# MELiSSA

TECHNICAL NOTE

Memorandum of Understanding TOS-MCT/2002/3161/In/CL




## TECHNICAL NOTE: 70.4

## PART A: TEST REPORTS FOR GENETIC STABILITY STUDY, RUN AS ANNOTATED TEST PROCEDURES

| | |
|---|---|
| prepared by/*préparé par* | L. Hendrickx, R. Wattiez, S. Baatout, A. Wilmotte, H. De Wever, P. Janssen, M. Mergeay |
| reference/*réference* | 15680/01/NL/ND |
| issue/*édition* | 1 |
| revision/*révision* | 1 |
| date of issue/*date d'édition* | 15/03/2004 |
| status/*état* | Draft |
| Document type/*type dedocument* | Technical Note |
| Distribution/*distribution* | C. Paillé, N. Dorval, R. Wattiez, A. Wilmotte, H. De Wever, M. Mergeay, S. Baatout, L. Hendrickx, P. Janssen |

CONFIDENTIAL DOCUMENT

## APPROVAL

| | | | |
|---|---|---|---|
| Title *titre* | Part A: Test reports for genetic stability study, run as annotated test procedures | issue *issue* 1 | revision *revision* 2 |
| author *auteur* | L. Hendrickx, S. Baatout, R. Wattiez, A. Wilmotte, P. Janssen, H. De Wever, M. Mergeay | date *date* | 15/03/ 2004 |
| approved by *approuvé by* | | date *date* | |

## CHANGE LOG

| reason for change */raison du changement* | issue/*issue* | revision/*revision* | date/*date* |
|---|---|---|---|
| | | | |

## CHANGE RECORD

Issue: 1 Revision: 2

| reason for change/*raison du changement* | page(s)/*page(s)* | paragraph(s)/*paragraph(s)* |
|---|---|---|
| | | |

# ABREVIATIONS

| | |
|---|---|
| AFLP | Amplified Fragment Length Polymorphysm |
| BLAST | Basic Local Alignment Search Tool |
| CFDA | carboxy-fluorescein diacetate |
| CHAPS | 3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate |
| CV | Coefficient of Variation |
| DGGE | Denaturing Gradient Gel Electrophoresis |
| DioC(6)3 | 3,3'-dihexyloxacarbocyanine iodide |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acid |
| dNTPs | Deoxynucleotide triphosphates |
| EDTA | Ethylene-Diamine Tetra Acetic acid |
| EtBr | Ethidium Bromide (2,7-diamino-10-ethyl-9-phenyl-phenanthridium bromide) |
| EWC | Engineering of the Waste Compartment |
| FCM | Flow Cytometry |
| FDA | fluorescein diacetate |
| HE | hydroethidine |
| HPLC | High-Pressure (or High-Performance) Liquid Chromatography |
| ICM-MS | Intact cell MALDI_TOF Mass Spectrometry |
| MALDI-TOF | Matrix Assisted Laser Desorption Ionization-Time of Flight |
| MCP | Micro-Channel Plate detector |
| MOPS | 4-Morpholinepropanesulfonic acid |
| MS | Mass Spectrometry |
| NCTC | National Collection of Type Cultures (Colindale, UK) |
| OD | Optical Density |
| OPA | One Phor All buffer |
| ORF | Open Reading Frame |
| PAGE | Pulsed Amplified Gel Electrophoresis |
| PCR | Polymerase Chain Reaction |
| PI | Propidium Iodide |
| PPM | parts per million |
| rDNA | rRNA encoding DNA |
| Rho-123 | Rhodamine-123 |
| RDP-II | Ribosomal Database Project (version 2) |
| RPM | Revolutions Per Minute |
| SAPD | Surface Accessible Proteins Detection |
| SCM-MS | Supernatant Culture MALDI-TOF MS |
| SDS | sodium dodecyl sulfate |
| TAE | Tris-Acetate-EDTA |
| TBE | Tris-Borate-EDTA |
| TCA | TrisChloroacetic Acid |

| | |
|---|---|
| TE | Tris EDTA |
| TEM | Transmission Electron Microscopy |
| TLF | Time Lag Focusing |
| Tris | Tris(hydroxymethyl)-aminomethane |
| UV | Ultra Violet |

