8). *MC4100(pTSHIC9)* was grown on LB medium to an optical density of 0.15, then induced with 0.5 mM IPTG, and harvested when the culture had reached an optical density of 1.0. *MC4100(pTSHIC9)* over expresses EI, HPr and IIA^{Glc}. The protein extract of this strain was later used for the production of a 2-D gel with EI, HPr and IIA^{Glc} as landmark proteins.

3.3. *Quality control of protein sample and optimisation of silver-staining procedure.* Before beginning 2-D electrophoretic separations on a large scale, the quality of the protein samples was controlled on a one-dimensional SDS-PAGE. A Coomassie blue stained gel with the three extracts, *MC4100, MC4100Δpts and MC4100(PTSHIC9)*, is shown in Figure 5. Lanes 1 and 2 show that the protein concentrations of the two samples are comparable. Sharp protein bands and the absence of smears indicate that the sample should be suitable for 2-D separation. In lane 3 the over expressed landmark proteins EI (66 kD), IIA^{Glc} (20 kD) and HPr (10 kD) are clearly visible.

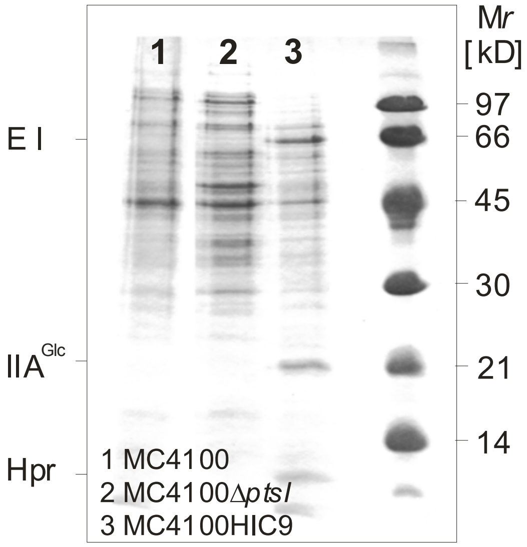

Figure 5. SDS-PAGE of lysed cells of MC4100 (lane 1), MC4100Δ*ptsI* (lane 2) and MC4100HIC9 (lane 3). Sharp protein bands indicate a comparable quality of the samples although some differences in the 50 kD range are already visible. In lane 3 over expressed enzyme I (66 kD), IIA^{Glc} (20 kD) and Hpr (10 kD) are visible.

The original silverstaining procedure from Swain and Ross that was the basis for this optimised procedures, was up-scaled from the “mini-gel” level it was originally designed for. An additional storing step was added at the end.

A precondition for a twin proteomic study is high sensitivity for the detection of a maximum number of well resolved protein spots. Exploratory experiments with metabolic labelling of proteins with S^{35}-methionine were soon abandoned after the initial experiments did not look promising. Instead the silverstaining procedure was standardized. Unlike Coomassie staining which is an equilibrium reaction, silverstaining is a time-dependent process which requires strict control if the protein staining intensities of several 2-D gels have to be compared. Because of this time-dependence, each gel in a set of multiple 2-D gels must be stained for exactly the same time under exactly the same conditions. Without this, averaging of multiple gels and a quantitative comparison of averaged gels is not feasible. The aim of the preliminary silver staining experiments was to find an optimal staining time a good balance between maximum sensitivity, low background and to avoid saturation staining of the spots from the most abundant proteins.

To quantify the kinetics of the silver-staining reaction, ten samples of three concentrations of purified IIAB^{Man} (2300 ng, 230 ng, or 23 ng per band) were each run on a 1-D SDS-PAGE and silver stained as described in section materials and methods. Spot volumes were calculated by numerical integration of the grey values of each pixel within a spot area with Phoretix 2-D advance. Saturation staining was reached already after 5 min with 2300 ng protein per lane and after 9 min with 230 ng. With the minimum concentration of 23 ng saturation was not reached (Figure 6).

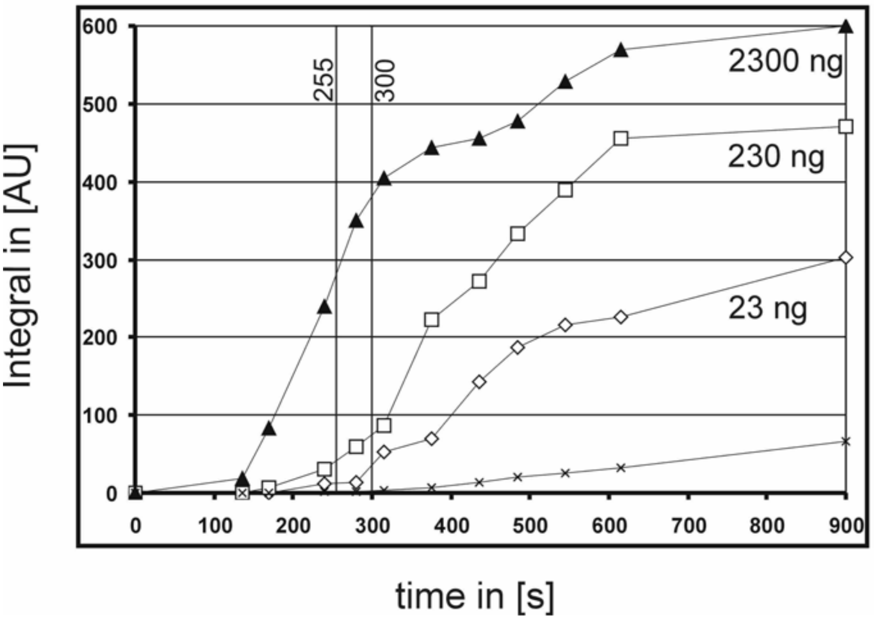

Figure 6. Plot of spot volume versus time to determine when saturation is reached, for 3 different protein concentrations (23 ng, 230 ng, 2300 ng). The symbols used are as follows: 2300 ng, triangles; 230 ng, squares; 23 ng, diamonds; background, crosses. Protein concentration 2300 ng reached saturation after 300 s (indicated by a vertical line. The vertical line at 255 s indicates the finally chosen time point to stop the development.

This result was compared to the average protein content per spot in a 2-D gel with 80 μg total protein load displaying 1000 spots. Spots were sorted according to ascending spot volumes. The total spot volume was set to one million. The theoretical protein content for each spot was calculated. The spots were then grouped in three classes (23 ng, 230 ng, 2300 ng) according to their staining intensities relative to the protein bands analysed in the time dependent silver staining experiment. Figure 7 shows hat the majority of the spots (740) have a protein contents from 0 to 80 ng followed by the group of spots with 80 to 320 ng protein content (220). Spots with more than 320 ng protein content represent the smallest group (40). Thus, only a small group of 10 spots may reach saturation staining at the chosen staining time of 300 s. Except for this staining intensity should be proportional to protein content for all spots.

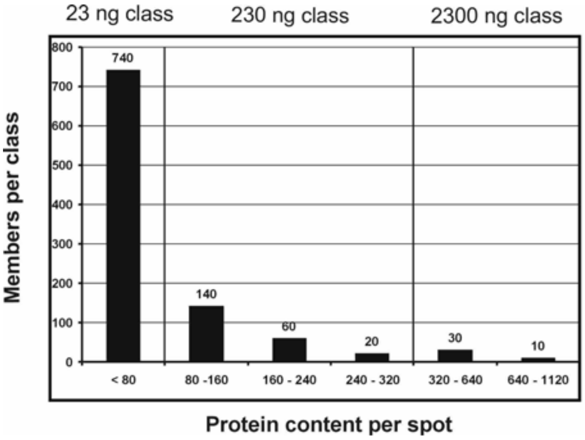

Figure 7. Frequency of protein spots of a given seize.

To test the detection limit of silverstaining for the given staining period of 300 s procedure, marker proteins were two-fold serially diluted and separated by SDS-PAGE. The concentration ranged from 4 μg down to 1 ng (see Figure 1 and 2 in part II of this work). The minimum amount of protein that could be detected after 255 s of staining was 1 ng per protein band.

3.4. Protein expression profiling. To address the question of how a knock-out of enzyme I and as a consequence the inactivation of the entire PTS affects the protein expression profile of *E. coli*, two cell extracts, one of MC4100 and one MC4100Δ*ptsI* were each separated on six analytical 2-D gels. The 12 gels were silverstained and the wet gels were scanned with a bench top scanner. Figure 8 shows the set of 6 2-D electropherograms of MC4100 and Figure 9 the 6 gels of MC4100Δ*ptsI*. After scanning the gels were conserved in dry form on Watman filter and sealed in laminate.

A total of 14111 spots were detected on the twelve gels shown in Figures. 8 and 9 which amounts to almost 1200 spots per gel. The gels obtained from MC4100 contained between 950 and 1482 spots, the gels from MC4100Δ*ptsI* between 1036 and 1319. Table 10 shows the exact number of spots detected in each gel.

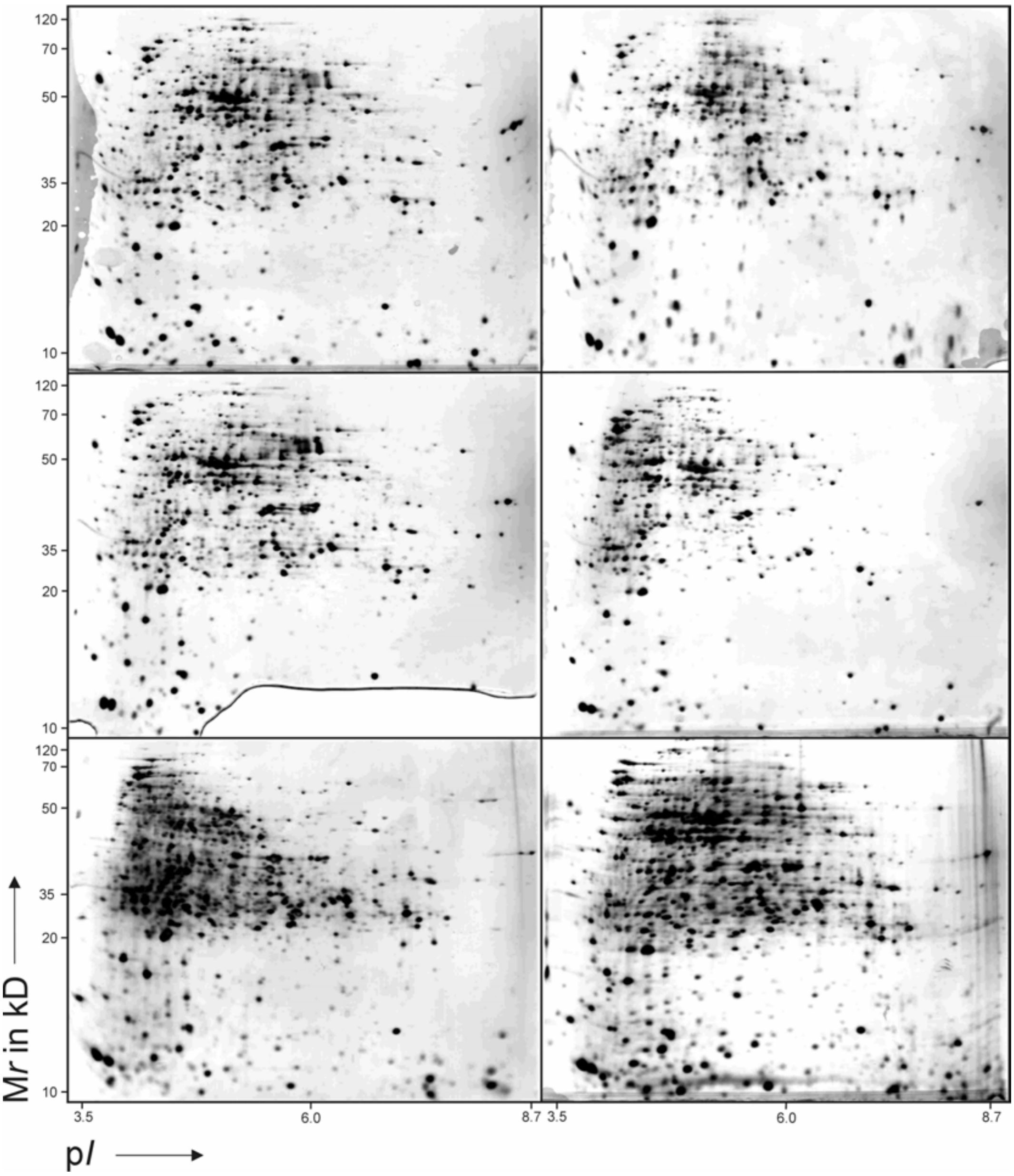
120
70
50
35
20
10
Mr in kD
3.5
6.0
8.7
pI

Figure 8. Six silverstained 2-D electropherograms of MC4100 each one of p*I* 3.5 to 9 and masses 10 kD to 120 kD that were used as slave gels.

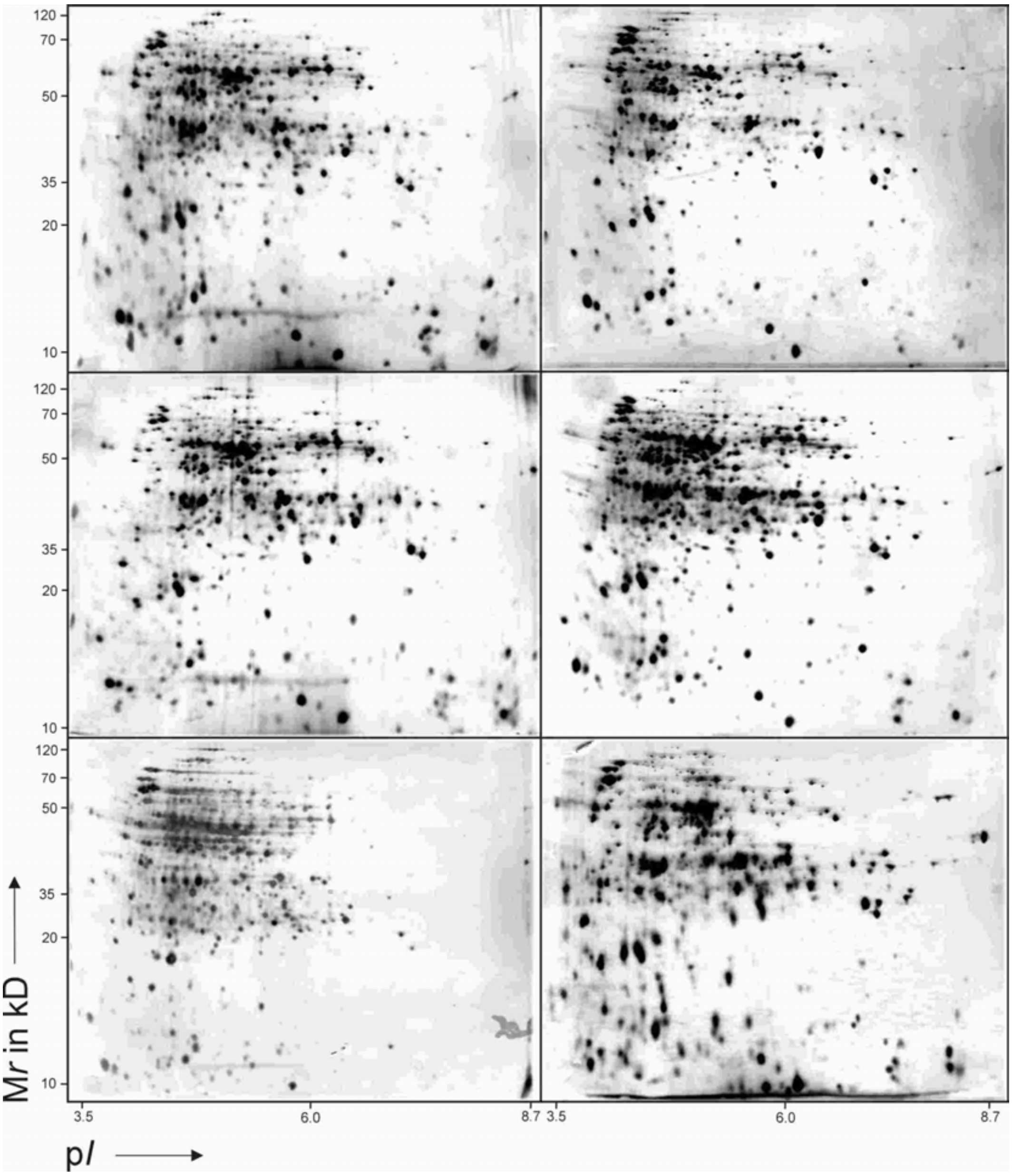
120
70
50
35
20
10
Mr in kD
3.5
6.0
8.7
pI

Figure 9. Six silverstained 2-D electropherograms of MC4100Δ*ptsI* each one of p*I* 3.5 to 9 and masses 10 kD to 120 kD that were used as slave gels.

Table 10. Detected number of spots in each gel of the experiment.

	Detected Spots	
Gel	**MC4100**	**MC4100Δ*ptsI***
1	950	1319
2	1151	1303
3	1482	1036
4	1341	1132
5	1134	1062
6	1137	1064
average	**1199**	**1153**
over all average	**1178**	

Spot volumes were quantified by numerical integration of the pixel values spot area. The background was subtracted in the “mode of non-spot” and the spot volumes were normalized as fractions of the total spot volume per gel. The spot volumes were converted to ppm by multiplying each normalized spot volume by 10^6. The gel images to be analyzed were assembled to sets of six “slave gels” (Figure 12) and one “master gel” (Figure 13).

For each gel set an average gel containing averaged spot volumes was calculated. The criteria for spots to be considered as reproduced in the average gel were set as follows: (i) For inclusion in a set of six gels, a protein spot must be present in at least 3 of the six gels of this set. (ii) For exclusion from a gel set, a spot must be absent in 5 of 6 gels of one set, and present in 3 of 6 gels of the other set.

According to these criteria, 935 spots were reproduced in the gels of MC4100 and 779 spots in the gels of MC4100Δ*ptsI*. 719 spots fulfilled the requirements of criteria one in both gel sets. 216 spots were present only in the average gel from MC4100, and 60 spots only in the average gel of MC4100Δ*ptsI*.

Taken together this shows, that of the on average 1200 spots per gel 935 (78%) could be reproducible detected on the gels of MC4100 and 780 (65%) on the gels of MC4100Δ*ptsI*. A total of 995 protein spots were found reproducibly in one or both gel sets. This is 22% of the theoretical proteome under the assumption that all spots represent different gene products and not posttranslational modifications (for instance proteolytic fragments) of a unique gene product To consider is that the following sub-proteomes are not accessible to standard 2-D electrophoresis: The approximately 300 low abundant proteins [124], approximately 700 membrane proteins [116], ca. 70 proteins with molecular mass above 120 kD [115] ca. 200 proteins of mass below 10 kD [115] and the 600 [144] proteins with an isoelectric point above 9 and below 3.5. This reduces the number of theoretically displayable protein per 2-D electropherogram to less the 2500. There is no clear explanation for the remaining 1300 “missing” proteins. Some may not be expressed under the conditions of cell growth (pH-buffered, amino acids medium at 37°C with aeration), others in too small amounts to be detected and yet others may be short lived, protease sensitive and degraded during sample preparation. For comparison, between 300 and 1000 proteins were displayed on 2-D electropherograms in the published twin proteome studies listed in Table 2**.**

3.5. *Construction of the expression profile.*

To derive the differential expression profile the spot volumes of the almost 1000 reproducibly detected protein spots were averaged, displayed in the comparison window of Phoretix, and exported to an Excel sheet. To find the absent and new appearing spots in the data sets the list was sorted according to ascending spot volumes. After manual inspection of each spot, 122 of the 216 spots which were only present in the gels from MC4100 turned out to be artefacts. Similarly, 42 of the 60 spots which were present exclusively on the gels from MC4100*ptsI* turned out to be staining artefacts (ghost spots), spots which after closer inspection appeared less than three times in the set, and wrongly matched spots. After these corrections, 94 protein spots remained which appeared to be present only in MC4100Δ*ptsI*, and 18 which appeared only in MC4100*.* From the other averaged spot volumes the difference between the means was calculated. The spots were grouped according to volume ratio into 5 classes (Table 11). The volume ratio was < 0.5 for 232 spots and >2 for 150 and it was between 0.5 and 2 for 459 spots which is considered unchanged (Figure 10).

Table 11. Regulation categories.

Regulation categories		
Category	Regulation (1)	Spot number
1	only in MC4100	94
2	< 0.5	138
3	0.5 - 2	459
4	>2	122
5	only in MC4100Δ*ptsI*	28

(1) Average spot volumes of MC4100Δ*ptsI* in times of MC4100

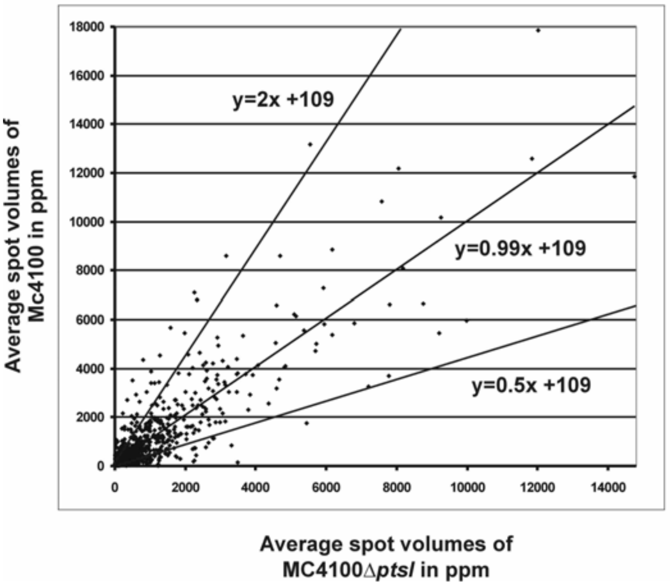

Figure 10. Plot of normalized average spot volumes of MC4100 versus MC4100Δ*ptsI*. The majority of the spots are within the factor two regulation sector.

3.6. Statistical analysis.

Protein loading, electrophoresis, protein staining, protein spot identification and possibly other experimental parameters all contribute to variations in spot pattern and spot size. Because of this analytical variance, the volumes (numerical integral per spot area minus background) of all matched spots (presumed to consist of the same protein) from replicate gels of wild-type and mutant are each averaged and a regulation factor is calculated from the ratio of the two means. Because of the analytical variance it is not always clear, whether a difference between two means is statistically significant or a coincidence. The two-tailed unpaired *t*-test was used to differentiate between significant and coincidental protein expression changes.

The standard deviations of the averaged spot volumes were calculated and these standard deviations were then used to estimate the minimal change of the average protein expression between wild-type and *ptsI* mutant that must occur for being considered statistically significant at 95% confidence level. Means of spot intensities and their variances were calculated from minimally three and maximally six replicate gels ($3 \leq n_i \leq 6$) and subsequently used to calculate the *t*-ratio. The *t*-ratio was then used to calculate the p-value, the probability that the difference between the sample means is not only a coincidence and therefore significant. The difference was considered significant if this probability (the probability of the null-hypothesis) was < 0.05 (95% confidence level). Proteins represented by spot pairs which passed the *t*-test were considered to be up- or down regulated as a consequence of the *ptsI* mutation. These spots are listed in Table 12.

Table 12. Spots that fulfilled statistical significance criteria and were cut out for peptide mass fingerprinting.

	Ref number	**Statistical data**						**regulation**	**number of spots Wt**	**number of spots Mu**
		MC4100		**MC4100Δ*ptsI***						
		Vol. in ppm	**SD in ppm**	**Vol. in ppm**	**SD in ppm**	**p-value**	**sign**			
ACNB	9	429	438	1461	502	0,00	**s**	3,9	4	6
ACNB	10	316	345	1031	400	0,01	**s**	3,4	5	6
ACEF	61	1635	446	714	552	0,01	**s**	0,4	6	6
SDHA	79	653	273	1257	652	0,02	**s**	2,1	4	6
DNAK	87	5259	2374	2914	1549	0,07	**ns**	0,6	6	6
MOPA	113	3697	2728	7778	3408	0,05	**s**	2,1	6	6
TNAA	222	6794	3939	2344	662	0,04	**s**	0,3	6	6
ASPC	275	344	148	2217	1314	0,03	**s**	6,4	5	5
HTRA	464	5573	1786	2936	2004	0,05	**s**	0,3	5	6
GAPA	557	635	326	106	89	0,02	**s**	0,2	5	5
MANX	583	1897	840	4645	1955	0,03	**s**	2,4	5	3
YEAD	584	389	176	2058	441	0,00	**s**	5,3	3	4
PYKF	585	231	231	2431	304	0,00	**s**	10,5	6	3
SUCD	656	1559	779	3640	1559	0,02	**s**	2,3	5	3
PFLC	858	59	38	1852	909	0,03	**s**	31,4	6	4
YCGS	971	294	525	951	1111	0,03	**s**	3,3	5	3
YHGI	980	1772	608	5435	2307	0,01	**s**	3,1	6	6
ENO	1020	2142	891	214	215	0,00	**s**	0,1	6	4
B2529	1172	4345	1911	792	239	0,01	**s**	0,2	6	6
USPA	1248	7109	759	2263	907	0,00	**s**	0,3	6	6
YFID	1251	621	150	1116	228	0.01	**s**	1,8	5	5
PGK	1256	13184	5545	2890	1817	0,00	**s**	0,4	6	6
GlVB	1306	4690	330	9985	2426	0,00	**s**	2,1	4	6

Table 13. Spots that were present in only one of the two gel sets.

PTSI	87	nd	nd	0	0	nd	nd	abs	6	0
GLYA	306	1357	572	0	0	nd	nd	abs	6	0
YBHE	427	2471	1454	0	0	nd	nd	abs	6	0
GLNH	917	2925	1240	0	0	nd	nd	abs	6	0
SODA	954	3797	909	0	0	nd	nd	abs	6	0
NUSG	1050	527	248	0	0	nd	nd	abs	6	0
TPX	1079	8580	2139	0	0	nd	nd	abs	6	1
YFIA	1237	305	515	0	0	nd	nd	abs	5	1
YDHD	1345	0	0	773	590	nd	nd	new	0	6

3.7. Protein identification.

For identification by MALDI-TOF and Edman degradation proteins had to be separated in preparative amounts. Instead of 80 µg protein 800 µg cell extract were loaded per gel. IPG stripes were rehydrated with cell extract under low voltage (30V), the focusing time was prolonged to 276000 Vh and a final high voltage step of 80000 Vh was added. Proteins were stained with Coomassie Blue which is compatible with mass spectrometric protein identification. An example of a preparative gel is shown in Figure 11.

Approximately 800 proteins could be displayed on a preparative gel. However, many spots identified by silver-staining in the analytical gels could not be unambiguously identified and recovered from the preparative gels. This can have several reasons. Firstly a loss of resolution. Abundant proteins form a ridge during isoelectric focussing and therefore are squeezed out of the IPG strip. This results in protein smears which obscure low abundant protein spots. Secondly, the spot intensities of silverstained proteins are different from the intensities obtained with Coomassie blue, so much different that some proteins may be visible only in a silverstained gel and others only in CBB stained gels. Taken together this resulted in approximately 400 spots which were present on silver stained gels but missing from the preparative gel. Of the 83 spots which showed statistically significant changes of expression level on the analytical gels (24) 28% could be detected on preparative gels (plus 7 spots which disappeared from wild type gels and one which newly appeared in the mutant).

These proteins were excised from five preparative 2-D gels. The spots were destained and digested with trypsine in the gel plugs for 12 h.

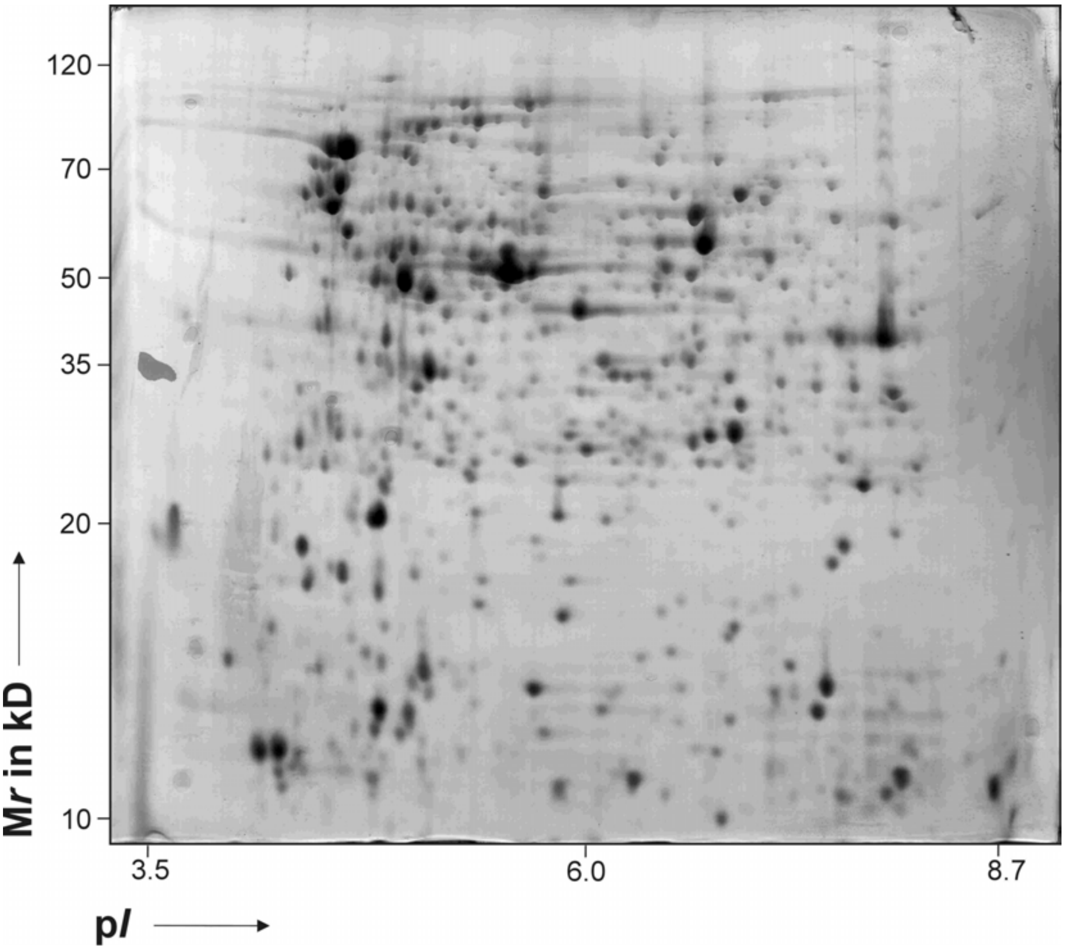

Figure 11. A preparative Coomassie Blue stained 2-D electropherogram displaying approximately 800 protein spots (p*I* 3.5–8.7, mass 10 – 120 kD). From this gel, samples were taken for MALDI-TOF mass spectrometric protein identification.

3.7.1. Protein identification with nanospray ESI-TOF. Six Coomassie blue stained protein spots (2 landmarks, 2 weakly regulated and 2 strongly regulated) were digested with trypsin and analysed with a Qstar-Pulsar mass spectrometer. They could be identified as D-ribose periplasmic binding protein (Figure 13, B4, RbsB), D-galactose binding periplasmic protein (A4, MglB), 6-phospho-beta-glucosidase (B4, AscB), superoxide dismutase (B5, SodA), periplasmic glutamine-binding protein (B5, GlnH).

The major problem using the Qstar-Pulsar mass spectrometer was contamination by keratin, a human epithelian protein which is omnipresent. Tryptic keratin peptides contaminated the samples which rendered the identification of an unknown protein difficult and mostly impossible.

3.7.2. Protein identification by MALDI-TOF. Because of the difficulties encountered with nanospray ESI-TOF the peptide mixtures were analysed by MALDI-TOF at Hoffmann-La Roche. The peptide mixtures obtained by in-gel trypsin digestion were extracted and loaded on a MALDI anker target plate, and analysed. Peptide spectra were submitted to peptide mass fingerprinting in the Roche in-house database. The results are presented in Table 14. 64 out of 72 different protein spots could be identified. Five of the differentially expressed proteins appeared in 13 different spots. Aconitate hydrase (AcnB, A1) was identified in three and glyceraldhyde-3-phosphate dehydrogenase (GapA, B4, B5, A6) in four different spots. Serine proteinase do (HtrA, B4), phosphoglycerate kinase (Pgk, A6, B4) and prolyl-t-RNA synthase (ProS, A2) were each identified in 2 spots.

Figure 12. Scheme of protein expression profiling and protein analysis that was done on *E.coli* MC4100 and MC4100Δ*ptsI*. The figure is divided in two parts, analytical (top) and preparative (bottom).

Analytical level: In the top centre of the figure the master gel that was used for x, y-coordinate matching of the protein spots in the slave gels (below) is shown. To the right and the left of the master gel 4 assistant gels are shown from where the proteins enzyme I, IIA^{Glc}, Hpr, universal stress protein and thiol peroxidase were localized and matched. Directly below the master gel two sets of 6 replicated slave gels are shown that were used for gel averaging.

Preparative level: In the in the lower section of the image preparative loaded 2-D electropherograms are shown and one PVDF membrane. Samples were taken from gels from for MALDI- and ESI-TOF mass spectrometric identification and from PVDF membranes for Edmann degradation.

Protein expression profiling and protein analysis *E.coli*

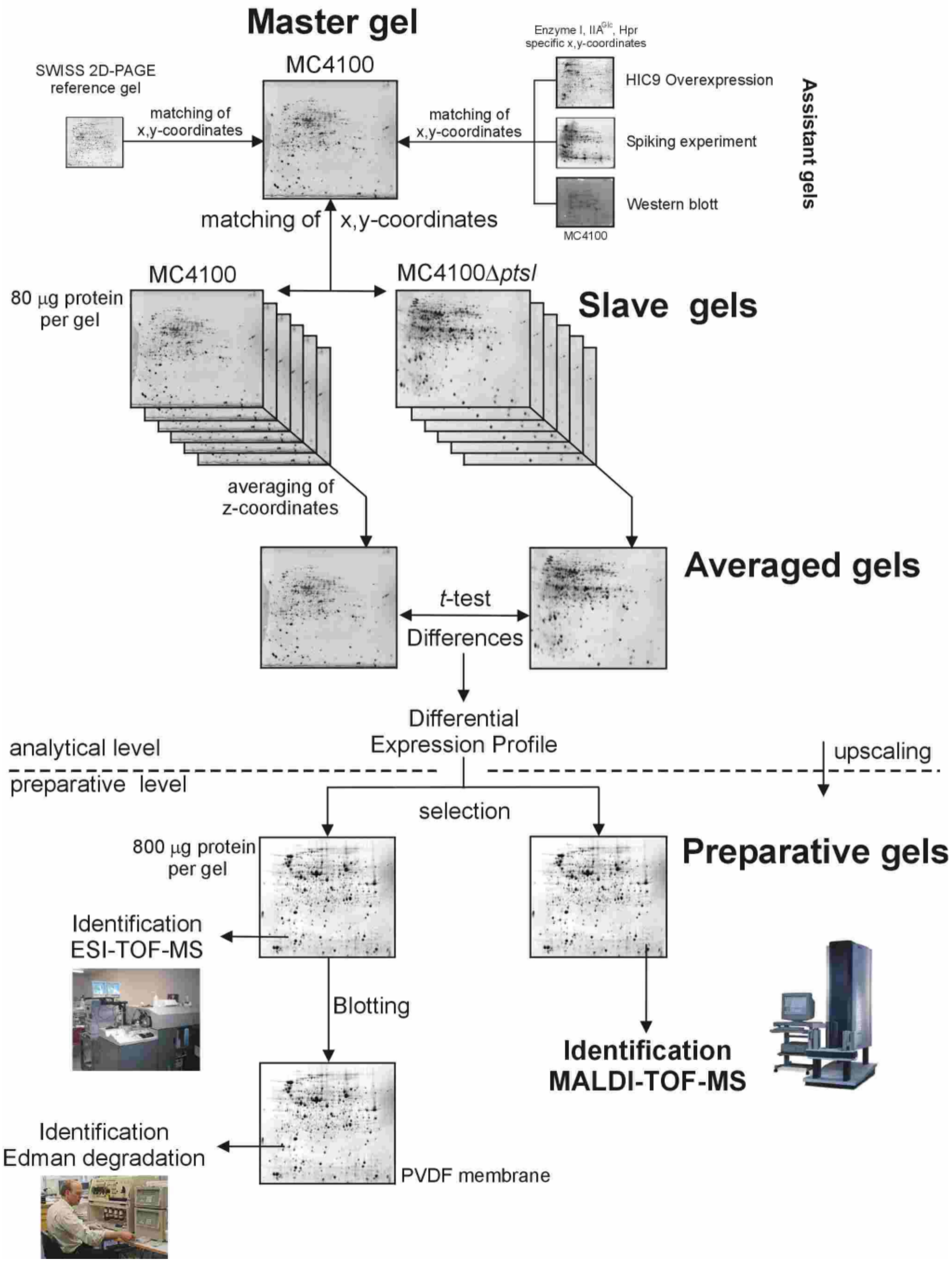
Master gel
MC4100
SWISS 2D-PAGE reference gel
matching of x,y-coordinates
Enzyme I, IIA^Glc, Hpr specific x,y-coordinates
matching of x,y-coordinates
HIC9 Overexpression
Spiking experiment
Western blott
MC4100
Assistant gels
matching of x,y-coordinates
MC4100
80 µg protein per gel
MC4100Δ*ptsI*
Slave gels
averaging of z-coordinates
Averaged gels
t-test
Differences
Differential Expression Profile
analytical level
preparative level
upscaling
selection
800 µg protein per gel
Preparative gels
Identification ESI-TOF-MS
Blotting
Identification MALDI-TOF-MS
Identification Edman degradation
PVDF membrane

Table 14. Identified proteins.

Identified proteins

					1	2	3	4
Gene	**Swissprot**	**Loc**	**Protein name**	**Reg**	**M*r***		***pI***	
					theor	**obs**	**theor**	**obs**
ACCC	P24182	B3	Biotin carboxylase	0,2	49320	52400	6,65	7,60
ACEF	P06959	A2	Pyruvate dehydrogenase	0,4	65965	71700	5,09	4,50
ACNB	P36683	A2	Aconitate hydratase 2	3,4	93498	85600	5,24	5,20
ACNB	P36683	A2	Aconitate hydratase 2	3,9	93498	83900	5,24	5,20
ACNB	P36683	A2	Aconitate hydratase 2	4,6	93498	83900	5,24	5,20
AHPC	P26427	A5/6	Alkyl hydroperoxide reductase C22 protein	0,7	20630	20400	5,03	4,70
ACSB	P24240	B4	6-phospho-beta-glucosidase	0,5	53934	44700	5,52	6,85
ALDA	P25553	A3	Aldehyde dehydrogenase	1,3	52142	58950	5,07	4,85
ASPA	P04422	A3	Aspartate ammonia-lyase	1,1	52356	57812	5,19	5,10
ASPC	P00509	A3	Aspartate aminotransferase (Transaminase A)	6,4	43573	51390	5,54	6,20
ATPD	P00824	A3	ATP synthase beta chain	1,00	50194	53450	4,90	4,60
B2529	P77310	A6	NifU-like protein (NifU)	0,2	13848	12000	4,82	4,30
CRR	P08837	A6	PTS system, glucose-specific IIA component	0,7	18120	17800	4,73	4,30
CSPC	P36996	B6	Cold shock-like protein CspC	1,1	7271	7220	6,82	8,20
DAPA	P05640	B4	Dihydrodipicolinate synthase	0,6	31270	32700	6	7,00
DEOA	P07650	A3	Thymidine phosphorylase	6,3	47207	48450	5,51	5,00
DKSA	P18274	A6	Dnak suppressor protein	0,1	17528	15220	5,06	4,80
DNAK	P04475	A2	Chaperone protein DnaK	0,6	68984	67640	4,83	4,40
ENO	P08324	A6	Enolase (2-phosphoglycerate dehydratase)	0,1	45523	19260	5,32	7,60
FABH	P24249	A4	3-oxoacyl-[acyl-carrier-protein] synthase III	0,6	33515	39000	5,08	4,70
FRR	P16174	B6	Ribosome releasing factor	0,7	20639	19260	6,44	7,20
FUSA	P02996	B6	Elongation factor G	0,3	77450	48450	5,24	5,30
GAPA	P06977	B4	Glyceraldehyde 3-phosphate dehydrogenase A	0,2	35401	36100	6,58	7,60
GAPA	P06977	B4	Glyceraldehyde 3-phosphate dehydrogenase A	0,3	35401	35400	6,58	7,50
GAPA	P06977	A6	Glyceraldehyde 3-phosphate dehydrogenase A	0,5	35401	8960	6,58	4,40
GAPA	P06977	B5	Glyceraldehyde 3-phosphate dehydrogenase A	0,9	35401	24370	6,58	7,30
GLNH	P10344	B5	Periplasmic glutamine-binding protein; permease	absent	27190	23440	8,44	7,70
GLYA	P00477	B3	Serine hydroxylmethyltransferase	absent	45316	49410	6,03	6,90
GPMA	P31217	B5	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	0,6	28425	26900	5,86	6,60

Table 14. continued

Gene	Swissprot	Loc	Protein name	Reg	M*r*		*pI*	
					theor	obs	theor	obs
GLVB	P31451	A6	PTS system, arbutin-like IIB component	2,1	17626	10000	5,92	6,10
HTRA	P09376	B4	Periplasmic serine protease do; heat shock protein HtrA	0,3	49354	39800	8,65	8,70
HTRA	P09376	B4	Periplasmic serine protease do; heat shock protein HtrA	0,6	49354	39000	8,65	8,60
MANX	P08186	B4	PTS system, mannose-specific IIAB component	2,4	34916	34700	5,74	6,50
MGLB	P02927	A4	D-galactose binding periplasmic protein	1,1	35713	32700	5,68	5,00
MIND	P18197	A4	Septum site-determining protein MinD (Cell division inhibitor MinD)	4,8	29482	29900	5,25	5,10
MODA	P37329	B5	Molybdate binding periplasmic protein	0.7	27364	23900	7,81	7,40
MOPA	P06139	A3	GroEL, (60 kDa chaperonin)	2,1	57198	63770	4,85	4,50
MOPB	P05380	A6	GroES (10 kDa chaperonin)	0,1	10387	11560	5,15	5,00
NDK	P24233	A6	Nucleoside diphosphate kinase	1,3	15332	10900	5,55	6,00
NUSG	P16921	B6	Transcription antitermination protein NusG	absent	20400	16150	6,33	7,20
PCKA	P22259	A3	Phosphoenolpyruvate carboxykinase [ATP]	0,6	59643	62500	4,46	5,50
PFLC	P32675	B5	Pyruvate formate-lyase 2 activating enzyme	31,4	32430	25850	8,03	7,80
PGK	P11665	A4	Phosphoglycerate kinase	0,4	40987	46580	5,08	4,80
PGK	P11665	A6	Phosphoglycerate kinase	3,00	40987	9880	5,08	4,10
PNP	P05055	A2	Polyribonucleotide nucleotidyltransferase	2,7	77100	77600	5,11	5,00
PPA	P17288	A5/6	Inorganic pyrophosphatase (Pyrophosphate phospho-hydrolase)	0,7	19772	19650	5,03	4,70
PPIB	P23869	A6	Peptidyl-prolyl cis-trans isomerase b (rotamase b)	1,6	18153	13800	5,51	5,80
PROS	P16659	A2	Prolyl-tRNA synthetase	1,1	63692	71740	5,15	4,80
PROS	P16659	A2	Prolyl-tRNA synthetase	1,3	63692	71740	5,15	4,80
PTSI	P08839	A2	Phosphotransferase system, enzyme I	absent	63561	66000	4,78	4,50
PYKF	P14178	A4	Pyruvate kinase I (formerly f)	10,5	50729	34700	5,77	5,20
RBSB	P02925	B4	D-ribose-binding periplasmic protein	1,0	30950	7080	6,85	6,70
RPE	P32661	A5	Ribulose-phosphate 3-epimerase	6,3	24554	22100	5,13	4,70
RPLI	P02418	B6	50s ribosomal subunit protein l9	0,8	15769	11800	6,17	7,20