# 1 INTRODUCTION

Many studies have already covered the effects of space conditions on micro-organisms. Mostly however, the results of these experiments have been inconclusive and/or contradictory. Observations as numerous and variable like increased cell density, changes in colony perimeter, growth reduction, biomass increase, increased resistance to minimum inhibitory concentrations of antibiotics, no effect on transformation or transduction, increased conjugation and increased resistance to high doses of radiation have been put forward. Nevertheless, it has been suggested that more extensive study needs to be applied to obtain reliable and conclusive evidence on the effects caused by space related environmental conditions on micro-organisms at the molecular, proteomic and genetic level.

Not only space related stress conditions will put pressure on the metabolism of the MELiSSA organisms, also the influence of processing conditions will have an effect on microbial metabolism. Heat stress caused by effluent spill of the first thermophilic compartment (run at 50°C) to the next (run at 30°C), shear stress, oxygen-, light- and nutrient-deprivation, supernatant conditions, accumulation of recalcitrant and even toxic intermediate metabolites and long-time continuous culturing will all have an influence on the health condition of the bacteria. Furthermore can the latter types of stresses even be mutagenic. It has been observed that prolonged mild temperature stress can mutagenise certain strains. Starvation conditions can likewise induce mutants to appear in single species populations.

Therefore, it is important to investigate the effect of stress on reactor stability, microbial condition and the nutritional quality of the food source. In addition it will be necessary to be able to detect stress as soon as possible for immediate intervention when this is required. Stress has already proved to be detectable in the MELiSSA organisms with a variety of methods.

In this technical note the methods to detect stress at the cellular, the proteomic, and the genetic level are validated and optimized.

# 2 DGGE ANALYSIS ON SLUDGE SAMPLES OF COMPARTMENT 1

## *2.1 Introduction*

Denatured Gradient Gel Electrophoresis (DGGE) analyses were performed to follow the evolution in the microbial population of compartment I and to compare it with population dynamics in an anaerobic reactor in which methanogenesis is not inhibited, and henceforth not desirable in the MELiSSA loop. DGGE is one of the most common fingerprinting methods and a lot of experience is available at Vito. The method was suggested in the EWC project for follow-up of acidogenic and methanogenic populations.

## *2.2 Materials and Methods*

### 2.2.1 SAMPLING

EPAS was running two laboratory-scale membrane bioreactors. One of the reactors was acidified to inhibit methanogenesis and is representative for compartment I of the MELiSSA loop. The second one is operated under neutral conditions. Hence, anaerobic transformations include methane production. Both reactors were continuously stirred. They were fed with fecal material collected from 8 different persons between age 24 and 40. The organic load was equal to 0.9 g organic matter.$l^{-1}$.$d^{-1}$. The hydraulic residence time was approximately 20 d. Three times a week, 150 ml of waste material was fed into the reactors. The same volume of liquid was removed from the neutral reactor. In the acidified reactor, 90 ml was removed directly from the reactor, 60 ml was removed via the ultrafiltration unit connected to it. This is not an additional step. It is needed because C1 is a membrane bioreactor system. To remove accumulated solids, some reactor content was also removed directly (and not via the membrane filtration unit). Using the membrane filtration, allowed to quantify and estimate the effluent quality for CII.

At regular time intervals, 5-ml sludge samples from both bioreactors were taken, centrifuged, resuspended in 2 ml of a solution containing 15% glycerol and 0.85% NaCl, and then stored frozen pending analysis.

### 2.2.2 DNA EXTRACTION AND DGGE COMMUNITY ANALYSIS

Total DNA was isolated from the samples using the protocol reported previously by el Fantroussi *et al.* (el Fantroussi et al. 1999) and modified by Dr.F. Faber, University of Groningen, The Netherlands (personal communication). This protocol was modified as follows: 2 g of sludge suspension was added to 4 ml of the extraction buffer (10mM Tris-HCl pH7/15%glycerol). To this, 2 g of glassbeads (0.10-0.11 mm) were added. The