Table 14. continued

Gene	Swissprot	Loc	Protein name	Reg	M*r*		*pI*	
					theor	obs	theor	obs
RPLL	P02392	A6	50s ribosomal subunit protein l7/l12	0,8	12164	9880	4,6	4.0
SDHA	P10444	A2	Succinate dehydrogenase, flavoprotein subunit	2,1	64422	68980	5,85	6,30
SODA	P00448	B5	Superoxide dismutase, manganese	absent	22966	22100	6,44	7,40
SUCD	P07459	B4	Succinyl-CoA synthetase alpha chain	2,3	29646	32100	6,31	7,20
TNAA	P00913	A3	Tryptophanase	0,3	52773	54590	5,88	3,80
TPX	P37901	A6	Thiol peroxidase	absent	17704	16460	4,75	3,80
TRXA	P00274	A6	Thioredoxin I	1,0	11675	8790	4,67	4,20
TSF	P02997	A4	Protein chain elongation factor ef-ts	1,6	30291	33370	5,22	5,00
TUFA	P21694	A4	Protein chain elongation factor ef-tu	0,3	43152	45680	5,3	5,40
USPA	P28242	A6	Universal stress protein A	0,3	15935	9320	5,12	4,60
YBHE	P52697	A4	Hypothetical protein ybhE (putative isomerase)	absent	36308	42230	5,06	4,80
YCGS	P76014	A5	Protein DhaL	3,3	22631	21670	5,31	5,20
YDHD	P37010	A6	Protein ydhD, (orf, hypothetical protein)	new	12879	8280	4,76	4,20
YEAD	P39173	B4	Unknown protein from 2-D PAGE (orf, hypothetical protein)	5,3	32665	34710	5,89	6,65
YFIA	P11285	B6	Protein yfiA, (orf, hypothetical protein)	absent	12653	9690	6,19	7,10
YFID	P33633	A6	Putative formate acetyltransferase (orf, hypothetical protein)	2,0	14284	10690	5,09	4,70
YHGI	P46847	A5	Protein yhgI, (orf, hypothetical protein)	3,1	20998	21250	4,52	4,10
YIBO	P37689	A3	Putative 2,3-bisphosphoglycerate-independent mutase	1,3	56194	58960	5,14	4,90

orf = open reading frame

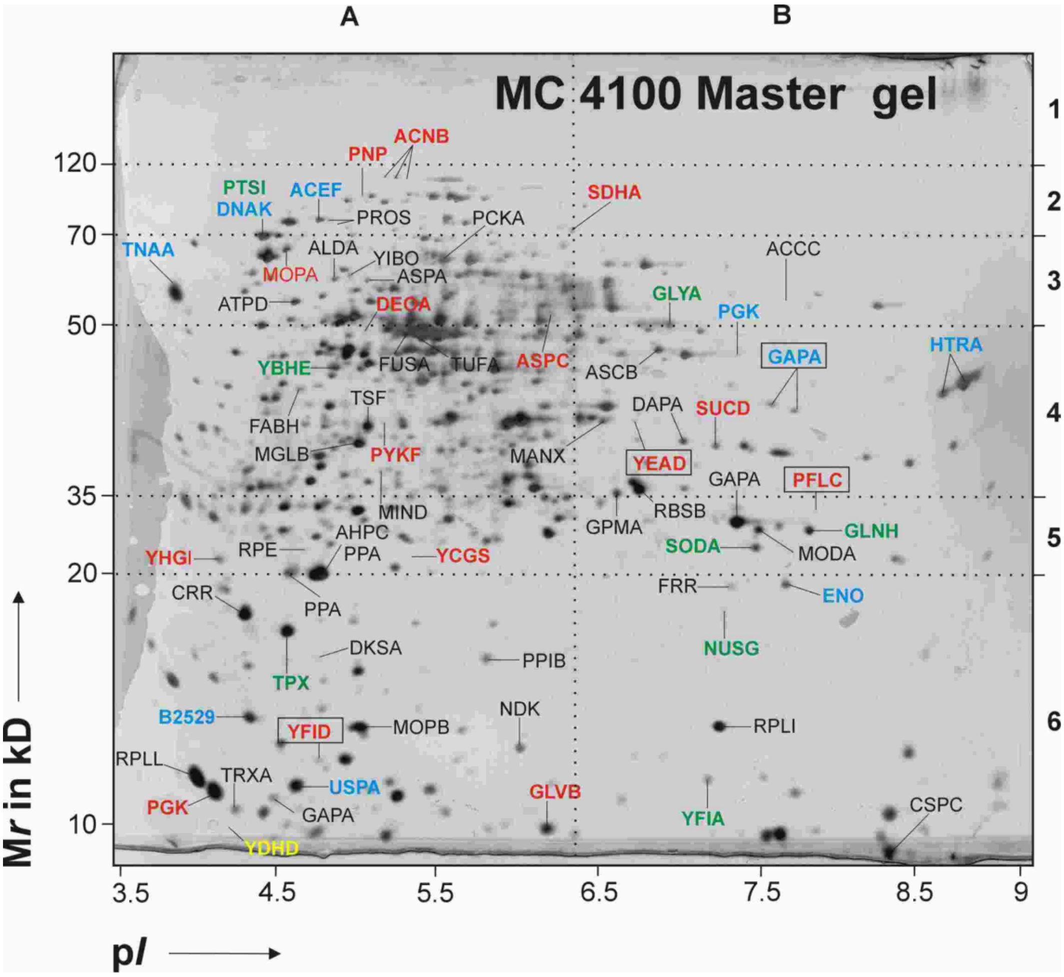
A
B
MC 4100 Master gel
1
2
3
4
5
6
120
70
50
35
20
10
ACNB
PNP
PTSI
ACEF
DNAK
PROS
PCKA
SDHA
TNAA
ALDA
YIBO
MOPA
ASPA
ACCC
ATPD
DEOA
GLYA
PGK
GAPA
HTRA
YBHE
FUSA
TUFA
ASPC
ASCB
TSF
DAPA
SUCD
FABH
MGLB
PYKF
MANX
YEAD
GAPA
PFLC
MIND
RBSB
AHPC
GPMA
GLNH
RPE
PPA
SODA
YHGI
YCGS
MODA
CRR
FRR
PPA
ENO
DKSA
NUSG
PPIB
TPX
NDK
B2529
YFID
MOPB
RPLI
RPLL
TRXA
USPA
GLVB
CSPC
PGK
GAPA
YFIA
YDHD
Mr in kD
3.5
4.5
5.5
6.5
7.5
8.5
9
pI

Figure 13. 2-D electropherogram of MC4100 displaying 1338 silverstained protein spots from p*I* 3.5 to 9 and masses 10 kD to 120 kD that was used as master gel for this study. Identified protein spots are labelled by their gene name. Black labels indicate unchanged protein expression. Red labels indicate proteins with more than 2 times increased protein expression in MC4100Δ*ptsI*, blue labels less than 0.5 times (50%) of the protein expression of MC4100. Green labels indicate spots which were absent in MC4100Δ*ptsI* and the yellow label new appearing spots in MC4100Δ*ptsI.* For convenient orientation the gel is split in two halves A and B and 6 zones of molecular weights.

3.8. Masses and isoelectric points.

For each identification the theoretical molecular weight and isolectric point was calculated by the Calculate pI/Mw software tool of Expasy (http://www.expasy.org/tools/pi_tool.html). Of all identified protein spots the observed masses and isoelectric points were added and compared with the theoretical data. The data obtained in this study were compared with the data from a study of Fontoulakis et al. [36] with the aim to detect possible systematic errors in the data sets. Protein found in both studies were therefore compared and plotted. The plots are shown in Figure 14. Correlation between predicted and theoretical masses are good and comparable to the study of Fontoulakis. In general p*I* correlate less well in both studies and between the studies than masses. In both studies the theoretical p*I*s below p*I* 5.5 are predicted too low. Above p*I* 5.5 theoretical values are predicted too high. This small systematic error arises possibly from the nonlinearity of the pH gradient applied during isoelectric focussing. However there was no general systematic error found that bias the data.

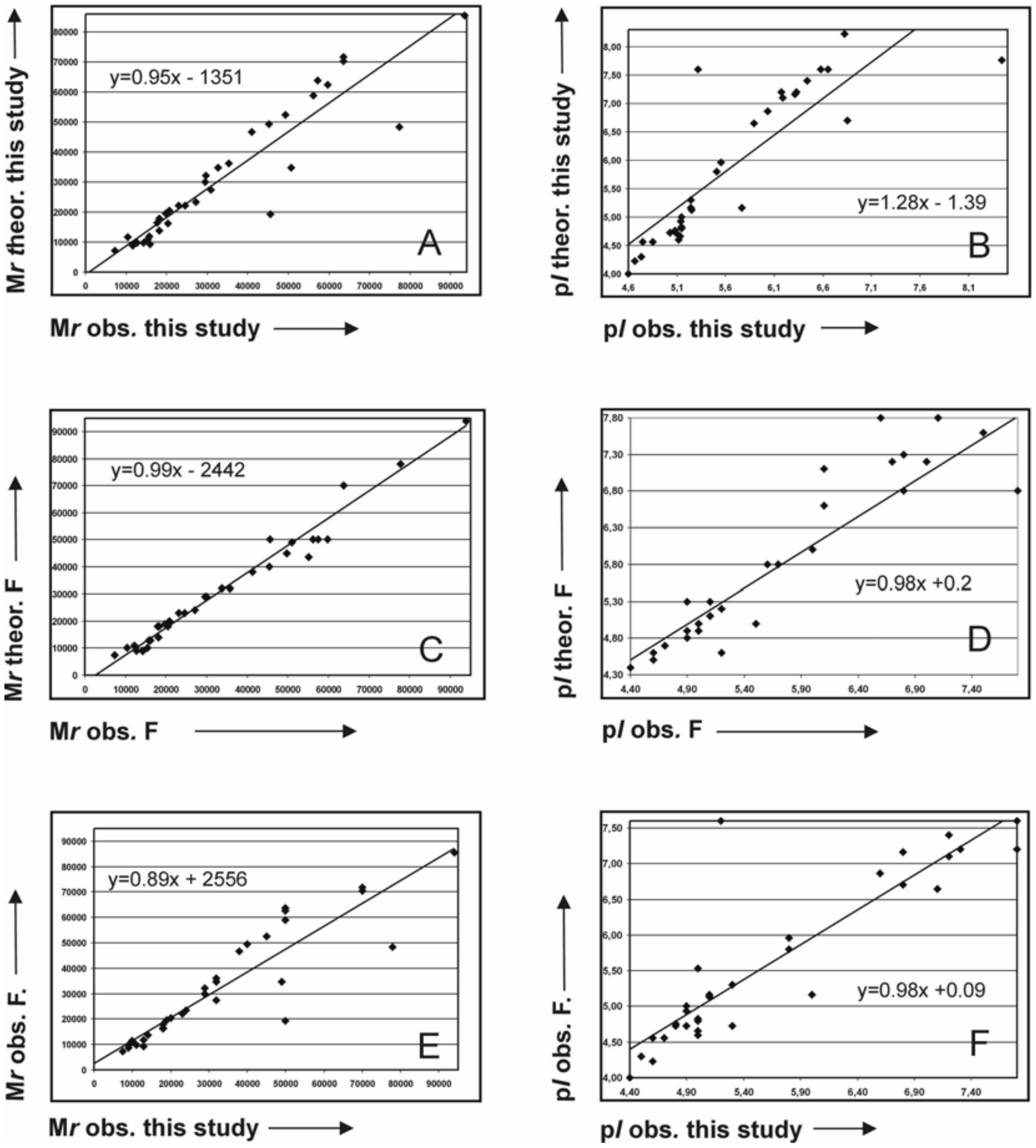

Figure 14. Comparison of observed p*I* and M*r* values with those predicted from the amino acid composition. Panel A and B: Masses and p*I* from, this work. Panel C and D: Masses and p*I* reported by Fountoulakis et al. [36]. Panel E and F: Values observed in this work and reported in [36].

3.9. Validation of the protein identification.

The protein identification has already been validated by the comparison of the experimental and the calculated p*I* values and apparent molecular masses (Figure 14). Because the PTS is primarily involved in carbohydrate catabolism, it can be expected that gene products from these catabolic pathways will be over represented among the identified proteins. Table 15 lists the identified genes according to their main function and compares their number to the total number of genes in these functional groups.

Table15. Comparison of orthologous groups

Class	**Members***	**in [%]**	**Rank**	**Members**	**in [%]**	**Rank**
unclassified	1063	25.6	1	2	6.7	4
Carbohydrate transport and metabolism	326	7.9	2	2	6.7	4
Amino acid transport and metabolism	**322**	**7.8**	**3**	**4**	**13.3**	**2**
General function prediction only	289	7.0	4			
Function unknown	263	6.3	5			
Energy production and conversion	**254**	**6.1**	**6**	**10**	**33.1**	**1**
Transcription	250	6.0	7	3	10	3
Cell wall/membrane biogenesis	217	5.2	8			
Replication, recombination and repair	215	5.2	9	1	3.3	5
Inorganic ion transport and metabolism	176	4.2	10			
Posttranslational modification, protein turnover, chaperones	122	2.9	11	4	13.3	2
Signal transduction mechanisms	122	2.9	12			
Coenzyme transport and metabolism	113	2.7	13			
Cell motility	96	2.3	14			
Nucleotide transport and metabolism	78	1.9	15			
Lipid transport and metabolism	74	1.8	16			
Secondary metabolites biosynthesis, transport and catabolism	61	1.5	17			
Defence mechanisms	43	1.0	18	4	13.3	2
Intracellular trafficking and secretion	35	0.8	19			
Cell cycle control, mitosis and meiosis	32	0.8	20			
Total	4151	100.0		30	100	

* http://www.ncbi.nlm.nih.gov/sutils/coxik.cgi?gi=115

Table 15 shows that by far the largest group (33%, rank 1) are proteins from the group of energy production and conversion. In the *E.coli* proteome this group only has rank 6 of seventeen. This group therefore is over represented. Rank 2 is held by proteins involved in amino acid transport and metabolism. This group also is over represented which may not be so surprising because amino acids are the only carbon and nitrogen source in this experiment. The number of proteins from the other groups is too small for further ranking. The result of this comparison can only point to the plausibility of our findings but not validate them for at least two reasons. (i) Enzymes of the central metabolism are generally expressed in larger amounts than proteins involved in replication or signal transduction. The probability to detect metabolic enzymes on a 2-D gel is therefore high. (ii) Some orthologous groups also contain membrane proteins. These proteins are not only expressed in smaller amounts than soluble proteins, they also are not detectable by standard 2-D gel electrophoresis. Therefore membrane proteins had to be removed form the list before the comparison. Nevertheless, from this data we conclude that a significant change in the central energy producing metabolism (glycolysis and citric acid cycle) had taken place in the enzyme I knockout mutant.

3.10. Validation of the regulation factors inferred from proteome analyses.

Several approaches are available to validate the experimental regulation factors: co-regulation of proteins encoded by a single operon, quantitative reverse transcription PCR, and reporter gene assays.

53 % of all the identified proteins are encoded in polycistronic operons. Of these only two proteins (SucD and SdhA) are gene products of a single operon. The other proteins are products of structurally independent operons. GroEL and GroES were detected both and they are encoded in a single operon. However the proteins were found to be inversely regulated. This is most likely an artefact, and unlikely to reflect a physiological effect because of the fixed stoichiometry of the GroEL chaperone.

Reporter gene assays with β-galacatosidase were used with five selected genes (Table 16). Glyceraldehyde-3-phosphate dehydrogenase (GapA) was chosen as a representative of the prominent group of glycolytic enzymes. Down regulation of *gapA* expression could be confirmed. The gene *yeaD* is immediately downstream of *gapA*. Unlike *gapA*, it was up regulated in the mutant. YeaD is predicted to be controlled by a promoter of its own. The reporter construct had only background β-galactosidase activity in both strains. Up regulation of could not be confirmed.

The putative pyruvate-formate lyase activating protein (PflC) was chosen as target for validation because it is encoded with a putative IIB-like PTS protein in a complex polycistronic operon (Figure 19). Two promoters were cloned, P_{frwC} which is almost 4 kb upstream of *pflC* and a 300 Bp region immediately upstream of the *pflC* coding region. The two reporter constructs displayed only weak activity. It appears (i) that *pflC* is not under the control of the P_{frwC} promoter and/or that the *pflC* promoter is cryptic and (ii) that there is no active promoter immediately upstream of *pflC*. Closer inspection of the region revealed a third possibility, namely that *pflC* forms a dicistronic operon with *pflD*, and that expression is controlled by the promoter upstream of *pflD*.

YfiD was chosen because of its functional relatedness to PflC. PflC activates pyruvate formate lyase, and YfiD is a short open reading frame encoding a peptide which is 48% identical with the C-terminus of pyruvate formate lyase. The region upstream of *yfiD* has no measurable promoter activity. It is known that *yfiD* has an unusual promoter architecture with two upstream binding sites for FNR [145]. Closer inspection of the *yfiD* gene indicated that *yfiD* could also be the second gene of a dicistronic operon with a promoter located upstream of *yfiK*.

Table 16. Results from β-galactosidase assays.

	β-galactosidase activity in Miller units					**T-Test**	
Plasmid	**MC4100**	**SD**	**MC4100Δ*ptsI***	**SD**	**reg**	**p-value**	**sign**
pBR gapA	7096	1879	692	262	0.1	< 0.05	s
pBR yeaD	3.56	0.22	4.2	0.17	1.2	< 0.05	s
pBR yfiD	2.3	0.26	4.1	0.15	1.8	< 0.05	s
pBR pflC (frwC)	27.7	3.7	5.9	1.2	0.2	< 0.05	s
pBR pflC2	16.8	1.8	11.8	1.2	0.7	< 0.05	s

Although direct validation of the results with reporter gene assays was unsatisfactory, there is still some indirect evidence that the regulation factors are real. This indirect evidence is based on the type of gene products that were regulated. Of the total of 30 up- or down regulated genes, 10 are known or predicted form the operator sequence to be under catabolite control. It was expected to find such genes because enzyme I is required for phosphorylation of IIA^{Glc} and because non-phosphorylated IIA^{Glc} activates catabolite repression.

3.11. Enzyme I related protein expression.

3.11.1. Controlled expression of genes under catabolite repression. The presented work was motivated by the idea that enzyme I not only plays a role in carbohydrate transport but could also play a direct role in signal transduction. For other components of the PTS, for example IIA^{Glc}, such a direct role is well documented. Because enzyme I phosphorylates all other PTS components, direct and indirect effects (via IIA^{Glc}) of enzyme I cannot be separated. However, because many IIA and IICB dependent effects have already been described, it should be possible to sort out the novel effects from those already known.

The proteins which are up- and down regulated in a *ptsI* mutant are listed in Table 17 grouped according to the functional classification of Riley.

Table 17. Identified proteins grouped according to Riley categories.

Category	Gene	R.S.	Protein	(+)Reg.
Metabolism				
Energy metabolism				
Tricarboxic acid cycle	*sucD*	Fnr (-) CAP	Succinyl-CoA synthetase alpha chain	2,3
Tricarboxic acid cycle	*sdhA*	Fnr (-) CAP (+)	Succinate dehydrogenase, flavoprotein subunit	2,1
Tricarboxic acid cycle	*acnB*	**CAP (+)**	Aconitate hydratase 2	3,4/3,9
Glycolysis	*pykF*		Pyruvate kinase	10,5
Central intermediary metabolism				
Amino acids	*aspC*		Aspartate aminotransferase (Transaminase A)	6,4
Information Transfer				
DNA replication	*ydhD*		Protein ydhD, (orf, hypothetical protein)	new
Chaperoning, folding	*mopA*		GroEL (60 kDa chaperonin)	2,1
Transport				
PTS	*glvB*		Arbutin-like IIB component, cryptic gene	2,1
PTS	*manX*	Mlc (-)	Mannose specific IIAB component	2,4
Regulation				
Transcriptional level	*yhgL*		Protein yhgI, (orf, hypothetical protein)	3,0
Activation of enzymes	*pflC*		Pyruvate formate-lyase 2 activating enzyme	31,4
Anaerobic respiration	*yfiD*	Fnr (-) (+/-)	Putative formate acetyltransferase activating enzyme 2	2,0
unknown				
	yeaD	CAP	Unknown protein from 2-D PAGE, (orf, hypothetical protein)	5,3
	ycgS		DhaL	3,3

R.S. = binding site for
Regulatory proteins

Category	Gene	R.S.	Protein	(-)Reg.
Metabolism				
Carbon compound utilization				
Carbohydrates/Carbon compounds				
bridge between glycolysis and TCA cycle	*aceF*	Fnr (+/-) CAP(+)	Pyruvate dehydrogenase (dihydrolipoyltransacetylase component)	0,4
Energy metabolism				
Glycolysis	*eno*		Enolase (2-phosphoglycerate dehydratase)	0,1
Glycolysis	*gapA*	CAP(+)	Glyceraldehyde 3-phosphate dehydrogenase A	0,2
Glycolysis	*pgk*		Phosphoglycerate kinase	0,4
Amino acids				
glycine cleavage	*glyA*		Serine hydroxymethyltransferase	absent
Tryptophan utilization	*tnaA*	CAP(+)	Tryptophanase (L-tryptophan indole-lyase, Tnase)	0,3
Information Transfer				
Transcription related	*nusG*		Transcription antitermination protein NusG	absent
Translation	*yfiA*		Protein yfiA, (orf, hypothetical protein)	absent
Protein related	*b2529*		NifU-like protein	0,2
Chaperoning, folding	*ybhE*		Hypothetical protein ybhE (putative isomerase)	absent
Chaperoning, folding	*dnaK*		Chaperone protein DnaK	0,6
Transport				
ATP-binding cassette (ABC) superfamily + ABC-type Uptake Permeases	*glnH*		Periplasmic glutamine-binding protein, permease	absent
Cell process				
adaption to stress	*sodA*	Fnr (-)	Superoxide dismutase, manganese	absent
adaption to stress	*uspA*		Universal stress protein A	0,3
other stresses	*htrA*		Periplasmic serine protease do, heat shock protein HtrA	0,3
Location of gene products				
Periplasmic space	*tpx*	Fnr (-) ArcA (-)	Thiol peroxidase	absent

A gross comparison of the two expression profiles (Table 17) indicates that 16 proteins are down regulated or completely absent in the enzyme I mutant This suggests that some kind of catabolite repression may be at work in the mutant. Catabolite repression is caused by IIA^{Glc} which is not phosphorylated if EI is missing.

Our results confirm that this is the case. We identified tryptophanase (L-tryptophan indole-lyase, TnaA), phosphoglycerate kinase (Pgk) and pyruvate dehydrogenase (AceF) with lower relative protein expression levels in the mutant strain.

Tryptophanase is encoded by *tnaA* which is located in the *tnaCAB* operon that is regulated by tryptophan-induced transcriptional antitermination and is subject to catabolite repression (Figure 15, panel A). Tryptophanase (TnaA) expression was reduced to 30% relative to the wild type. *TnaB* encodes the tryptophan transporter and *tnaC* a leader peptide [122]. Tryptophanase catalyses the degradation of tryptophan to indole, pyruvate and ammonia. Tryptophan is present in excess in the medium that was used in this study therefore tryptophanase synthesis is subject to antitermination in both strains. In the wild type *tnaCAB* is activated by cAMP-CAP but not in the enzyme I knockout mutant were *tnaCAB* is under catabolite repression. We therefore explain the reduced tryptophanase expression in the mutant by catabolite repression.

Phosphoglycerate kinase (Pgk) is an enzyme of the glycolysis pathway. It converts 3-phospho-D-glyceroyl phosphate into 3-phospho-D-glycerate. In the same reaction one molecule of ATP is formed from ADP. In the mutant Pgk expression was reduced to 40% of the wild type. Phosphoglycerate kinase is encoded together with erythrose 4-phosphate dehydrogenase (*epd)* (Figure 15, panel B) in a transcriptional unit that is under catabolite controlled activation by the cAMP-CAP complex and repressed by Cra (formerly FruR) in presence of fructose-1-phosphate. As there is no fructose present in the medium we conclude that the small reduction of Pgk expression we detect in the mutant is due to catabolite repression only. Since no other sugars are present glycolysis is not expected to be active, and down regulation of a glycolytic enzyme is compatible with the metabolism. A basal activity is however necessary for glucogneogensis, for instance for the synthesis of precursors of cell wall (peptidoglycan) biosynthesis. That enzymes involved in catabolism of sugars are under catabolite control, has been observed before. The reason for this is not clear but it may be a fine tuning reaction to prevent an overshoot of carbohydrate uptake.

Dihydrolipoamide acetyltransferase (AceF) is a subunit of the pyruvate dehydrogenase complex which catalyses the oxidative decarboxylation of pyruvate to Acetyl-CoA. This reaction links glycolysis with the TCA cycle. AceF is part of an operon which also contains the two other subunits of the pyruvate dehydrogenase complex, AceE (*aceE*), E3 (*lpd*), and in addition a gene for the PdhR transcription regulator. AceF was reduced to 40% in the enzyme I mutant. There are two cAMP-CAP binding sites by which *aceF* expression is activated in the absence of glucose (Figure 15, panel C). AceF expression is reduced due to catabolite repression. We do not expect a repression by PdhR, which binds to DNA in the absence of pyruvate. Although no pyruvate is formed from carbohydrates a significant amount could be formed during catabolism of alanine, cysteine, thyrosine, serine and tryptophan which are present in the growth medium.

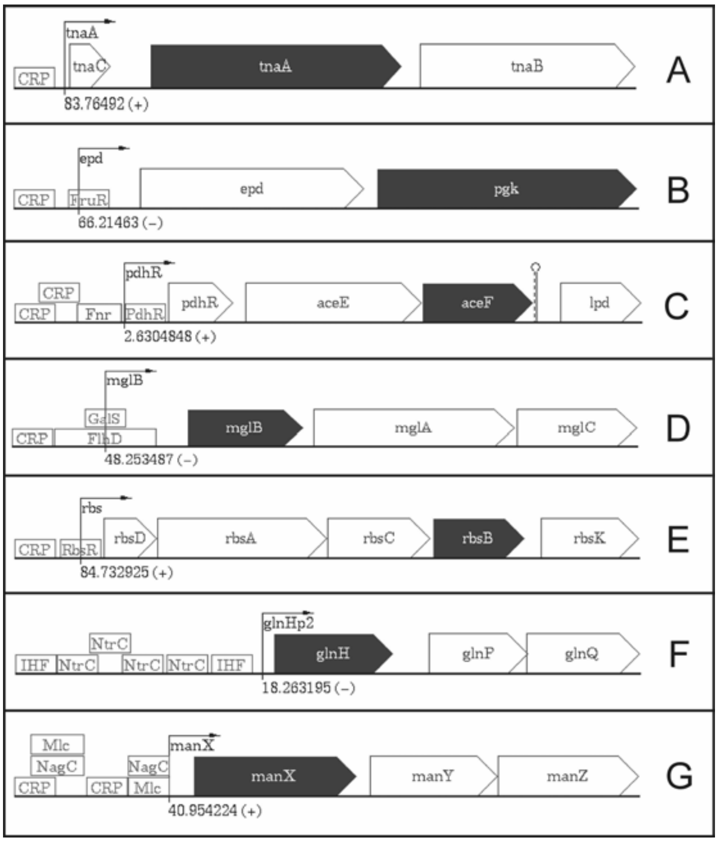
tnaA
CRP
tnaC
tnaA
tnaB
A
83.76492 (+)
epd
CRP
FruR
epd
pgk
B
66.21463 (−)
pdhR
CRP
CRP
Fnr
PdhR
pdhR
aceE
aceF
lpd
C
2.6304848 (+)
mglB
GalS
CRP
FlhD
mglB
mglA
mglC
D
48.253487 (−)
rbs
CRP
RbsR
rbsD
rbsA
rbsC
rbsB
rbsK
E
84.732925 (+)
glnHp2
NtrC
IHF
NtrC
NtrC
NtrC
IHF
glnH
glnP
glnQ
F
18.263195 (−)
Mlc
NagC
CRP
manX
NagC
CRP
Mlc
manX
manY
manZ
G
40.954224 (+)

Figure 15. Operons under cAMP-CAP control.

3.11.2. Controlled expression of genes which are not under catabolite repression. The glutamine binding periplasmic protein (GlnH) was strongly down regulated in the enzyme I knockout strain (Figure 16). GlnH is part of the GlnHPQ high-affinity glutamine transport system which is a member of the ATP-Binding cassette (ABC) super family of transporters. The gene *glnH* is part of the *glnHPQ* operon. The operon belongs to the nitrogen regulated operons and does not have a consensus cAMP-CAP binding site in its promoter region (Figure 15, panel F). A similar observation was reported before by Eppler [123]. The presence of glycerol or glycerol-3-phosphte induced catabolite repression in *E.coli* grown in tryptic broth. And like in our study GlnH was repressed in the presence of glycerol-3-phosphate. Eppler showed that glycerol-3-phosphate only weakly dephosphorylated IIA^{Glc} suggesting that GlnH repression is regulated indirectly by a cAMP-CAP independent regulator [123]. Therefore our result gives additional evidence that the *glnHPQ* operon might be in deed under the control of a cAMP-CAP independent regulator.

YeaD is a gene which was up regulated 5.3 times. *YeaD* is located 84 base pairs downtream of the glyceraldehydes-3-phosphate dehydrogenase gene *gapA* and there is no evidence for a CAP binding consensus sequence in its promoter region. YeaD is a hypothetical protein that was discovered by Fountoulakis in a proteomic screen for low abundant proteins of *E.coli* [124]. YeaD displays sequence similarity with aldose-1-epimerase of *Salmonella typhimyrium*. Aldose-1-epimerase converts alpha-aldoses to their beta-anomer. It is active on D-glucose, L-arabinose, D-xylose, D-galactose, maltose and lactose. The reaction links the metabolism of lactose and galactose and in this sense is a component of carbohydrate catabolism. It must be recalled here, that the apparent up regulation of *yeaD* could not be validated in a reporter gene assay because only background β-galactosidase activity could be detected in mutant and wild-type background.

3.11.3. Constant expression of genes expected to be under catabolite repression. D-galactose-binding periplasmic protein (MglB) and the D-ribose binding periplasmic protein (RbsB) were used as landmarks because they had similar spot intensities in both gel sets. Both proteins are encoded by operons which are known to be under catabolite repression.

The gene of MglB (*mglB*) belongs to *mglBAC* operon which encodes the beta-methylgalactoside transport system, a member of the ATP-binding cassette (ABC) superfamily of transporters. *MglB* encodes the galactose-binding component *mglC* encodes the integral membrane component, and *mglA* the ATP-binding component of the ABC transporter. The *mgl* operon is negatively controlled by the transcriptional repressor GalS. A cAMP-CAP binding site is located upstream of the transcription start (Figure 15, panel D). Expression of *mglB* is unchanged because there is no galactose present in the medium which would be necessary for induction in the wild-type.

The second identified protein from this small subset is the D-ribose binding periplasmic protein. RbsB is encoded by *rbsB* that belongs to the *rbsABC* operon which is cAMP-CAP sensitive expressed and under control of the repressor protein RbsR (Figure 15, panel E). Transcription is induced is when cAMP-CAP and ribose are present which is not the case in our experiment which is the reason that the expression level remains constant.

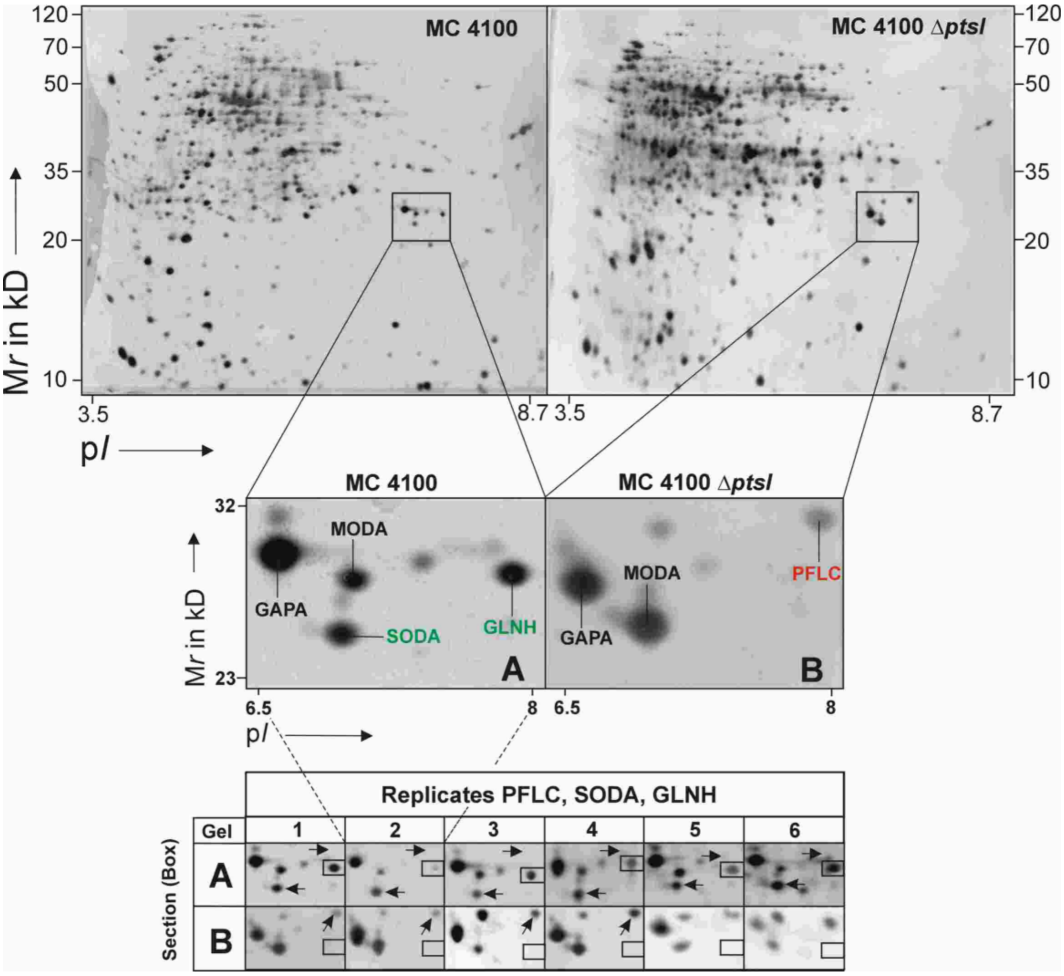

Figure 16. Two 2-D electropherograms of MC4100 (left) and MC4100Δ*ptsI* (right) displaying proteins from p*I* 3.5 – 8.7 and masses 10 – 120 kD. The two enboxed regions (p*I* 6.5 – 8 and masses 23 – 32 kD) in the gel are shown enlarged below. Mass spectrometric identified proteins are labelled with their gene name. Protein expression regulation is indicated by the colour of the label. Green labels (SodA, GlnH) stand for superoxide dismutase and glutamine binding periplasmic protein that both are strong repressed and are absent in MC4100Δ*ptsI*. The red label shows the position of pyruvate formate lyase activating enzyme which shown shows 31 time increased protein expression. Below the enlarged sections, the corresponding regions of each gel, from both gel sets is shown. Letter A indicates replicates of MC4100, letter B for replicates of

MC4100Δpts. The vanishing spots of SodA and strong up regulated PflC are indicated by arrows, vanishing GlnH is enboxed.

3.11.4. The effect of EI on glycolytic and tricarboxylic acid cycle enzymes.

Deleting *ptsI* has a strong impact on enzymes of glycolysis and the TCA cycle. Eight out of the thirty identified proteins belong to these two pathways.

Glyceraldehyde-3-phosphat dehydrogenase A (GapA), phosphoglycerate kinase (Pgk) and enolase (2-phosphoglycerate dehydratase, Eno) were down regulated in the *ptsI* mutant, pyruvate kinase I (Pykf) was up regulated. The genes for these proteins are encoded in four different operons.

G*apA* is transcribed from at least four promoters. Three are well identified: P1 and P3 are σ^{70} dependant transcribed, P2 is transcribed by the heat shock polymerase σ^{32} and P3 is regulated by catabolic repression [125]. Induced transcription from P1 was observed during growth on glucose but only in presence of $IICB^{Glc}$ [125]. Based on this Charpentier and Branlant proposed regulatory mechanism similar to a two component system with $IICB^{Glc}$ acting as sensor (kinase) and an unknown regulatory protein as the response regulator acting on P1 of *gapA* [125]. The $IICB^{Glc}$/Glc dependent activation of *gapA* in addition also is cAMP-CAP dependent. The $IICB^{Glc}$/Glc dependent activation could for instance be mediated by the repressor protein Mlc, which is sequestered from the operator when $IICB^{Glc}$ is dephosphorylated by glucose. However, if this mechanism were operative, *gapA* should have been up regulated (like *manX*, see below) in the *ptsI* mutant grown in amino acids medium. But *gapA* was down regulated under these conditions suggesting that the cAMP-CAP dependent promoter P3 is in charge of control under these conditions and not P1.

Phosphoglycerate kinase (Pgk) is the second glycolytic gene which is down regulated in the *ptsI* mutant. Pgk and GapA share diphosphoglycerate as substrate, and a direct physical interaction between GapA and Pgk is not unlikely. Nevertheless, Pgk is encoded independently of *gapA* in a biscistronic operon with *epd (gapB)*, the gene for erythrose-4-phosphate dehydrogenase. Epd and GapA have 40 % sequence identity. The *epd pgk* operon is activated in a $IICB^{Glc}$/Glc dependent reaction like *gapA* [125]. For full activation $IICB^{Glc}$ and Glc and cAMP-CAP are required. There is no glucose in the amino acid medium, and *pgk* expression should be low. If in addition cAMP is low, as in the *ptsI* mutant, *pgk* may be repressed even further, explaining the observed down regulation.

The third enzyme, enolase (2-phosphoglycerate dehydratase, Eno). was repressed to 10% wild-type. Enolase converts 2-phospho-D-glycerate to phosphoenolpyruvate, the fourth step in second phase of glycolysis. Enolase is encoded in a bicistronic operon together with a gene of unknown function. There are no upstream binding sites for regulatory proteins known.

2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (GpmA) is a fourth glycolytic enzyme which was picked up in the twin proteome comparison. It catalyses the conversion of 3-phosphoglycerate to 2-phosphoglycerate. It shares one of its substrates, 3-phosphoglycerate, with Pgk and the other, 2-phosphoglycerate, with enolase. Physical interactions between Pgk and GpmA and between GpmA and enolase are not unlikely. Although the extent of down regulation of GpmA is small and statistically not significant, the correlation with its neighbors is noteworthy. It underscores the impression that a *ptsI* mutation results in an overall down regulation of glycolytic activity.

The exception to this observation is pyruvate kinase I (Pykf). Pkf also is a glycolytic enzyme but its expression was up regulated 10.5 fold in the *ptsI* mutant. Pykf catalyses the interconversion of phosphoenolpyruvate to pyruvate, the final step in glycolysis. In this substrate level phosphorylation one molecule of ATP is generated per pyruvate. The *pykf* gene is subject to Cra dependent stimulation. As there is no fructose in the medium we do not expect a stimulation of *pykF*. In cells growing in an amino acids medium pyruvate is formed as an intermediate in the amino acid catabolism. Only small amounts of PEP are formed from oxalacetate for gluconeogenesis. We therefore expect that Pfk activity would be repressed to avoid a futile cycle of ATP hydrolysis. Why do we observe the opposite? The activity of Pfk controls the distribution of PEP between the two pathways of substrate level phosphorylation and PTS dependent carbohydrate uptake. The higher the Pkf activity the lower the PEP concentration and this in turn can slow down PTS activity and PTS-sugar uptake. It is conceivable, that increasing Pfk activity feed back inhibits carbohydrate uptake if carbohydrate supply is abundant. If uptake activity were not controlled by restriction of the PEP supply, the cell could be overloaded with sugar-phosphates which are highly toxic. In the *ptsI* mutant all PTS proteins are desphosphorylated and thus in the same state as during vigorous PTS-dependent sugar uptake. Seen from this perspective, up regulation of Pgk could be interpreted as “misinterpretation” of a signal from the PTS.

Pyruvate dehydrogenase connects glycolysis with the TCA cycle. Of the three gene products, E1, E2 and E3, the E2 subunit (AceF) was down regulated in the *ptsI* mutant. The gene is controlled by glucose and oxygen availability. The promoter region contains CAP binding sites, and down regulation can thus be explained by *ptsI* induced catabolite repression. The possibility, that AceF is down regulated as a consequence of insufficient oxygenation will be discussed below in the context of the up regulation of pyruvate lyase related functions. Pyruvate lyase and pyruvate dehydrogenase are two alternative pathways of pyruvate catabolism.

Unlike the glycolytic enzymes which – with the exception of Pgk - were down regulated, the three enzymes which were picked up from the TCA cycle were up regulated in the *ptsI* mutant. Because cells were grown aerobically in an amino acid medium, it is likely, that the amino acids were catabolised by respiration and oxidative phosphorylation. The expression of TCA cycle enzymes therefore should be increased. Succinate dehydrogenase (Sdha) and succinyl-CoA synthetase (SucD) are encoded in a complex operon of 10 genes which are cotranscribed on a 10 kb transcript from the *sdh* promoter [126]. An additional, internal promoter is in front of the *suc* gene cluster. The *sdhP* promoter is the dominant one. It is repressed by glucose and in the absence of oxygen [127]. But this control was shown to be independent of the *crp* (CAP) gene product. The internal *sucP* is controlled differently. It is derepressed in a *crp* mutant and does not respond to glucose as shown with reproter gene assays by Cunningham and Guest [126]. Up regulation of *sucA* would agree with their finding, and up regulation of *sdhA* does at least not contradict the results of Park and Gunsalus [127]. Aconitase B (AcnB) was 4 times more abundant in the mutant than in the wild type. AcnB is a citric acid enzyme that interconverts citrate to cis-aconitate and water. In addition to the catalytic activity of the holo-protein, the apo-protein serves as repressor of superoxide dismutase (SodA) translation based on the oxidative disassembly the Acn iron-sulphur clusters [128]. Our results agree with the findings of Guest as we were not able to detect the *sodA* gene product in the mutant strain with elevated AcnB level (Figure 16).

3.11.5. The effect of EI on anaerobiosis-responsive enzymes. Pyruvate is the endproduct of glycolysis. It is also produced as intermediate in the catabolism of amino acids. There exist several pathways for how pyruvate can further be metabolized. In the presence of a terminal oxygen acceptor such as oxygen (aerobic conditions) or nitrate, pyruvate is oxidatively decarboxylated by pyruvate dehydrogenase (see above) and acetyl-CoA is then oxidized in the TCA cycle. In the absence of an external electron acceptor, pyruvate can either be reduced to lactate (homolactic fermentation) or cleaved into formic acid and acetyl-CoA by pyruvate formate lyase. Formic acid is secreted. Acetyl-CoA is either utilized to support ATP production in a substrate level phosphorylation reaction or it is reduced to ethanol. Two proteins of the pyruvate formate lyase pathway are up regulated in the *ptsI* mutant. The pyruvate formate lyase activating enzyme 2 (PflC) is 31 times increased (Figure 16). YfiD, a homologue of pyruvate formate lyase 2 (PflD) is increased two times.

YfiD has a predicted mass of 14’284 Da and a predicted p*I* of 5.1. It exhibits a strong amino acid sequence homology (50% identity over a 64 amino acid overlap) to the C-terminal region of PflD and PflB. PflB is the active pyruvate formate lyase 1, PflD has 21% sequence identity with PflB [146]. The identity is strict for the C-terminal a 22mer containing the radical-bearing Gly 734. At first sight the observed mass of 10700 Dalton does not correspond well to the theoretical mass while the observed isoelectric point 4.70 is close to the theoretically predicted (5.1) (Figure 17). Wyborn observed on Western blots, a second cross-reacting species YfiD’ of higher mobility and thus lower mass of 11000 Da [129]. The estimated mass of this second species corresponds well with the observation in this study. The reason why YfiD is detected at too low mass in our experiment can be explained by the participation of YfiD in the reaction of pyruvate formate lyase (Pfl).

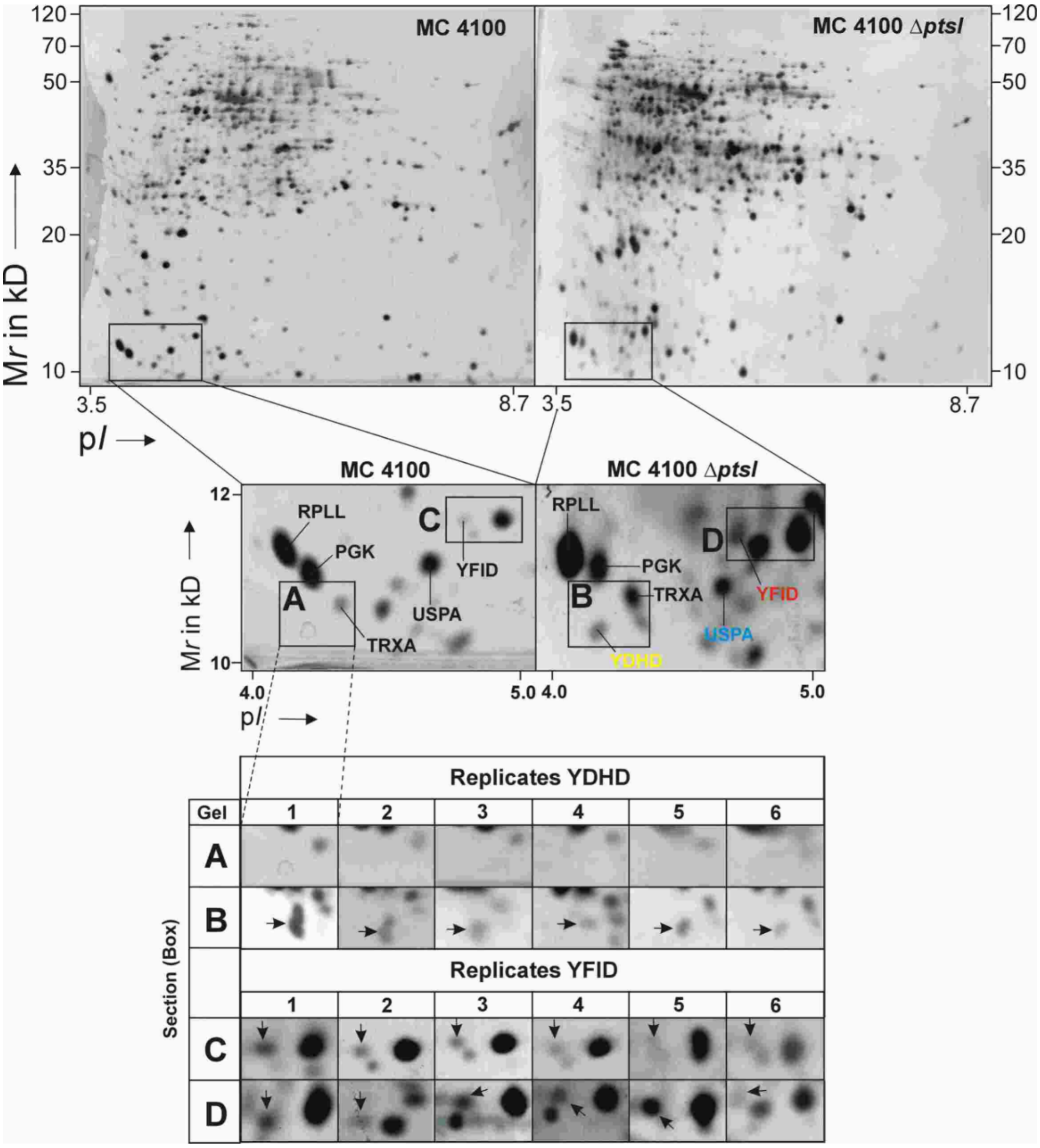
MC 4100
MC 4100 Δptsl
120
70
50
35
20
10
Mr in kD
3.5
8.7
pI
12
RPLL
PGK
C
YFID
USPA
A
TRXA
D
B
YDHD
4.0
5.0
Replicates YDHD
Gel
1
2
3
4
5
6
Replicates YFID
Section (Box)

Figure 17. Two 2-D electropherograms of MC4100 (left) and MC4100Δ*ptsI* (right) displaying proteins from p*I* 3.5 – 8.7 and masses 10–120 kD. The two enboxed regions (p*I* 4.0 – 5.0 and masses 10 – 12 kD) in the gel are shown enlarged below. Mass spectrometric identified proteins are labelled with the gene name. Protein expression regulation is indicated by the colour of the label. The red label shows the position of pyruvate formate lyase activating enzyme homologue YfiD which is 2 time more abundant in the mutant than in the wild type. The yellow label indicates the spot position of YdhD which appears new in MC4100Δ*ptsI*. The blue label indicates the position of the universal stress protein which shows a decreased expression in MC4100Δ*ptsI.* Below the enlarged sections, the corresponding regions of each gel from both gel sets, are shown. Letter A and C indicate replicates of MC4100, letter B and D for replicates of MC4100Δ*ptsI*. The new appearing spot of YdhD is indicated by an arrow (row B) as well as the strong unregulated spot of YfiD (row D).