# ABREVIATIONS

| | |
|---|---|
| AFLP | Amplified Fragment Length Polymorphysm |
| BLAST | Basic Local Alignment Search Tool |
| CFDA | carboxy-fluorescein diacetate |
| CHAPS | 3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate |
| CV | Coefficient of Variation |
| DGGE | Denaturing Gradient Gel Electrophoresis |
| DioC(6)3 | 3,3'-dihexyloxacarbocyanine iodide |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acid |
| dNTPs | Deoxynucleotide triphosphates |
| EDTA | Ethylene-Diamine Tetra Acetic acid |
| EtBr | Ethidium Bromide (2,7-diamino-10-ethyl-9-phenyl-phenanthridium bromide) |
| EWC | Engineering of the Waste Compartment |
| FCM | Flow Cytometry |
| FDA | fluorescein diacetate |
| HE | hydroethidine |
| HPLC | High-Pressure (or High-Performance) Liquid Chromatography |
| ICM-MS | Intact cell MALDI_TOF Mass Spectrometry |
| MALDI-TOF | Matrix Assisted Laser Desorption Ionization-Time of Flight |
| MCP | Micro-Channel Plate detector |
| MOPS | 4-Morpholinepropanesulfonic acid |
| MS | Mass Spectrometry |
| NCTC | National Collection of Type Cultures (Colindale, UK) |
| OD | Optical Density |
| OPA | One Phor All buffer |
| ORF | Open Reading Frame |
| PAGE | Pulsed Amplified Gel Electrophoresis |
| PCR | Polymerase Chain Reaction |
| PI | Propidium Iodide |
| PPM | parts per million |
| rDNA | rRNA encoding DNA |
| Rho-123 | Rhodamine-123 |
| RDP-II | Ribosomal Database Project (version 2) |
| RPM | Revolutions Per Minute |
| SAPD | Surface Accessible Proteins Detection |
| SCM-MS | Supernatant Culture MALDI-TOF MS |
| SDS | sodium dodecyl sulfate |
| TAE | Tris-Acetate-EDTA |
| TBE | Tris-Borate-EDTA |
| TCA | TrisChloroacetic Acid |

| | |
|---|---|
| TE | Tris EDTA |
| TEM | Transmission Electron Microscopy |
| TLF | Time Lag Focusing |
| Tris | Tris(hydroxymethyl)-aminomethane |
| UV | Ultra Violet |

# 1 INTRODUCTION

Many studies have already covered the effects of space conditions on micro-organisms. Mostly however, the results of these experiments have been inconclusive and/or contradictory. Observations as numerous and variable like increased cell density, changes in colony perimeter, growth reduction, biomass increase, increased resistance to minimum inhibitory concentrations of antibiotics, no effect on transformation or transduction, increased conjugation and increased resistance to high doses of radiation have been put forward. Nevertheless, it has been suggested that more extensive study needs to be applied to obtain reliable and conclusive evidence on the effects caused by space related environmental conditions on micro-organisms at the molecular, proteomic and genetic level.

Not only space related stress conditions will put pressure on the metabolism of the MELiSSA organisms, also the influence of processing conditions will have an effect on microbial metabolism. Heat stress caused by effluent spill of the first thermophilic compartment (run at 50°C) to the next (run at 30°C), shear stress, oxygen-, light- and nutrient-deprivation, supernatant conditions, accumulation of recalcitrant and even toxic intermediate metabolites and long-time continuous culturing will all have an influence on the health condition of the bacteria. Furthermore can the latter types of stresses even be mutagenic. It has been observed that prolonged mild temperature stress can mutagenise certain strains. Starvation conditions can likewise induce mutants to appear in single species populations.

Therefore, it is important to investigate the effect of stress on reactor stability, microbial condition and the nutritional quality of the food source. In addition it will be necessary to be able to detect stress as soon as possible for immediate intervention when this is required. Stress has already proved to be detectable in the MELiSSA organisms with a variety of methods.

In this technical note the methods to detect stress at the cellular, the proteomic, and the genetic level are validated and optimized.