Pfl (PflB(1), PflD(2), PflF(3)) catalyse the non-oxidative cleavage of pyruvate to acetyl-CoA and formate in anaerobically growing cells. The enzymes are transformed from their inactive form to the active form by pyruvate formate activating enzymes (PflA, PflC, PflE). PflC was identified in our study. The activated pyruvate formate lyase contains a glycine radical which is essential for catalysis. This glycyl radical is highly sensitive to oxygen, Upon exposure to oxygen the Pfl polypeptide backbone is cleaved at N–Cα of Gly734 into a fragment (γPfl) and inactive pyruvate formate lyase Pfl'.

It was shown by Wagner et al. that YfiD can associate with a recombinant protein comprising the core of Pfl (Ser1–Ser733) to form a heterooligomeric Pfl(1-733)::YfiD complex. This complex can be activated by the Pfl activase and then has full catalytic activity. It is proposed that YfiD acts as "spare part" for Pfl's glycyl radical domain, and that YfiD allows rapid recovery of Pfl activity (and thus ATP generation) in cells that have experienced oxidative stress (Figure 18) [130].

If the Pfl(1-733)::YfiD complex is activated by the PflC activase and then exposed to oxygen, the YfiD subunit is cleaved into an 11 kD (YfiD') and a 3 kD (YfiD'') polypeptide. And this 11 kD polypeptide was picked up in this twin proteome analysis.

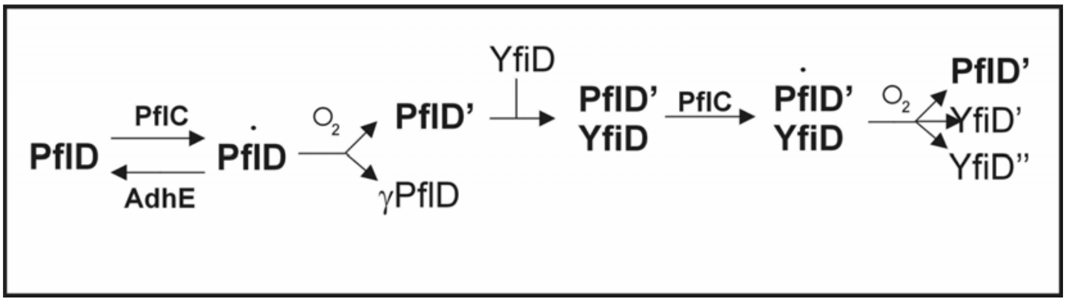

Figure 18. Scheme of Pfl (PflD) activation by PflC and inactivation by oxygen, formation of 3 kD γPflD, association of inactive PflD' to PflD'-YfiD complex with YfiD, activation by PflC and deactivation by oxygen with formation of inactive PflD', 11 kD YfiD' and 3 kD YfiD''.

PflC is one of three Pfl activating enzymes encoded in the *E. coli* genome. *PflC* is in a bicistronic operon together with *pflD*, the gene for a pyruvate formate lyase. This operon is inserted between open reading frames for IICB (*frwCfrwB*) and a IIB (*frwD*) like PTS proteins. Up regulation of PflC could not be confirmed by a reporter assay, probably because *pflC* has no promoter of its own but is transcribed as bicistronic mRNA from the *pflD* promoter. The reason for this dramatic up regulation is therefore unclear. The genes *pflB* and *pflA* for the primary pyruvate formate lyase are essential and they also cannot be complemented by *pflD* and *pflC* 1 [146]. The biological function of this secondary Pfl complex therefore is not clear.

A strong increased expression of the primary pyruvate formate lyase B was observed in *E.coli* cells which were under glycerol-3-phosphate induced catabolite repression and grown aerobically (Eppler [123]). But also in this instance it remains unclear why a highly oxygen sensitive enzyme should be induced under aerobic conditions.

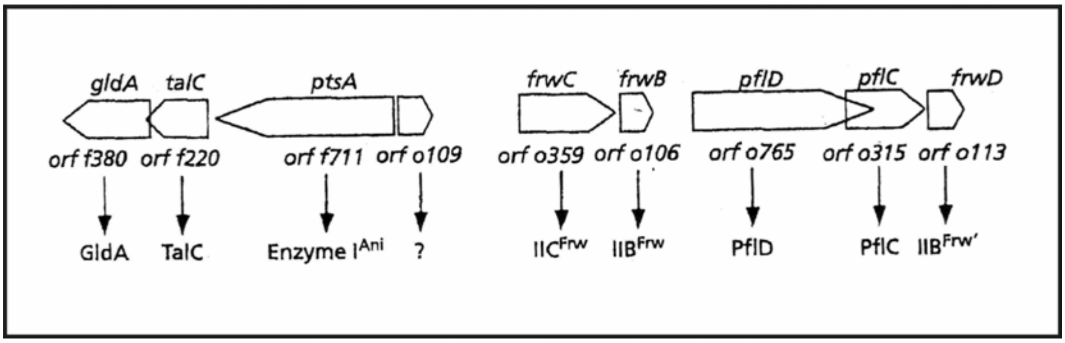
gldA
talC
ptsA
frwC
frwB
pflD
pflC
frwD
orf f380
orf f220
orf f711
orf o109
orf o359
orf o106
orf o765
orf o315
orf o113
GldA
TalC
Enzyme I^Ani
?
IIC^Frw
IIB^Frw
PflD
PflC
IIB^Frw'

Figure 19. *PtsA* gene cluster (the figure was taken from Reizer J., et al., *Microbiology* 1995, 141, 961-971).

3.11.6. The effect of EI on proteins of the PTS**.** Expression of the mannose specific IIAB component was increased by factor 2.4 in the Δ*ptsI* strain. This result is consistent with earlier observations from R. Beutler [131]. According to Plumbridge [132] expression of *manX* ($IIAB^{Man}$) is under the control of Mlc, a repressor protein which is itself controlled by $IICB^{Glc}$. The dephosphorylated form of $IICB^{Glc}$ binds and thereby sequesters Mlc from its operators. In the Δ*ptsI* strain $IICB^{Glc}$ is permanently dephosphorylated and Mlc remains sequestered. Therefore the *man* operon is expressed constitutively which leads to the increased protein expression that we were able to measure.

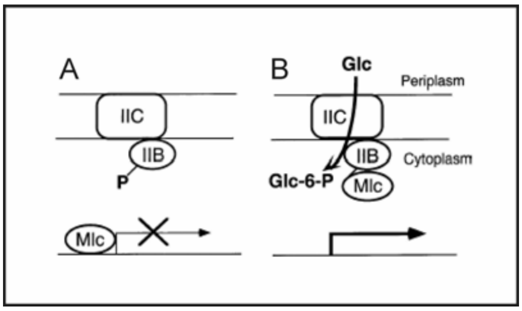
A
IIC
IIB
P
Mlc
B
Glc
Periplasm
IIC
IIB
Cytoplasm
Glc-6-P
Mlc

Figure 20. Operon unter Mlc repression. The figure was taken from J. Plumbridge, *Microbiology* 2000, 146, 2655–2663.

Also 3.3 fold increased in the *ptsI* mutant is the expression of DhaL (YcgS), a subunit of the dihydroxyacetone kinase. DhaL has previously been identified in 2-D gels as strongly up regulated by Beutler [131]. DhaL is encoded in the *dhaKLM* operon. The dephosphorylated form of DhaL functions as coactivator of *dhaKLM* expression. The phosphorylated form of DhaL is inactive. DhaL is phosphorylated by the PTS protein DhaM which in turn is phosphorylated by enzyme I. In a *ptsI* mutant DhaL cannot be inactivated and gene expression becomes constitutively high [133]. Detailed studies on the transcriptional regulation of dihydroxyacetone kinase are done by C. Bächler.

The third PTS protein up regulated in the *ptsI* mutant is GlvB, a IIB component of a PTS permease (GlvCB) of unknown specificity but with sequence similarity to the arbutin specific permease and also to the B-domain of $\mathrm{IICB}^{\mathrm{Glc}}$. GlvB is encoded in *glvCBG*, reported to be a cryptic operon by Reizer in 1994 [134, 135] (Figure 21). The appearance of GlvB indicates that the *glv* operon is not crytptic. The *glv* operon encodes a putative PTS permease with detached enzymes IIB and IIC but no enzyme IIA. If GlvB and GlvC are expressed and have a function in the uptake of carbohydrates they might work in conjunction with a IIA protein such as $\mathrm{IIA}^{\mathrm{Glc}}$.

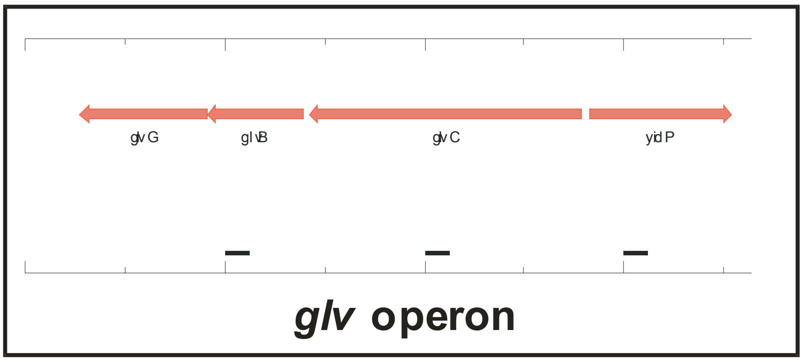

Figure 21. The *glv* operon, including pesudogene *glvG*, cryptic gene *glvC*, hypothetical open reading frame *yidP* and the gene *glvG* of the newly discovered protein GlvB.

3.11.7. The effect of EI on stress proteins. Changes in environment, the genome and growth conditions place *E.coli* cells under stress. *E.coli* has a cellular stress response to cope with many different forms of stress. Heat shock is the best known inducer of the stress response but the stress response can also be activated by many other conditions or agents. When not responding to stress induced damage, most stress proteins also perform important "housekeeping" functions. For example, many stress proteins are molecular chaperones that help new synthesized polypeptides assume their proper conformation. Another physiological stress is the oxidative stress which builds up when cells are respiring and thereby producing reactive oxygen as inevitable side reaction.

In this study a number of stress inducible proteins were identified. The heat shock inducible periplasmic serine protease (HtrA), DnaK (protein chaperone) and the universal stress protein A (UspA), and several oxidative stress response proteins were down regulated in the *ptsI* mutant. Chaperon GroEL, from the Hsp 60 family was induced.

3.11.7.1. The effect of EI on "general" stress proteins. UspA is a member of all starvation and stress stimulons so far studied in *E. coli* [136]. It is up regulated under carbon starvation conditions. Wild-type and *ptsI* mutant were exponentially growing at the time of cell harvesting and therefore not under carbon starvation. It is not clear why UspA should be down regulated in the *ptsI* mutant. But it could be speculated that a strongly dephosphorylated PTS is a sign of vigorous carbohydrate uptake, and hence the opposite of starvation. In the *ptsI* mutant the PTS is in the same state and thus signals carbohydrate excess and hence down regulation.

The cold shock protein YfiA was only detected in the wild type. This protein was first identified by Fountoulakis on a 2-D electropherogram of total *E.coli* protein [124]. YfiA binds to the small ribosomal subunit and stabilizes ribosomes against dissociation when bacteria experience environmental stress [137]. We have no explanation why this protein is under repression in the mutant.

3.11.7.2. The effect of EI on proteins related to oxidative stress. Four proteins with significantly different expression levels were related to oxidative stress. Superoxide dismutase and thiol peroxidase were only detectable in the wild-type, while the glutaredoxin-like protein YdhD was present only in the mutant. The thioredoxin like protein YhgI was three times increased in the mutant. No change of TrxA abundance was seen between wild type and enzyme I knockout mutant. TrxA was identified in a landmark spot.

SodA and Tpx belong to the thioredoxin and glutaredoxin system of *E.coli*. YdhD has homology to glutaredoxin GrxC. Of YghI it is not known whether it belongs to the thioredoxin and glutaredoxin system.

The so far known eleven constituents of the thioredoxin and glutaredoxin system are organized in three branches (Figure 22). The thioredoxin branch, the glutaredoxin branch and the thiolperoxidase branch.

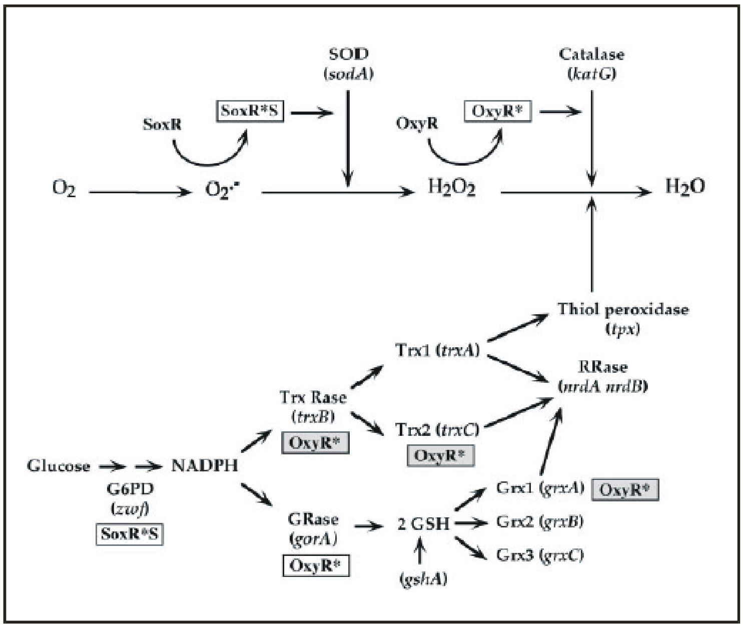
SOD
(sodA)
Catalase
(katG)
SoxR
SoxR*S
OxyR
OxyR*
O2
O2·⁻
H2O2
H2O
Thiol peroxidase
(tpx)
Trx1 (trxA)
RRase
(nrdA nrdB)
Trx Rase
(trxB)
OxyR*
Trx2 (trxC)
OxyR*
Glucose
NADPH
G6PD
(zwf)
SoxR*S
GRase
(gorA)
OxyR*
2 GSH
(gshA)
Grx1 (grxA)
OxyR*
Grx2 (grxB)
Grx3 (grxC)

Figure 22. The thioredoxin and glutaredoxin system of *E.coli*. Image taken from C. Pueyoet al., *J Biol Chem* 2000, 275 (18) 13398–13405.

Tpx reduces H_2O_2 and alkyl hydroperoxides with electrons provided by thioredoxin (Trx), Trx reductase, and NADPH [138]. Tpx and SodA act along the oxygen detoxification pathway to prevent the harmful effects of reactive oxygen species that are produced by many physiological processes (e.g., incomplete reduction of molecular oxygen during aerobic respiration).

The gene product of the open reading frame *yhgI* is identified for the first time in our study. A sequence alignment of the predicted open reading frame shows strong homology to a thioredoxin-like protein in *Salmonella typhimyrium*. The function of YhgI is unknown. The gene *yhgI* has no known upstream binding site for regulatory proteins. 3-D structure for YhgI shows that it has an N-terminal HesB and a C-terminal NifU domain. The HesB domain is also found in IscN a protein from Rhizobium etli [139]. Closer inspection of the YhgI HesB domain revealed that YhgI has a similar fold like the Hes B domain of IscA (b2528).

A second iron sulphur cluster (isc) protein, the NifU-like protein of *E. coli* was reduced to 20% in the mutant. NifU (IscU, P77310, B2529) belongs to the *iscSUA* operon, The *iscSUA* operon plays a role in assembly of iron-sulfur clusters in *E.* coli. Cysteine desulfurase (IscS) interacts with IscU and transfers sulfur directly to IscU [143]. Transcription of the *isc* operon is repressed by the

IscR protein, no additional binding sites for regulatory protein are known. YfhF (IscA, B2528, P36539) which is located downstream of NifU belongs to the HesB family. It therefore appears that YhgI is a two-domain version of the NifU and YfhF subunits and that the three proteins have a related function

The gene *b1654* encoding for the glutaredoxin like protein (YdhD) was first mentioned by Bacher in 1995 [140]. *b1654* was found downstream of the gene encoding for the long helicase related protein (*lhr)*, the longest known gene in *E.coli* [140]. YdhD was discovered by Sanchez in 2001 [141] during a proteomic screen. Sequence alignment of the theoretical gene product revealed a similarity to glutaredoxin GrxC. In figure 23 the new appearing protein spot in 2-D gels from cell extracts of the enzyme I lacking mutant is showed in row B. There are no known upstream binding sites for sites for regulatory proteins.

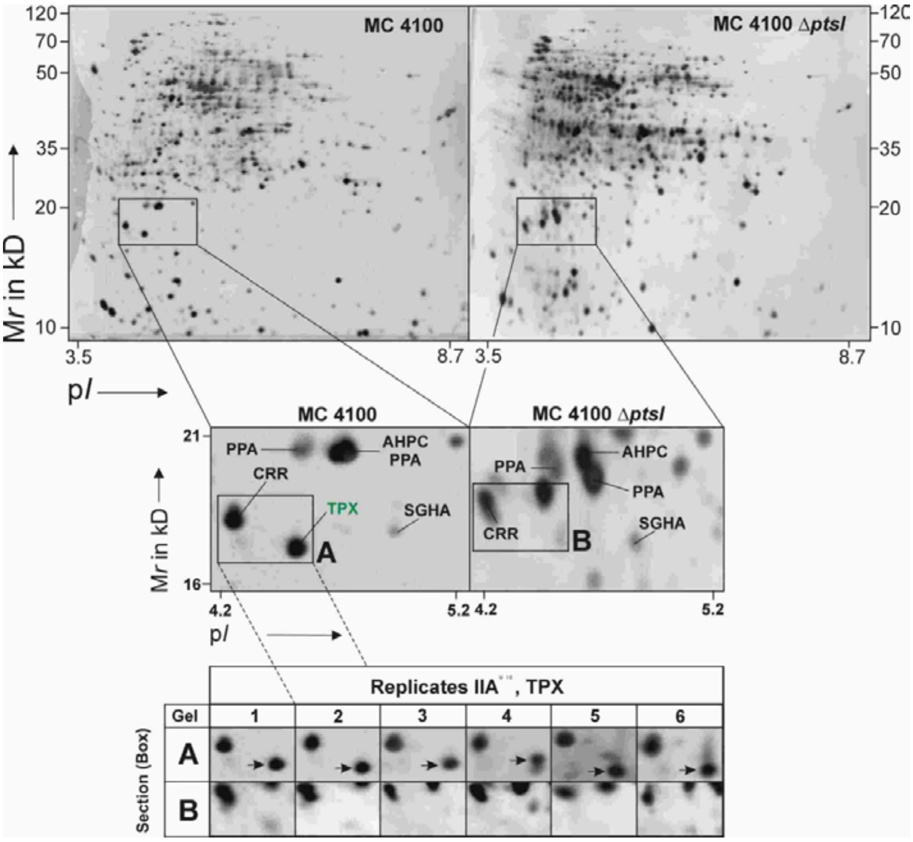

Figure 23. Two 2-D electropherograms of MC4100 (left) and MC4100Δ*ptsI* (right) displaying proteins from p*I* 3.5 – 8.7 and masses 10 – 120 kD. The two enboxed regions (p*I* 4.2 – 5.2 and masses 16 – 21 kD) in the gel are shown enlarged below. Mass spectrometric identified proteins are labelled with their gene names. Protein expression regulation is indicated by the colour of the label. Green labels (Tpx) for thiol peroxidase that is absent in MC4100Δ*ptsI*. The two rows A and B below contain the corresponding regions of each gel from both gel sets enlarged. Letter A indicates replicates of MC4100, letter B for replicates of MC4100Δ*ptsI*. Tpx in wild type gels is indicated by an arrow.

3.11.8. The effect of EI on proteins related to information transfer. The transcription antitermination protein NusG is absent in the enzyme I knockout mutant. NusG is a transcriptional elongation factor in *E.coli* that aids transcriptional antitermination by the phage lambda N protein [142]. There are no known binding sites for regulatory protein upstream of this operon. We do not know why the protein is not expressed in the mutant strain.

The gene product of *b067* is the hypothetical protein YbhE*.* Until now it was not clear if this protein is expressed by the cell. Now this gene product is reported as identified for the first time in this work. YbhE could only be detected in the wild type strain but not in the mutant. Sequence comparison shows that this protein exhibits 95% homology to a putative isomerase from *Shigella flexneri* and a 26% homology to muconate cycloisomerase I of *Trichosporon cutaneum* (Yeast). The gene *ybhE* is not part of an operon and has no known binding sites for regulatory proteins. Also for this protein we no explanation what it is not expressed in the mutant.

4. References

1. Postma P., Lengeler J., Jacobson G., *Microbiological Reviews* 1993, 57 (3), 543-594.
2. Siebold C., Flükiger K., Beutler R., Erni B., *FEBS Letters* 2001, 1-8.
3. Erni B., *Bacterial Transport Systems*, Viley-VCH Verlags GmbH, Weinheim, Chapter 5, 115-138, 2001.
4. Amster-Choder O., Houman F., Wright A*., Cell* 1989, 58(5), 847-55.
5. Reizer J., Reizer A., Saier J., *Protein Science* 1992, 1, 722-726.
6. Reizer J., Reizer A., Merrick M., Plunkett Saier M., *Gene* 1996, 181(1-2), 103-8.
7. Reizer J., Michotey V., Reizer A., Saier M., *Protein Sci* 1994, 3(3), 440-50.
8. Reizer J., Reizer A., Saier M., *Microbiology* 1997, 143 (Pt 8), 2519-20.
9. Hall B., Xu L., *Mol Biol Evol* 1992, 9(4), 688-706.
10. Feldheim D., Chin A., Nierva C., Feucht B., et al., *J Bacteriol* 1990, 172(9), 5459-69.
11. Ramseier T., Neger D., Cortay J., Scarabel M., et al., *J Mol Biol* 1993, 234(1), 28-44.
12. Schnetz K., Toloczyki C., Rak B., *J Bacteriol* 1987, 169 (6), 2579-90.
13. Schnetz K., Rak B., *Proc Natl Acad Sci U S A* 1992, 89(4), 1244-8.
14. Vick J., Von Bredow J., *J Appl Toxicol* 1996, 16(6), 509-16.
15. Reizer J., Mitchell W., Minton N., Brehm J., et al., *Curr Microbiol* 1996, 33(5), 331-3.
16. Yamada M., Saier M., *J Mol Biol* 1988, 203(3), 569-83.
17. Buhr A., Erni B., *J Biol Chem* 1993, 268(16), 11599-603.
18. Chen Y., Reizer J., Saier M., Fairbrother W., et al., *Biochemistry* 1993, 32(1), 32-7.
19. Eberstadt M., Grdadolnik S., Buhr A., Erni B., et al., *Biochemistry* 1996, 35(35), 11286-92.
20. Liao D., Kapadia G., Reddy P., Saier M, et al.,*Biochemistry* 1991, 30(40), 9583-94.
21. Reidl J., Boos W., *J Bacteriol* 1991, 173(15), 4862-76.
22. Figge R., Ramseier T., Saier M., *J Bacteriol* 1994, 176(3), 840-7.
23. Kroon G., Grotzinger J., Dijkstra K., Robillard G., et al., *Protein Sci* 1993, 2(8), 1331-41.
24. Lee C., Saier M., *J Biol Chem* 1983, 258(17), 10761-7.
25. Sugiyama J., Mahmoodian S., Jacobson G., *Proc Natl Acad Sci U S A* 1991, 88(21), 9603-7.
26. Sprenger G., *Biochim Biophys Acta* 1993, 1158(1), 103-6.
27. Huber F., Erni B., *Eur J Biochem* 1996, 239(3), 810-7.
28. Nunn R., Markovic-Housley Z., Genovesio-Taverne J., Flükiger K., *J Mol Biol* 1996, 259(3), 502-11.
29. Plumbridge J., *Mol Microbiol* 1998, 27(2), 369-80.
30. Seip S., Lanz R., Gutknecht R., Flükiger K., et al., *Eur J Biochem* 1997, 243 (1-2), 306-14.
31. Stolz B., Huber M., Markovic-Housley Z., Erni B., *J Biol Chem* 1993, 268 (36), 27094-9.
32. Keyhani N., Roseman S., *Proc Natl Acad Sci U S A* 1997, 94(26),14367-71.
33. Parker L., Hall B., *Genetics* 1990, 124(3), 455-71.
34. Parker L., Hall B., *Genetics* 1990, 124(3), 473-82.
35. Reizer J., Reizer A., Saier M., *Res Microbiol* 1990, 141(9), 1061-7.
36. van Montfort R., Pijning T., Kalk K., Reizer J., et al., *Structure* 1997, 5(2), 217-25.
37. Reizer J., Ramseier T., Reizer A., Charbit A., et al., *Microbiology*, 1996,142 (Pt 2), 231-50.
38. Peri K., Waygood E., *Biochemistry* 1988, 27(16), 6054-61.
39. Plumbridge J., *Mol Microbiol* 1989, 3(4), 505-15.
40. Plumbridge J., *Mol Microbiol* 1991, 5(8), 2053-62.
41. Plumbridge J., Kolb A., *J Mol Biol* 1995, 249(5), 890-902.
42. Boos W., Ehmann U., Forkl H., Postma P., et al., *J Bacteriol* 1990, 172(6), 3450-61.
43. Klein W., Horlacher R., Boos W., *J Bacteriol* 1995, 177(14), 4043-52.
44. Marechal L., *Arch Microbiol* 1984, 137(1), 70-3.
45. Rimmele M., Boos W., *J Bacteriol* 1994, 176(18), 5654-64.
46. Nobelmann B., Lengeler J., *Biochim Biophys Acta* 1995, 1262(1), 69- 72.
47. Nobelmann B., Lengeler J., *J Bacteriol* 1996, 178(23), 6790-5.
48. Postma P., Lengeler J., Jacobson G., *Microbiol Rev* 1993, 57(3),543-94.

49. Reizer J., Charbit A., Reizer, A., Saier M., *Genome Sci. Technol*, 11996, 53-75.
50. Reizer J., Reizer A., Saier M., *Microbiology* 1997, 143 (Pt 8), 2519-26.
51. Reizer J., Michotey V., Reizer A., Saier M., *Protein Sci* 1994, 3(3), 52.
52. Reizer J., Reizer A., Saier M., *Microbiology* 1995, 141 (Pt4), 961-71.
53. Kirkpatrick C., Maurer L., Oyelakin N., Yoncheva Y., et al., *Bacteriol* 2001,183, (21), 6466-77.
54. Peng L, Shimizu K., *Appl Microbiol Biotechnol* 2003, 61(2),163-78.
55. Lee J., Lee D., Lee B., Kim., *J. Bacteriol* 2003, 185 (18), 5442-5451.
56. Stancik L., Stancik D., Schmidt B., Barnhart D., et al., *J. Bacteriol* 2002,184(15), 4246-58.
57. Han M., Yoon S., Lee Y., *J. Bacteriol* 2001, 171, 301-308.
58. Perrot F., Hebraud M., Charlionet R., Junter A., Jouenne T., *Electrophoresis* 2000, 21, 645-653.
59. Blankenhorn D., Phillips J., Slonczewski J., *J. Bacteriol* 1999, 181, 2209-2216.
60. Wasinger C., Cordwell J., Cerpa-Poljak A., Yan X., et al. *Electrophoresis* 1995, 16, 1090-1094.
61. Westermeier R., Naven T., *Proteomics in Practice*, Wiley-VCH, Verlags GmbH, Weinheim, 2002.
62. Hamdan M., Righetti, G., *Mass Spectrom Reviews* 2002, 21, 287-302.
63. Rose K., GeneProt inc. Geneva, *personal communication* 2001.
64. Righetti G., Castagna A., Hamdan M., *Advances in Chromatography*, Vol. 42, New York, 2003, pp. 270-321.
65. Ikeda K., Suzuki S., *U.S. Patent* 1912, 1’015’891.
66. Tiselius A*., Trans Faraday Sic* 1937, 33, 524-531.
67. Kenrick K., Margolis J., *Anal. Biochem* 1970, 33, 204-207.
68. Laemmli K., *Natur*e 1970, 227, 680-685.
69. O’Farrell H., *J. Biol. Chem* 1975, 250, 4007-4021.
70. Scheele A., *J. Biol. Chem* 1975, 250, 5375-5385.
71. Klose J., *Humangenetik* 1975, 26, 231-243.
72. Bjellqvist B., Righetti P., Giannazza E., et al., *J. Biochem Biophys Methods* 1982, 6, 317-339.
73. Görg A., Postel W., Günther S., Hanash S., et al., *Electrophoresis* 1988, 9, 37-46.
74. Hamdan H., Righetti G., *Mass Spectrom Reviews* 2003, 22, 272-284.
75. Hoving S., Voshol H., van Oostrum J., *Proteomics* 2004, submitted.
76. Hoving S., Voshol H., van Oostrum J., *Electrophoresis* 2000, 21, 2617-2621.
77. Langen H., USGEB Meeting Davos, March 2003.
78. Langen H., Lecture “*The Human Proteome*” University of Berne, May 2003.
79. Tastet C., Luche S., van Dorsselaer A., Rabilloud T., et al., *Electrophoresis* 2003, 11, 1787-94.
80. Rabilloud T., *Electrophoresis* 1998, 19, 758-760.
81. Hartinger J., Stenius K., Hogemann D., Jahn R. *Anal. Biochem* 1996, 240, 126-133.
82. Macfarlane E., *Anal. Biochem* 1989, 176, 457-463.
83. Langen H., Evers S., Berndt P., Fountoulakis M., *Electrophoresis* 2000, 21, 411-429.
84. Braun R., Ueffing M., GSF Research Center, Neuherberg, *personal communication* 2003.
85. Oguri T., Nomura E., Hidaka M., Inagaki N., et al*., Proteomics* 2002, 6, 666-72.
86. Zischka H., Braun R., Ueffing M., Eckerskorn C., et al., *Proteomics* 2003, 3, 906-16.
87. Zischka H., Keller H., Kellermann J., Eckerskorn C., et al., *Proteomics* 2003, 3, 78-86.
88. Lee C., Gupta S., Morozova I., *Analytical Biochemistry* 317, 2003, 271-275.
89. Gygi P., Rist B., Gerber A., Turecek F., *Nat Biotechnol* 1999, 17, 994-999.
90. Cagney C., Emili A., *Nat Biotechnol* 2002, 20, 163-170.
91. Link J., Eng J., Schieltz M., Carmack E., *Nat Biotechnol* 1999, 17, 676-682.
92. McDonald H., Yates R.**,** *Dis Markers* 2002, 2, 99-105.
93. Edman P., Begg G., *Eur J Biochem* 1967,1, 80-90.
94. Langen H., *Habilitationsschrifft* 2000, p.18.
95. Karas M., Hillenkamp F., *Anal Chem* 1988, 20, 2299-2301.
96. Henzel W., et al., *Proc natl Acad Sci USA* 1993, 11, 5011-5015.
97. Fountoulakis, M., Langen H., *Anal Biochem* 1997 250, 153-156.
98. Langen H., Fountoulakis M., Berndt P., Suter L., *Electrophoresis* 2002, 23, 311-328.
99. Berndt P., Hobohm U., Langen H., *Electrophoresis* 20, 1999, 3521-3526.
100. Casadaban M., *J Mol Biol* 1976, 104(3), 557-66.

101. Hesterkamp T., Erni B., *J Mol Microbiol Biotechnol* 1999, 1(2), 309-17.
102. Swain M., Ross W., *Electrophoresis* 1995, 16, 948-951.
103. Berggren K., Chernokalskaya E., Steinberg H., Kemper C., et al. *Electrophoresis* 2000, 21, 2509-2521.
104. Pappin D., Hojrup P., Bleasby A., *Curr. Biol* 1993, 3(6), 327-32.
105. http://us.expasy.org/cgi-bin/map2/def?ECOLI.
106. Bartsch, *Taschenbuch Mathematischer Formeln*, Verlag Harry Deutsch p.534, 9. Auflage, 1986.
107. www.graphpad.com/articles/interpret/principles/power.html.
108. www.graphpad.com/www/book/Choose.html.
109. http//trochim.human.cornell.edu/kb/stat_t.html.
110. Riedwyl, *Angewandte Statistik*. 2. Auflage, Haupt. Bern (1992) 91-131.
111. Bartsch, *Taschenbuch Mathematischer Formeln*, Verlag Harry Deutsch p. 530, 9. Auflage, 1986.
112. http://us.expasy.org/sprot/.
113. http://us.expasy.org/tools/pi_tool.html.
114. http://cib.nig.ac.jp/dda/backup/taitoh/ecoli.operon.html.
115. http://www.ncbi.nlm.nih.gov/genomes/framik.cgi?db=Genome&gi=115.
116. http://www.ecocyc.org.
117. http://uqbar.rockefeller.edu/regulons/ClustersWadsworth/index_new.html.
118. Lamanda A., *Diploma Thesis* 1999.
119. NEB Catlogue 2003, p242-243.
120. Miller F., Hershberger C**.,** *Gene*, 1984 29(1-2), 247-50.
121. Bächler C., University of Berne, *personal communication* 2003.
122. Deeley M., Yanofsky C., *J Bacteriol* 1981, 147(3), 787-96.
123. Eppler T., Postma P., Schütz A., Boos W., et al., *J Bacteriol* 2002, 184(11), 3044-3052.
124. Fountoulakis M., Takacs M., Berndt P., Langen H., *Electrophoresis* 1999 20, 2181-2195.
125. Charpentier B., Bardney V., Robas N., Branlant C., *J Bacteriol* 1998, 180(24), 6476-83.
126. Cunningham L., Guest J., *Microbiology* 1998, 144 (Pt 8), 2113-23.
127. Gunsalus R. and Parker S., *Res Microbiol* 1994, 145(5-6), 437-50.
128. Tanf Y., Quali A., Artymiuk P., Guest J., et al., *Microbiology* 2002, 148, 1027-37.
129. Wyborn N., Messenger S., Green J., et al., *Microbiology* 2002, 148, 1015-1026.
130. Wagner A., Schultz S., Bomke J., Pils T., et al., *Biochem biphys Res Commun*, 2001, 285, 456- 462.
131. Beutler R., *PhD Thesis*, 2000.
132. Plumbridge J., *Molecular Microbiology* 1999, 33 (2), 260-273.
133. De Reuse H., Danchin A., *J Bacteriol* 1991, 173(2), 727-33.
134. Postma P., Lengeler J., Jacobson G. *Microbiol Rev* 1993, 57(3), 543-94.
135. Michotey V., Reizer A., Saier M., *Protein Sci* 1994, 3(3), 440-50.
136. Farewell A., Diez A., DiRusso C., Nystrom T., *J Bacteriol* 1996, 178(22), 6443-50.
137. Agafonov D., Kolb A., Nazimov V., Spirin S., *Proc Natl Acad Sci US* 1999, 96(22), 12345-9.
138. Cha M., Kim H., Kim I., *J Bacteriol* 1996, 178(19), 5610-5610.
139. Huang C., Lin F., Chu K., Chen M., *Microbiology* 1999, 145, 743-753.
140. Bacher N., Koonin E., Rudd K., Deutscher M., *J. Bacteriol* 1995, 177 (19), 5393-5400.
141. Binz P., Appel R., Hochstrasser D., Sanchez J., et al., *Proteomics* 2001, 409-423.
142. Li J., Mason S., Greenblatt J., *Genes Dev* 1993, 7, 161-172.
143. Hoff K., Cupp-Vickery J., Vickery L, *J Biol Chem* 2003, 26, 278(39), 37582-9.
144. Langen H., Röder D., Juranville J., Fountoulakis M., *Electrophoresis* 1997, 18, 2085-90.
145. Marshall F., Messenger S., Wyborn N., Guest J., et al., *Molecular Microbiology* 2001, 39(3), 747-753.
146. Zhu J. and Shimizu K., *Appl Microbiol Biotechnol* 2003, Epub ahead of print.

Part II

Improvement of in-gel protein visualisation

Part III

Appendices

Mein herzlicher Dank geht an:

Gertrud und Andreas Lamanda

Prof. Dr. B. Erni

Prof. Dr. U. Jenal

Franziska Buchmann

Stephan Reber

PD Dr. H. Langen

Außerdem danke ich den folgenden Leuten: (Reihenfolge zufällig)

Daniel Röder, Dr. Stefan Evers, Dr. Michel Fountoulakis, Dr. Peter Berndt von F. Hoffmann-La Roche, Basel, für die vielen guten Ratschläge, die durchgeführte Analytik, den freundlichen Empfang und die Mittagessen. Dr. Gerd Grenner (Chief Technology Officer), für die interessanten Gespräche. Dr. Reiner Westermeier von Amersham Pharmacia Biotech Europe, Freiburg für die interessanten Gespräche. Prof. Keith Rose (Chief Science Officer) GeneProt Inc, Genf, für den freundlichen Empfang, die Gespräche und die Besichtigung von GeneProt. Dr. Jan Van Oostrum, (Unithead of Proteomics), Dr. Hans Voshol, Dr. Sjouke Hoving von Novartis Pharma, Basel, für die interessanten Diskussionen, die Nomination für den Poster Award in München, die vielen Informationen und die Mittagessen. Dr. B. Grünenfelder, Gabriele Rummel vom Biozentrum der Uni Basel für die Unterlagen, den “mini 2-D Gel Kurs“, den freundlichen Empfang. Prof. Dennis Hochstrasser, Maged Saad, Peter Skorpil, Severine Hughes, vom Universitätsspital Genf für die informativen Gespräche die wir führten. Dr. Jan Sanders (European Sales Manager), Nonlinear Dynamics für die freundliche Beratung und die Blitzschulung in Phoretix 2D Advance. Dr. Said Modaressi, Dr. M. Liebler (Vicepresident) und G. Keck, Raytest Isotopen Messgeräte GmbH (Fuji), für die langjährige gute Beratung, die Blitzschulung in AIDA, den Webauftritt und das Interesse an meiner Arbeit. Dr. Johannes Hewel (für die ESI-TOF Analytik), Alain Zahn (für das Synthetisieren von RuBP), Dr. Rudolf Beutler, PD Dr. Johann Schaller (für die Analytik , das Material, die interessanten Gespräche und das Verteilen meiner Flyers in den USA), Urs Kämpfer für die interessanten Gespräche und die guten Ratschläge, Prof. R. Häner für sein Interesse an meinem research proposal, Martin Zwahlen (computertechnische Unterstützung), Herr Oetliker (Bau von Acrylamid-Teilen und Schweissarbeiten) Priska Bähler, Christoph Bächler, Ingrid Arnold, Sandra Christen, Philipp Schneider, Dr. Luis Garcia, vom Departement für Chemie und Biochemie der Universität Bern. Frau Prof. Lanzrein, Martha, Rames Pillai vom Zellbiologisches Institut der Universität Bern. Prof André Häberli (Vizedirektor Kinderklinik), Daniel Molina dipl.phil.nat, PD. Dr. Andreas Schaller vom Inselspital. Prof. Steven Leib, Denis Grandgirard von Institut für Infektionskrankheiten der Universität Bern für ihr interesse an meiner Arbeit und unser gemeinsames Projekt bei der Menigitis Research Foundation. Prof. Lussi von der Zahnmedizinischen Klinik der Universität Bern für sein Interesse am Spucke Projekt. Dr. Margarita Esquinas, Dr. Hans Zischka von BioRad für die “Müsterli“ Dr. Pierre Nording, Dr.M. Schöneberger, Dr. M. Müller von Fluka AG Buchs für die langjärige Zusammenarbeit. Marcel Salin, Peter Dollack, Andrea Schütz und Michael Dixon von Disetronic Medical Systems (Roche) in Burgdorf für unsere gemütlichen Feierabendbiere. Eliane Küttel (Patentanwältin), Davis A. Blumenthal and Christina Flores (Foley and Lardner Patentanwälte USA), Dr. Müller-Graf (Fürsprecher Uni Bern, Rechtskonsulent), Laura Ezquerra (Rechtskonsulentin), Sutter Treuhand (Rechtskonsulent). Dr. sc. nat. Cornelia Boesch von der Uitectra AG für das Magement während des Technologietransfers. Dr. Beatrice Michel (stv. Leiterin der Stelle für Oeffentlichkeitsarbeit der Uni Bern) vom Unilink Magzin, für die Redaktion meiner Artikel für das Unilink. Frau Prof. Etter vom Unipress Magzin, für die Redaktion meines Artikel für das Unipress .H. Wüthrich von der BZ (Ihr Artikel über mich hat eingeschlagen wie eine Bombe!). H. Beer von der InnoBE für den Internet Auftritt, Dr. Hans-Joachim Kraus (Senior Publishing Editor, Proteomics, Wiley-VCH Verlag, Weinheim) und Chris Trand für das Interesse an meinem Gel-Bildern und die Redaktion unseres Artikel in “Proteomics“. S. von May-Granelli (Feusi Humboldtianum AG), H. Richard (Bankrat DC Bank), Willy Michel (Techpharma Burgdorf), Prof. P. Mürner (Akademischer Direktor Uni Bern), Heinrich Mühlemann (Mühlemann Management Consulting) für die Gelegenheit mein Projekt beim DC Bank Preis präsentieren zu dürfen. Und aussederm danke ich all jenen welchen ich vergessen habe hier zu danken.

Proteomics 2004, *4,* 599–608 DOI 10.1002/pmic.200300587

Andreas Lamanda[1]
Alain Zahn[1]
Daniel Röder[2]
Hanno Langen[2]

[1]Departement für Chemie und Biochemie, Universität Bern, Bern, Switzerland
[2]F. Hofmann-La Roche, Basel, Switzerland

Improved Ruthenium II tris (bathophenantroline disulfonate) staining and destaining protocol for a better signal-to-background ratio and improved baseline resolution

In proteomics the ability to visualize proteins from electropherograms is essential. Here a new protocol for staining and destaining gels treated with Ruthenium II tris (bathophenantroline disulfonate) is presented. The method is compared with the silver-staining procedure of Swain and Ross, the Ruthenium II tris (bathophenantroline disulfonate) stain described by Rabilloud (Rabilloud T., Strub, S. M. Luche, S., Girardet, S. L. *et al.*, *Proteomics* 2001, *1*, 699–704) and the SYPRO Ruby gel stain. The method offers a better signal-to-background ratio with improved baseline resolution for both sodium dodecyl sulfate-polyacrylamide gels and two-dimensional gels.

Keywords: Baseline resolution / Destaining / Ruthenium II tris (bathophenantroline disulfonate) / Signal-to-background ratio PRO 0587

1 Introduction

There exist different staining methods to visualize proteins in polyacrylamide gels. The most sensitive is autoradiography of proteins radiolabelled biosynthetically or by radio-iodination. This method is unsurpassed for analytical purposes. It is inconvenient if proteins have to be isolated because the correlation between signals on a film and the invisible protein spot in the gel is difficult without special equipment. Silver-staining is the most efficient method for direct in-gel visualization. The dynamic range of this method however, is limited due to saturation effects. Staining is time-dependent, not an equilibrium reaction, and color as well as intensity of staining vary with type of protein in an idiosyncratic way [1]. Silver-stained proteins are difficult to analyze by MS. Staining with Coomassie Blue and Colloidal Coomassie Blue does not have any of these drawbacks but is less sensitive.

It has long been known that proteins covalently labelled with fluorescent dyes can be detected with high sensitivity. More recently noncovalent fluorescent stains have been introduced such as SYPRO Ruby, other dyes of the SYPRO family and ruthenium II tris-bathophenantroline disulfonate (RuBP) as described by Rabilloud *et al.* [2, 3]. While SYPRO Ruby is a ready to use formula, RuBP offers several possibilities for improvement and is cheaper. A number of modified staining protocols have been published for RuBP [4, 5] and SYPRO Ruby [4–7], and destaining protocols for iron II tris (bathophenantroline disulfonate) [8], SYPRO Ruby (Voshol, H, Novartis, private communication) [5] and europium tris (bathophenantroline disulfonate) [9]. The RuBP staining procedure by Rabilloud *et al.* was improved mainly by optimizing reagent concentration, pH, and solvent composition for both the staining and destaining steps. Here it is compared with silver-staining according to Swain and Ross (modified) [10], SYPRO Ruby staining according to Berggren *et al.* [4] and ruthenium tris (bathophenanthroline disulfonate) staining according to Rabilloud *et al.* [2].

2 Materials and methods

2.1 Synthesis of RuBP

RuBP was prepared as published by Rabilloud *et al.* [2] and used without further purification. The UV/visible spectrum of the product was identical to the published spectrum [3].

2.2 SYPRO® Ruby protein stain

SYPRO Ruby was purchased from Bio-Rad Laboratories (170–3125; Richmond, CA, USA). Check of the lot number confirmed the use of the new SYPRO Ruby formulation as described in [11].

Correspondence: Andreas Lamanda, Departement für Chemie und Biochemie, Freiestr. 3, Labor S162, Universität Bern, 3012 Bern, Switzerland
E-mail: andreas.lamanda@ibc.unibe.ch
Fax: +41-031-631-48-87

Abbreviations: AIDA, advanced image data analyzer; **LAU**, linear arbitrary units; **RuBP**, ruthenium II tris-bathophenantroline disulfonate

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

2.3 PAGE

Marker proteins (17–0446–01; Amersham Biosciences, Uppsala, Sweden) were serially diluted 1:2 (phosphorylase b 3350 ng to 0.82 ng, albumin 4150 ng to 1 ng, ovalbumin 7350 ng to 1.8 ng, carbonic anhydrase 4150 ng to 1 ng, trypsin inhibitor 4000 ng to 1 ng, lactalbumin 5800 ng to 1.4 ng) and separated on a 17.5% gel prepared as described [12] in a Protean 3 Mini cell system (Bio-Rad).

2.4 2-DE

(i) Analytical gels: For the first dimension, IPG strips (pH 3–10; Amersham Biosciences) were rehydrated for 10 h with 300 μL sample buffer (8 M urea, 2% CHAPS, 18 mM dithioerythritol (DTE), 0.5% IPG buffer 3–10 and traces of Bromophenole Blue) containing 80 μg *Escherichia coli* proteins. IEF was carried out in an IPGphor (Amersham Biosciences) with the following settings: 20°C, 200 μAmp *per* strip, 150 Vh (1 h, 150 V, step-n-hold), 300 Vh (1 h, 300 V, step-n-hold) 17 500 Vh (5 h, 3500 V, step-n-hold) 63 250 Vh (gradient). (ii) Preparative gels: IPG strips were rehydrated with 300 μL sample buffer containing 800 μg *E. coli* proteins for 10 h on the IPGphor under 30 V. Focussing was carried out with the following settings: 20°C, 200 μm Amp *per* strip, 150 Vh (1 h, 150 V, step-n-hold), 300 Vh (1 h, 300 V, step-n-hold) 17 500 Vh (5 h, 3500 V, step-n-hold) 27 6000 Vh (gradient), 80 000 Vh (10 h, 8000 V, step-n-hold). Second dimension (for i and ii): The IPG strips were equilibrated for 12 min with 5 mL/strip of solution I (50 mM Tris-HCl, pH 7.0, 6 M urea, 30% v/v glycerol, 2.5% w/v DTE) and 12 min with 5 mL/strip of solution II (50 mM Tris-HCl, pH 7.0, 6 M urea, 30% v/v glycerol, 2.5% w/v iodoacetamide). Before loading on the second dimension 1.5 cm of the basic end from the IPG strip was cut off. The second dimension was run on a 12% polyacrylamide gel (15 × 16 cm^2) at 2°C for 5 h at 50 mAmp *per* gel (500 V) in a Hoefer Dalt 600 chamber (Hoefer Scientific Instruments, San Fransico, CA, USA).

2.5 Staining procedures

A volume of 200 mL in each step of staining or destainig was used. Analytical gels were stained as detailed in protocols 1–4 (see Tables 1–4) with (i) silver nitrate as described by Swain and Ross [10] (with modifications), (ii) with SYPRO Ruby as described by the manufacturer, (iii) with RuPB according to the procedure of Rabilloud *et al.* [2], and (iv) with RuPB as follows: Gels were incubated in 30% ethanol and 10% acetic acid for 15 h (overnight), washed four times for 30 min each with 20% ethanol, and then stained with 1 μM RuBP in ion-exchange water. Staining was done in a stainless steel tray, in the dark, on a shaker for six h. The gels were destained by rinsing with water two times for 10 min, and then incubated in 40% ethanol/10% acetic acid for 15 h or overnight. The destained gels can be stored in water for several days without loss of signal intensity and can be silver-stained for further analysis [13].

Table 1. Protocol 1: Silver-staining protocol (Swain and Ross [4] modified)

Step	Procedure	Comments
1	Fix the gel in 40% EtOH/10% acetic acid for 2 h	For the sake of convenience the gel can be fixed overnight
2	Incubate the gel in 40% EtOH, containing 10 μL of 37% formaldehyde for 5 min	
3	Wash with 40% EtOH for 20 min	
4	Rinse for 20 min with water	
5	Incubate the gel for 1 min in sodiumthiosulfate (20 mg/100 mL)	
6	Stain with 1 g silvernitrate *per* 1000 mL for 20 min	
7	Rinse for one min with water and repeat three times	Do not rinse longer than one min each time
8	Develop with a solution of 25 g sodiumcarbonate and 400 μL 37% formaldehyde in 1000 mL water for 4 min and 15 s	After 1 min and 15 s the first spots become visible. If the solution turns yellow, it must be removed and replaced with an equal amount. Note: the best is to use a system that can remove at least 50 mL solution *per* second. Work has to be done hand free on all stages!
9	Stop developing with 10% acetic acid containing 0.6% Tris base for 5 min	
10	Rinse the gel with water for 5 min	Scan the gel immediately after development. Color varies even after a short time
11	Store the gel in 1% acetic acid at 4°C	

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2. Protocol 2: SYPRO Ruby gel stain [5]

Step	Procedure	Comments
1	Fix the gel in 10% MeOH, 7% acetic acid for 30 min	
2	Incubate the gel in SYPRO Ruby staining solution for 3 h-overnight	
3	Wash the gel in 10% MeOH, 7% acetic acid for 30 min and scan	
	all % are in V/V	

Table 3. Protocol 3: Ruthenium II tris (bathophenantroline disulfonate) staining protocol according to T. Rabilloud *et al.* [2]

Step	Procedure	Comments
1	Fix the gel in 30% EtOH, 10% acetic acid overnight	
2	Rinse the gel in 20% EtOH for 30 min and repeat 3 times	
3	Incubate the gel in 100 nM RuBP solution for 6 h	The concentration of the stock solution is 20 mM. Dilute 5 μL in 1000 mL 20% EtOH just prior to use
4	Equilibrate the gel in water for 10 min, repeat once and scan	
	all % are in V/V	

Table 4. Protocol 4: Ruthenium II tris (bathophenantroline disulfonate) new staining/destaining protocol [13]

Step	Procedure	Comments
1	Fix the gel in 30% EtOH, 10% acetic acid overnight	
2	Rinse the gel in 20% EtOH for 30 min and repeat 3 times	
3	Incubate the gel in 1 μM RuBP solution for 6 h	The concentration of the stock solution is 20 mM. Dilute 50 μL in 1000 mL water just prior to use
4	Equilibrate the gel in water for 10 min and repeat once	A first scan is possible at this stage
5	Destain the gel with 40% EtOH/10% acetic acid for 15 h	
6	Equilibrate the gel in water for 10 min repeat once and scan	
	all % are in V/V	

2.6 Imaging

Silver-stained gels were scanned on a flatbed scanner (HP Deskscan, DeskScanII V2.3) with the following scanning parameters: 300 × 300 dots *per* inch, eight bit black and white picture (256 grey shades, two times sharpened), contrast 125, brightness 125. SYPRO Ruby and RuBP-stained gels were scanned with a Phosphorimager (Fuji FLA-3000 from Raytest with software BASReader V3.01 Straubenhardt, Germany) using the following scanning parameters: resolution 50 μm, 16 bit picture (65 536 grey shades), sensitivity 1000, excitation wavelength 473 nm and detection filter O580. Images were processed with advanced image data analyzer (AIDA) V3.10. Profiles were drawn with the 1-D evaluation tool of AIDA V3.11. 2-D Densitometric analysis was done with the 2-D densitometry module of AIDA 3.11.002. First the protein spots were captured. Then peak volumes

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

over background were calculated by numeric integration of the grey values of each pixel within a spot area. The background was measured in a rectangular box of 5 mm^2 near the spot and was subtracted from the numerical integral *per* area.

2.7 Protein identification

Coomassie or RuBP-stained protein spots were cut out of the gel. The gel pieces were destained with 100 mM ammonium bicarbonate in 30% ACN. Proteins were digested with trypsin [14] and peptide masses were identified by MALDI TOF as described [15]. The probability of a false positive match of an observed MS-spectrum was determined for each analysis [16].

3 Results and discussion

3.1 SDS PAGE

Two-fold serially diluted marker proteins were separated by SDS-PAGE (Fig. 1) and stained with silver, SYPRO Ruby and RuBP respectively, according to the protocols listed in Tables 1–4. Gels stained with RuBP and SYPRO Ruby were scanned in the phosphor-imager FLA-3000 (Raytest). Silver-stained gels were scanned on a flatbed scanner, and images were transformed to linear Arbitrary Units (LAU) with AIDA. Silver-staining produces uneven staining. Protein bands stained with fluorescent dyes have an overall balanced appearance. Peak volumes over background were used to measure the quantity of protein. A plot of signal intensities (integral/area-background) *versus* protein amount obtained from 2-D densitometric evaluation is shown for each protein and staining technique in Fig. 2. The threshold of detection was 2 ng for the silver-stain procedure (Lactalbumin, 14.1 kD), 4 ng for SYPRO Ruby (Trypsin inhibitor, 20.1 kD), 16 ng for the RuBP stain (Carbonic anhydrase, 30 kD) according to [2] and 8 ng for the modified RuPB procedure (Trypsin inhibitor, 20.1 kD). Silver-staining was the most sensitive method for all quantified proteins. SYPRO Ruby, which was second best in sensitivity, did not show as good linear dose/response behavior and had less intense signals than the modified RuBP stain which gave the strongest signals and best linear behavior (Fig. 2). Staining of 1-D gels with silvernitrate, according to the modified RuBP procedure, gave an artefact caused by the gel electrophoresis unit. It occured in the 60 kD region. This artefact shows up frequently in gels in our laboratory. It occurs in all types of analysis for all types of samples from different species in varying intensities. It can not be avoided. It may originate from traces of detergent used to cleanse the glass plates.

The behavior of signal-to-background was investigated by profiling and 2-D densitometry. Profile curves of lanes 1, 4 and 7 from each gel were recorded (Fig. 3) by scanning the lanes from edge to edge along a 1.05 mm wide rectangular field at their centre. Areas at the upper and lower edge of the lanes containing no protein were chosen to determine the background. The peak heights obtained with silver-staining are unrelated to the protein amount and therefore cannot be used for quantification. Peak heights obtained with fluoresencent dyes are proportional to the protein amounts but the relative intensities vary from scan to scan as is obvious from a comparison of the scans of lanes 1, 4 and 7. The signal-to-background ratio was 60/0.2 for the improved RuBP procedure, and 7/0.8 and 23/3 for SYPRO Ruby and

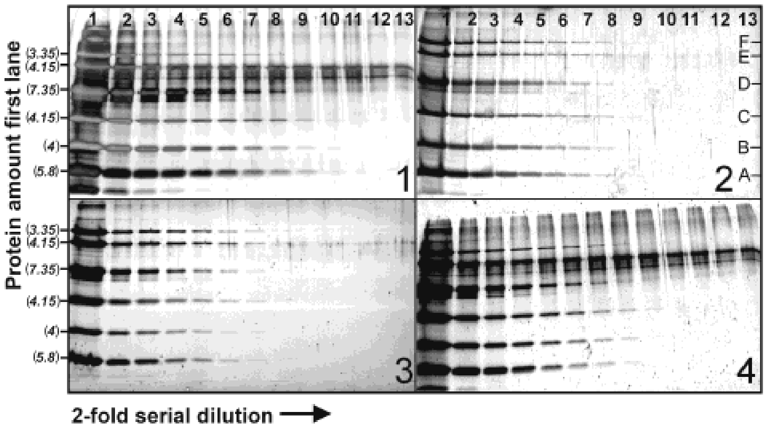

Figure 1. Signal strength and quality and background staining by different protein detection procedures. (1) Silver-stain; (2) SYPRO Ruby Protein Gel Stain; (3) Ruthenium bathophenantroline disulfonate according to the procedure of Rabilloud *et al.;* (4) modified Ruthenium bathophenantroline disulfonate procedure. Marker proteins were subjected to two-fold serial dilution. (A) Lactalbumin, 14.4 kD (5800 ng), (B) Trypsininhibitor, 20.1 kD (4000 ng), (C), Carboanhydrase, 30 kD (4150 ng), (D) Ovalbumin, 45 kD (7350 ng), (E) Albumin, 67 kD (4150 ng), (F) Phosphorylase b, 94 kD (3350 ng). The exact amount of each protein in lane 1 is indicated in parenthesis. Notice, that staining with silver (1) and RuBP (4) produces an electrophoresis artefact in the 60 kD region of the gel.

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

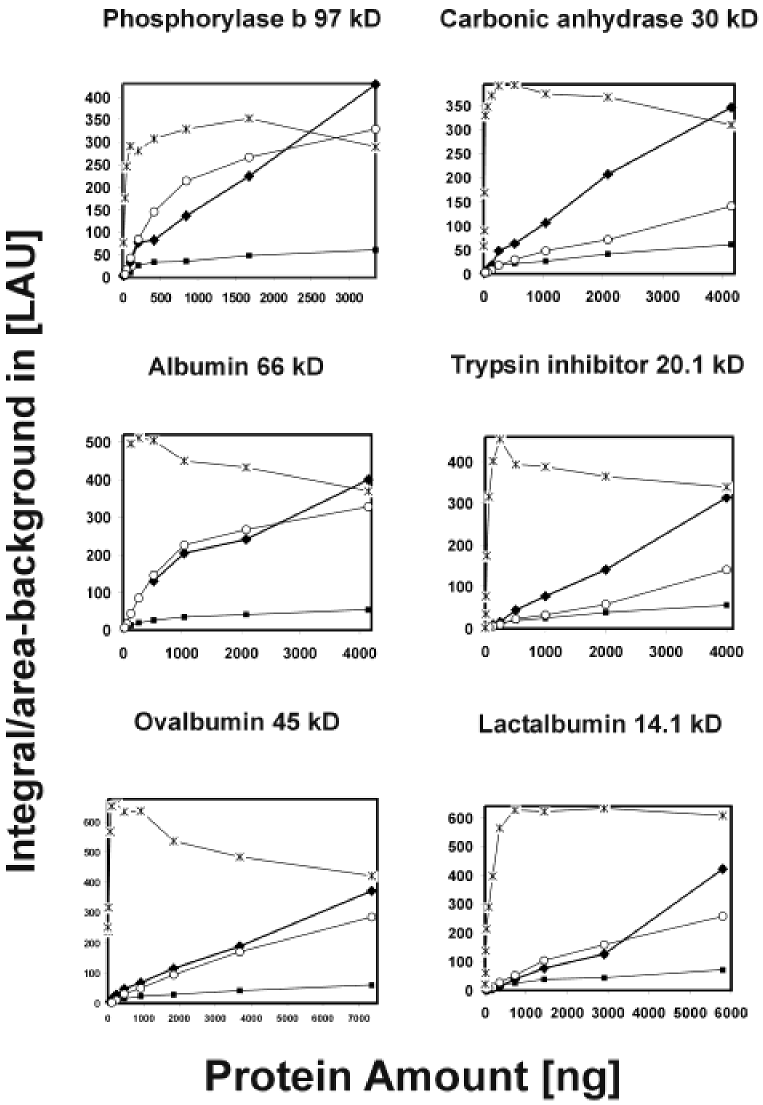

Figure 2. Plot of signal intensities (integral/area-background) *versus* protein amount obtained from 2-D densitometric evaluation. The symbols used are as follows: modified RuBP stain, diamonds; silver-stain, stars; SYPRO Ruby, squares; RuBP according to the procedure of Rabilloud *et al.,* open circles.

RuBP staining [2] respectively. Even at the highest protein amount/band, the fluorescent stains did not produce flat peaks (Fig. 3, lane 1). The artificial band mentioned above produces a peak of 4 LAU in all lanes of the gel stained with the new method. As it comigrates with albumin, both peaks of the profile overlap. In lanes 1 to 4 the albumin peak is dominant. In lane 7 the peak caused by the artefact is dominant over the albumin peak which is not visible.

Peak height can be affected by local anomalies and bands often are unevenly distributed across a lane. Peak height measurement is therefore potentially erroneous. To circumvent this, the peak volumes over background were compared to background using the 2-D densitometry Module of AIDA. Figures 4, 5, and 6 show plots of the ratio of signal intensities-to-background (integral/area-background) *versus* protein amount. The modified RuBP staining gave the best ratio with a maximum ratio of 74.2 for Lactalbumin, which was up to 12-fold stronger than the second best ratio (Fig. 4). The good linear behaviour becomes visible here. In Figure 5 the signal intensity scale is restricted to 10 LAU and below. SYPRO Ruby and RuBP staining according to the procedure of Rabilloud *et al.* produced signals up to 6.1 LAU. Silver-staining showed the typical curve of saturation. The plots of protein amounts below 700 ng for Lactalbumin, 450 ng for Ovalbumin and 250 ng for the other proteins tested are shown in Fig. 6. For protein loads below 50 ng silver-staining always produced the best signal-to-background ratios. The only exception was seen for Phosphorylase b where the modified RuBP procedure

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

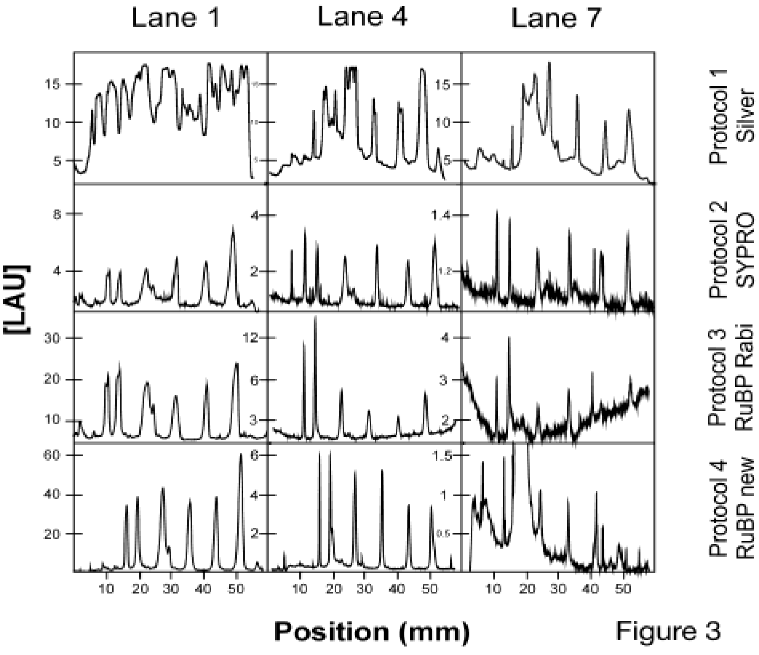

Figure 3

Figure 3. Profile scans of lanes 1, 4, and 7 from gels shown in Fig. 1. Different staining procedures produce markedly different signal intensities (y-axis). The signal-to-background ratio is close to 200:1 for the modified RuBP staining procedure, 5:1 for SYPRO Ruby, 8:1 for the RuBP procedure by Rabilloud *et al.,* and 6:1 for the silver-stain procedure.

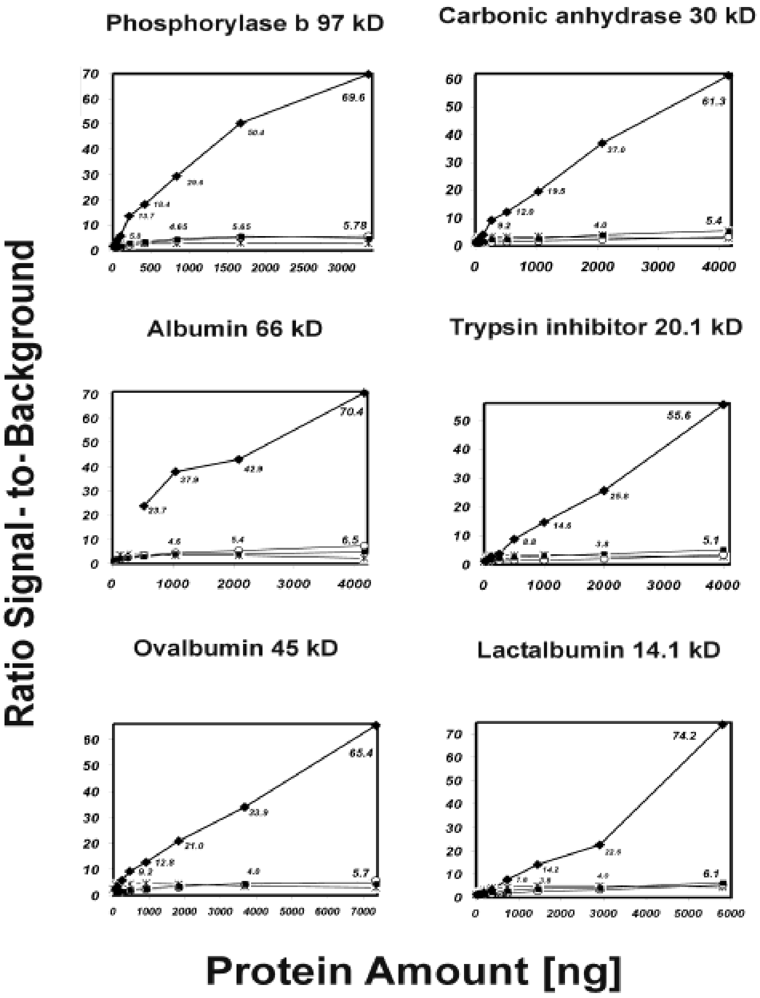

Figure 4. Plot of the ratio of signal intensities-to-background (integral/area-background) *versus* protein amount obtained from 2-D densitometric evaluation. The symbols used are as follows: modified RuBP stain, diamonds; silver-stain, stars; SYPRO Ruby, squares; RuBP according to the procedure of Rabilloud *et al.,* open circles. The signal-to-background values are indicated for the new RuBP stain and the second best stain. The Modified RuBP stain generated a maximum ratio of 74.2 compared to the second best maximum ratio of 6.5 from the other staining methods. The modified RuBP staining showed the best linearity.

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

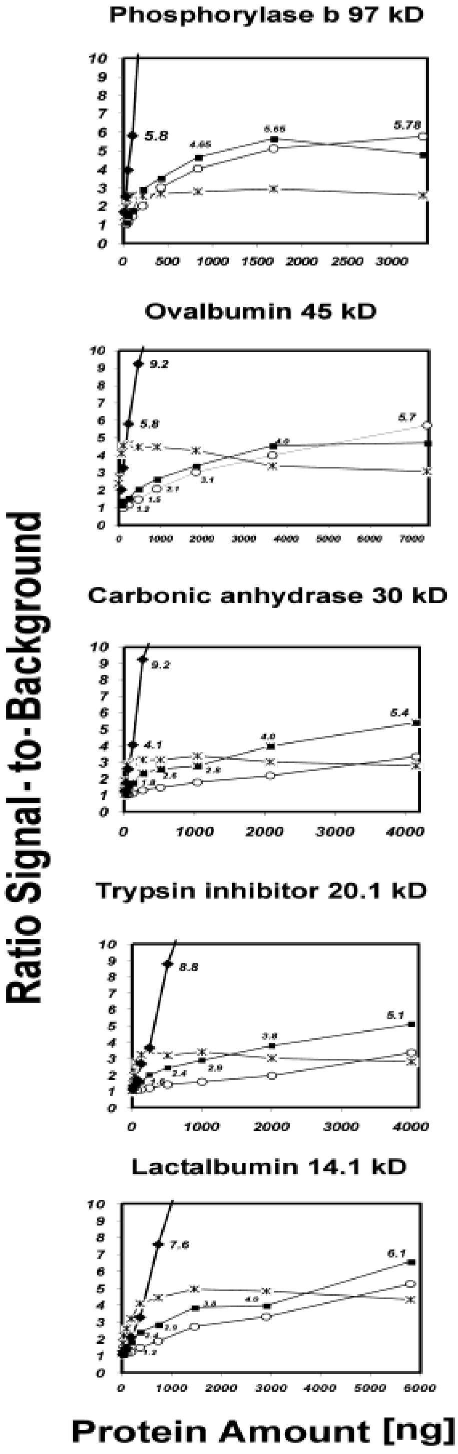

Figure 5. Plot of the ratio of signal intensities-to-background *versus* protein amount for ratios below 10. The symbols used are as follows: modified RuBP stain, diamonds; silver-stain, stars; SYPRO Ruby, squares; RuBP according to the procedure of Rabilloud *et al.,* open circles. Silver-staining reached saturation. The RuBP stain according to Rabilloud *et al.* and the SYPRO Ruby Protein stain showed a slow increase in ratio whereas the modified RuBP stain was out of scale.

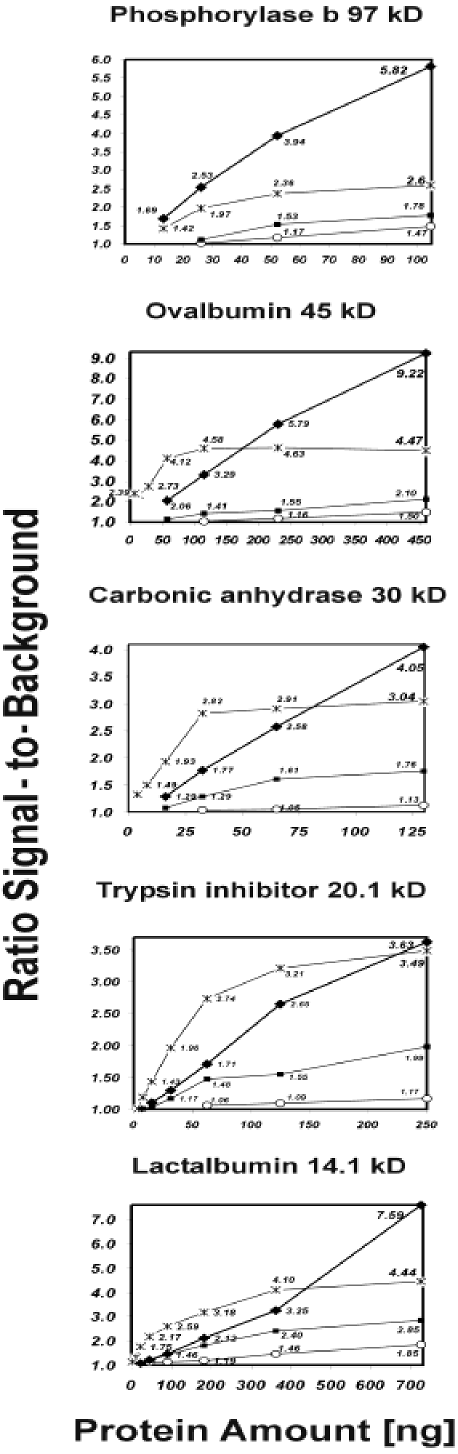

Figure 6. Plot of the ratio from signal intensities-to-background *versus* protein amount for protein quantities below 700 ng and ratios below 9. The symbols used are as follows: modified RuBP stain, diamonds; silver-stain, stars; SYPRO Ruby, squares; RuBP according to the procedure of Rabilloud *et al.,* open circles. For low protein amounts, silver-staining showed the best sensitivity but early onset of saturation whereas the modified RuBP staining has better linearity.

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

was better. For all other proteins, the second best values were obtained with the modified RuBP procedure. SYPRO Ruby can only compete for Lactalbumin at the two lowest concentrations.

Compared to SYPRO Ruby protein staining, the modified RuBP procedure gives better contrast but is less sensitive. This is due to the final destaining step. During destaining RuBP molecules are removed selectively from the gel matrix but not from proteins. For low protein concentrations the destaining time might be too long. Therefore, these proteins are destained too much and loose their fluorescence. Shorter destaining times may optimize the procedure resulting in greater sensitivity.

3.2 2-D gels

Protein expression proteomics requires the quantitative comparison of samples that differ by multiple small variables. The significance of small differences can be assessed only by statistical treatment of the raw data, which in this case is protein spot intensity and/or area. Therefore, systematic error caused by protein staining should be kept as low as possible. The highly sensitive silver-staining is time-dependent. Therefore, data derived from different gel batches are difficult to compare. Even more difficult is the comparison of data from different laboratories. In contrast, fluorescent staining is an equilibrium reaction, and, being time-independent, more suitable for comparison. For comparison, four 2-D gels (Fig. 7) of urea soluble *E. coli* proteins were stained according to the protocols listed in Tables 1–4.

First, the number of protein spots was compared. 1000 spots were visible by silver-staining, 650 by SYPRO-Ruby staining, 430 by RuBP staining according to Rabilloud *et al.* and 750 spots by the optimized RuBP staining procedure. Second the signal/background ratio was determined. Visual inspection of the four gels shows that the modified RuBP staining affords a picture with high contrast. The intensity of six protein spots (labelled A–F in Fig. 7), selected from the four quadrants of the 2-D gels, was determined as the integral of grey shade *per* scanned line in the profiled area. The background was measured over a protein free area of the same size. Results are shown in Fig. 8A. The signal/background ratios are comparable or better for protein spots stained with the modified RuBP procedure than with any of the other three methods. The next best method was RuBP staining according to Rabilloud *et al.* and SYPRO Ruby staining. Silver-staining was less good.

Thirdly a profile was drawn over a train of three protein spots in the acidic 70 kD region of the gels. The profile boxes are shown in the enlarged region of each gel (Fig. 7). The first protein (Fig. 7A) was identified as chaperone protein Dnak (P04475). The second protein (Fig. 7B) was identified as enzyme I (PtsI, P08839) of the bacterial phosphoenolpyruvate-protein phosphotransferase system. Four Peptides of the third protein matched Dnak.

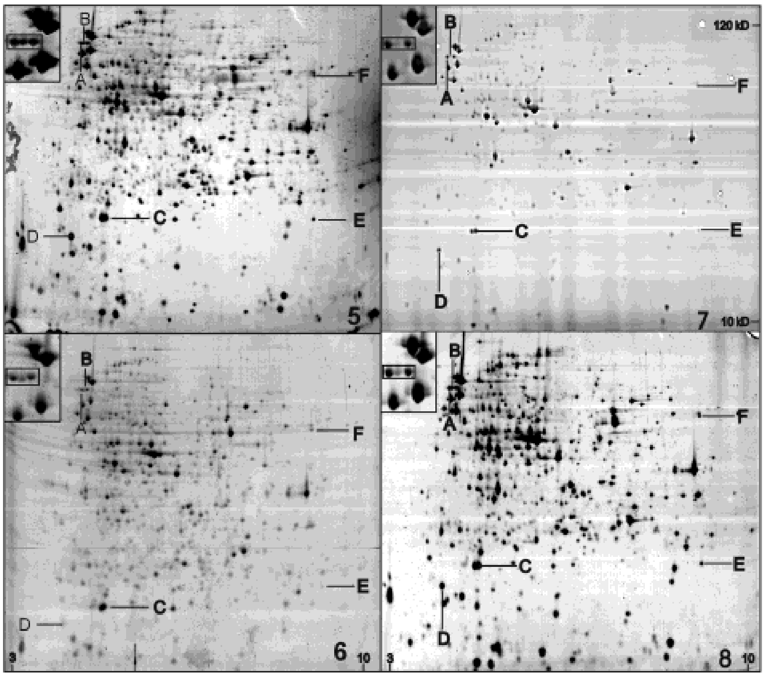

Figure 7. 2-D gels of *E. coli* proteins stained with silver nitrate (5), SYPRO Ruby (6), RuBP according to the procedure of Rabilloud (7), and with RuBP with modifications (8). Eighty μg of protein were loaded *per* gel. For details see Section 3.2. The inserts (top left) show the three spots used for the profile scans shown in Fig. 8A. A–F indicate protein spots used for the calculation of signal/background ratios given in Fig. 4B. A: Dnak, B: PtsI, C: ppa, D: crr, E: Eno, F: Rho.

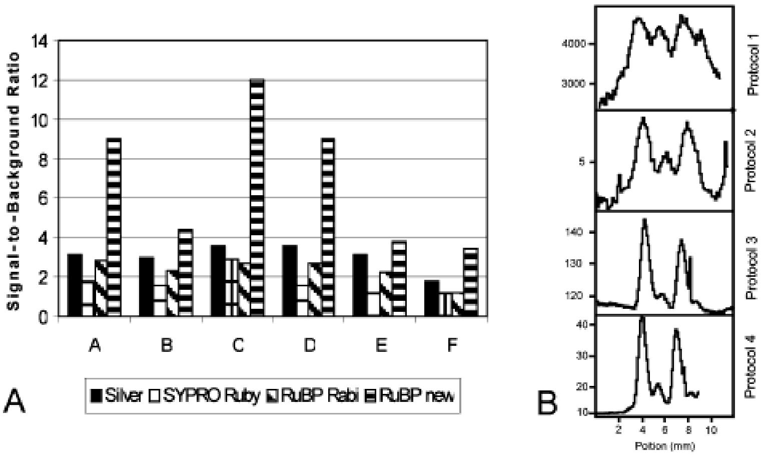

Figure 8. (A) Signal strength-to-background intensity of protein spots A–F from the 2-D gels shown in Fig. 7. Plotted is the ratio of peak intensity/baseline intensity. Intensities in LAU were obtained from scans across entire spots. The proteins were identified by MALDI-TOF. (B) Profile scans over three adjacent protein spots (Dnak, PtsI and unknown), from the 2-D gels shown in Fig. 3., stained with silvernitrate (5), SYPRO Ruby (6), RuBP according to the procedure of Rabilloud *et al.* (7), the modified RuBP procedure (8). Notice differences in peak resolution and signal-to-background intensities.

The identification was not considered safe as it did not reach the value (negative log $p = 5$) needed to be excluded as a false positive match. The profile scans are shown in Fig. 8B. Silver-staining does not allow for baseline resolution between the three spots. The signal/background ratio for Dnak, the spot on the left, is around 3. Fluorescent staining affords consistently better resolution. SYPRO Ruby staining affords partial, but not baseline resolution and a signal/background ratio of 1.8 for Dnak. RuBP staining according to the procedure of Rabilloud *et al.* and modified as described in this communication gave baseline resolution and signal/background ratios of 2.8 and 9, respectively for Dnak.

Reproducibility was checked by comparison of the spot volumes of six reproduced spots from six independently run and RuBP stained 2-D gels (Table 5). Spot volumes were determined by 2-D densitometry and were then normalized using 2-D Advance Software V6.01 from Phoretix (Nonlinear Dynamics, Durham, NJ, USA). The raw data show SDs of 17 to 31% in spot volume and from 10 to 21% in the normalized data set. The overall SD was 17.4% for the raw data and 15% for the normalized data. As little as 10 ng of protein *per* spot can be analyzed with the current methods of high-throughput protein identification by MS (Röder, D., F. Hofmann-La Roche, private communication). The sensitivity threshold of RuBP staining is between 6 and 8 ng/spot and is therefore sufficient for the detection of protein spots which can be analyzed by current state-of-the-art technologies.

4 Concluding remarks

The modified RuBP staining is suitable for protein identification by MALDI-TOF. Proteins from spots A (Dnak), C (ppa) and D (IIA^{Gluc}) (Fig. 7) were cut out of 2-D gels giving a total load of 300 μg of *E. coli* protein and stained according to the modified RuBP staining procedure. Proteins were identified *via* peptide mass fingerprinting (data not shown). Detailed studies to compare the sequence coverage with spectra generated by proteins visualized by SYPRO Ruby and the Rabilloud technique are in progress. In conclusion, the modified RuBP staining method has considerable advantages. Staining intensity is linear with protein amount and baseline resolution between

Table 5. 2-D densitometry results

	Gel			Integral/Area-background in (LAU)					AIDA	2-D Advance
Spot	1	2	3	4	5	6	Average	SD	SD in%	SD in%
1	0.35	0.18	0.23	0.2	0.18	0.18	0.22	0.07	30.29	13.94
2	0.31	0.24	0.21	0.24	0.19	0.18	0.23	0.05	20.62	14.48
3	0.18	0.17	0.17	0.23	0.16	0.14	0.18	0.03	17.24	10.35
4	0.31	0.17	0.18	0.23	0.17	0.17	0.21	0.06	27.55	10.10
5	0.27	0.26	0.22	0.21	0.17	0.15	0.21	0.05	22.32	21.87
6	0.3	0.24	0.28	0.24	0.18	0.18	0.24	0.05	20.99	19.10
									Averages	
								0.04	17.37	14.97

closely adjacent protein spots is excellent. The signal-to-background behavior is better than for any of the compared fluorescent staining methods. This fact becomes visible in the excellent contrast of gel images. Good contrast enables protein spots to be detected more easily by image processing software. The method is suitable for staining of analytical gels which can be quantitatively analyzed on a fluorescent scanner or a CCD camera as well as for preparative gels which can be inspected on an UV transiluminator at 254 nm for spot excision. SYPRO Ruby stained gels are known to break (Esquinas, M., Bio-Rad, private communication) easily. This is not the case for RuBP-stained gels. Unlike silver nitrate, which is used at a concentration of 1 g *per* litre, the RuBP concentration needed is only 1 μM. Ruthenium II tris (bathophenantroline disulfonate*)* is not listed as a environmental toxic compound (Bundesamt für Gesundheit Sektion Gewässerschutz, private communication). Last but not least, RuBP is easy to prepare and significantly cheaper to use than other commercially available fluorescent stains.

Said Modaressi from Raytest Isotopenmessgeräte GmbH is acknowledged for his help with the Phosphorimager and AIDA. Cornelia Boesch from Unitectra AG and Pierre Nording from Fluka Productions GmbH are acknowledged for their exellent mangement. Stephan Reber from the ETH Zürich is acknowledged for his keen interest and the ideas he contributed, Franziska Buchmann for her support and Prof. Bernhard Erni from the University of Bern for his careful reading and correction of the manuscript. This work was financed by the Swiss National Science Foundation (grant 31-45838.95). This Project was part of a technology transfer project. It was awarded the DC Bank Prize for Research and Development 2002 (1st rank). Since November 2002 Ruthenium II tris (bathophenantroline disulfonate) and additional reagents can be purchased from Fluka Productions GmbH (Catalogue Number 03038, p. 1578; Buchs, Switzerland).

Received April 7, 2003
Accepted July 28, 2003

5 References

[1] Mortz, E., Krogh, T. N., Vorum, H., Görg, A., *Proteomics* 2001, *1*, 1359–1363.

[2] Rabilloud, T., Strub, J. M., Luche, S., Dorsselaer, A. *et al.*, *Proteomics* 2001, *1*, 699–704.

[3] Rabilloud, T., Strub, J. M., Luche, S., Girardet, J. L. *et al.*, *Proteome*, http://link.springer.de/link/service/journals/10216/contents/00/00002/

[4] Berggren, K., Chernokalskaya, E., Steinberg, T. H., Kemper, C. *et al.*, *Electrophoresis* 2000, *21*, 2509–2521.

[5] Bhalgat, M. K., Diwu, Z., Haugland, R. P., Patton, W. F., *United States Patent*, Patent No. US 6, *316*, 267 B1, 2001, Example 7.

[6] Berggren, K., Steinberg, T. H., Wendy, M., Lauber, W. M. *et al.*, *Anal. Biochem.* 1999, *276*, 129–143.

[7] Berggren, K., Chernokalskaya, E., Lopez, M. F., Beechem, J. M. *et al.*, *Proteomics* 2001, *1*, 54–65.

[8] Graham, G., Nairn, R. S., Bates, G. W., *Anal. Biochem.* 1978, *88*, 434–441.

[9] Patton, W. F., Lim, M. J., Shepro, D., in: Link, A. J. (Ed.), *2-D Proteome analysis Protocols*, Humana Press, Totowa, NJ, USA 2001, pp. 353–362.

[10] Swain, M., Ross, N. W., *Electrophoresis* 1995, *16*, 948–951.

[11] Berggren, K., Schulenberg, B., Lopez, M. F., Steinberg, T. H. *et al.*, *Proteomics* 2002, *2*, 486–498.

[12] Laemmli, U. K., *Nature* 1970, *227*, 680–685.

[13] Lamanda, A. USA: *Patent Application No. 60/454, 979*, 2003.

[14] Fountoulakis, M., Langen, H., *Anal. Biochem.* 1997, *250*, 153–156.

[15] Langen, H., Fountoulakis, M., Berndt, P., Suter, L., *Electrophoresis* 2002, *23*, 311–328.

[16] Berndt, P., Hobohm, U., Langen, H., *Electrophoresis* 1999, *20*, 3521–3526.

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Mein herzlicher Dank geht an:

Gertrud und Andreas Lamanda

Prof. Dr. B. Erni

Prof. Dr. U. Jenal

Franziska Buchmann

Stephan Reber

PD Dr. H. Langen

Außerdem danke ich den folgenden Leuten (Reihenfolge zufällig)

Daniel Röder, Dr. Stefan Evers, Dr. Michel Fountoulakis, Dr. Peter Berndt von F. Hoffmann-La Roche, Basel, für die vielen guten Ratschläge, die durchgeführte Analytik, den freundlichen Empfang und die Mittagessen. Dr. Gerd Grenner (Chief Technology Officer), für die interessanten Gespräche. Dr. Reiner Westermeier von Amersham Pharmacia Biotech Europe, Freiburg für die interessanten Gespräche. Prof. Keith Rose (Chief Science Officer) GeneProt Inc, Genf, für den freundlichen Empfang, die Gespräche und die Besichtigung von GeneProt. Dr. Jan Van Oostrum, (Unithead of Proteomics), Dr. Hans Voshol, Dr. Sjouke Hoving von Novartis Pharma, Basel, für die interessanten Diskussionen, die Nomination für den Poster Award in München, die vielen Informationen und die Mittagessen. Dr. B. Grünenfelder, Gabriele Rummel vom Biozentrum der Uni Basel für die Unterlagen, den "mini 2-D Gel Kurs", den freundlichen Empfang. Prof. Dennis Hochstrasser, Maged Saad, Peter Skorpil, Severine Hughes, vom Universitätsspital Genf für die informativen Gespräche die wir führten. Dr. Jan Sanders (European Sales Manager), Nonlinear Dynamics für die freundliche Beratung und die Blitzschulung in Phoretix 2D Advance. Dr. Said Modaressi, Dr. M. Liebler (Vicepresident) und G. Keck, Raytest Isotopen Messgeräte GmbH (Fuji), für die langjährige gute Beratung, die Blitzschulung in AIDA, den Webauftritt und das Interesse an meiner Arbeit. Dr. Johannes Hewel (für die ESI-TOF Analytik), Alain Zahn (für das Synthetisieren von RuBP), Dr. Rudolf Beutler, PD Dr. Johann Schaller (für die Analytik , das Material, die interessanten Gespräche und das Verteilen meiner Flyers in den USA), Urs Kämpfer für die interessanten Gespräche und die guten Ratschläge, Prof. R. Häner für sein Interesse an meinem research proposal, Martin Zwahlen (computertechnische Unterstützung), Herr Oetliker (Bau von Acrylamid-Teilen und Schweissarbeiten) Priska Bähler, Christoph Bächler, Ingrid Arnold, Sandra Christen, Dr. Luis Garcia, vom Departement für Chemie und Biochemie der Universität Bern. Frau Prof. Lanzrein, Martha, Rames Pillai vom Zellbiologisches Institut der Universität Bern. Prof André Häberli (Vizedirektor Kinderklinik), Daniel Molina dipl.phil.nat, PD. Dr. Andreas Schaller vom Inselspital. Prof. Steven Leib, Denis Grandgirard von Institut für Infektionskrankheiten der Universität Bern für ihr interesse an meiner Arbeit und unser gemeinsames Projekt bei der Menigitis Research Foundation. Prof. Lussi von der Zahnmedizinischen Klinik der Universität Bern für sein Interesse am Spucke Projekt. Dr. Margarita Esquinas, Dr. Hans Zischka von BioRad für die "Müsterli" Dr. Pierre Nording, Dr.M. Schöneberger, Dr. M. Müller von Fluka AG Buchs für die langjärige Zusammenarbeit. Marcel Salin, Peter Dollack, Andrea Schütz und Michael Dixon von Disetronic Medical Systems (Roche) in Burgdorf für unsere gemütlichen Feierabendbiere. Eliane Küttel (Patentanwältin), Davis A. Blumenthal and Christina Flores (Foley and Lardner Patentanwälte USA), Dr. Müller-Graf (Fürsprecher Uni Bern, Rechtskonsulent), Laura Ezquerra (Rechtskonsulentin), Sutter Treuhand (Rechtskonsulent). Dr. sc. nat. Cornelia Boesch von der Uitectra AG für das Magement während des Technologietransfers. Dr. Beatrice Michel (stv. Leiterin der Stelle für Oeffentlichkeitsarbeit der Uni Bern) vom Unilink Magzin, für die Redaktion meiner Artikel für das Unilink. Frau Prof. Etter vom Unipress Magzin, für die Redaktion meines Artikel für das Unipress .H. Wüthrich von der BZ (Ihr Artikel über mich hat eingeschlagen wie eine Bombe!). H. Beer von der InnoBE für den Internet Auftritt, Dr. Hans-Joachim Kraus (Senior Publishing Editor, Proteomics, Wiley-VCH Verlag, Weinheim) und Chris Trand für das Interesse an meinem Gel-Bildern und die Redaktion unseres Artikel in "Proteomics". S. von May-Granelli (Feusi Humboldtianum AG), H. Richard (Bankrat DC Bank), Willy Michel (Techpharma Burgdorf), Prof. P. Mürner (Akademischer Direktor Uni Bern), Heinrich Mühlemann (Mühlemann Management Consulting) für die Gelegenheit mein Projekt beim DC Bank Preis präsentieren zu dürfen. Und aussederm danke ich all jenen welchen ich vergessen habe hier zu danken.

Curriculum vitae

Name: **Andreas Lamanda**

Born: 28.06.1968
From : Oberburg, Switzerland

Education/Training

Institution and location	Degree	Year(s)	Field of study
Primary school		1975-1980	
Secondary school		1980-1984	
Secondary school, Humboldtianum, Bern, Switzerland		1984-1985	
Humboldtianum, Bern, Switzerland		1985-1986	
Sanitized Inc., Burgdorf, Switzerland	eidg. dipl.	1986-1989	Lab technician (apprentice), microbiology and analytical chemistry
International House College, Sydney, Australia	diploma	1989	Language intensive course
Feusi-Humboldtianum, Bern, Switzerland	eidg. Matura	1990-1994	Gymnasium education (languages typ D)
L.I.F.E. Institute, Fairleigh Dickinson University, Rutherford, New Jersey, USA	diploma	1992	Language intensive course
University, Bern, Switzerland		1994-1995	Geology
University, Bern, Switzerland	dipl.phil.nat	1995-1999	Chemistry
University, Bern, Switzerland	Dr.phil.nat.	2000-2004	Biochemistry (Proteomics)

Positions

1989	Lab technician, Sanitized Inc. (Sandoz), Burgdorf, Switzerland.
1994-1996	Lab technician, Phytomed Inc., Hasle, Switzerland.
1995-2000	Private teacher for Chemistry, Bern, Switzerland.
2000-2004	Assistant and system administrator in the group of Prof. B. Erni.

Professional memberships

2000	USGEB Swiss Societies for Experimental Biology
2001	CHEL Communauté Helvetique d'Electrophorese
2002	SPS Swiss Proteomic Society DEG Deutsche Elektrophorese Gesellschaft
2003	SCG Schweizerische Chemische Gesellschaft

Honors

2002	DC Bank Price for R&D