# 2 DGGE ANALYSIS ON SLUDGE SAMPLES OF COMPARTMENT 1

## 2.1 Introduction

Denatured Gradient Gel Electrophoresis (DGGE) analyses were performed to follow the evolution in the microbial population of compartment I and to compare it with population dynamics in an anaerobic reactor in which methanogenesis is not inhibited, and henceforth not desirable in the MELiSSA loop. DGGE is one of the most common fingerprinting methods and a lot of experience is available at Vito. The method was suggested in the EWC project for follow-up of acidogenic and methanogenic populations.

## 2.2 Materials and Methods

### 2.2.1 SAMPLING

EPAS was running two laboratory-scale membrane bioreactors. One of the reactors was acidified to inhibit methanogenesis and is representative for compartment I of the MELiSSA loop. The second one is operated under neutral conditions. Hence, anaerobic transformations include methane production. Both reactors were continuously stirred. They were fed with fecal material collected from 8 different persons between age 24 and 40. The organic load was equal to 0.9 g organic matter.$l^{-1}.d^{-1}$. The hydraulic residence time was approximately 20 d. Three times a week, 150 ml of waste material was fed into the reactors. The same volume of liquid was removed from the neutral reactor. In the acidified reactor, 90 ml was removed directly from the reactor, 60 ml was removed via the ultrafiltration unit connected to it. This is not an additional step. It is needed because C1 is a membrane bioreactor system. To remove accumulated solids, some reactor content was also removed directly (and not via the membrane filtration unit). Using the membrane filtration, allowed to quantify and estimate the effluent quality for CII.

At regular time intervals, 5-ml sludge samples from both bioreactors were taken, centrifuged, resuspended in 2 ml of a solution containing 15% glycerol and 0.85% NaCl, and then stored frozen pending analysis.

### 2.2.2 DNA EXTRACTION AND DGGE COMMUNITY ANALYSIS

Total DNA was isolated from the samples using the protocol reported previously by el Fantroussi *et al.* (el Fantroussi et al. 1999) and modified by Dr.F. Faber, University of Groningen, The Netherlands (personal communication). This protocol was modified as follows: 2 g of sludge suspension was added to 4 ml of the extraction buffer (10mM Tris-HCl pH7/15%glycerol). To this, 2 g of glassbeads (0.10-0.11 mm) were added. The

mixtures were beaten 2 times for 30 seconds using a bead beater (B.Braun Biotech International, Melsungen, Germany). After this, lysosyme (Invitrogen) was added to a final concentration of 2mg/ml, followed by an incubation of the samples at 37°C for 30 min on a rotary mixer. To achieve a complete lysis 0.6% SDS and 160µg/ml proteinase K ( Invitrogen) was added and incubated at 50°C for 30 min on a rotary mixer. After that, 2 ml of a 120 mM $Na_2HPO_4.2H_2O$/120 mM $KH_2PO_4$ pH8 was added and the samples were again beaten 2 times for 30 seconds. The supernatant was collected after centrifugation of the samples at 14000 rpm for 3 min, A phenol:chloroform:isoamylalcohol extraction (25:24:1) was applied, followed by two extractions with chloroform:isoamylalcohol (24:1). To the supernatant 100 mg polyvinylpyrrolidone (Aldrich) was added and left on ice for 30 min with gentle shaking from time to time. After centrifugation at 11000 rpm for 5 min at 4°C, the supernatant was precipitated overnight at -20°C with 2 volumes 100% ethanol and 0.1 volume 3 M sodiumacetate pH 5.2. After centrifugation at 11000 rpm for 5 min at 4°C the pellet was washed with 70% ethanol and vacuum dried. The pellet was resuspended in 400 µl milliQ-$H_2O$ and the DNA was purified using Wizard DNA clean-up system (Promega). For the PCR reaction, 10 ng of the purified DNA was used as template.

A 455 bp eubacterial 16S rDNA from the extracted DNA was amplified by PCR as described by Marchesi *et al* (1998) using the forward primer 63F (Marchesi *et al*, 1998) and the reverse primer 518R (Felske et al, 1996) in order to study the bacterial community composition. Amplification reactions were performed in a total volume of 100 µl. The reaction mixture contained 10 µl of 10x *Ex Taq*™ reaction buffer (TaKaRa, BioWhittaker Europe), 200 µM of each deoxynucleotide triphosphate (BioWhittaker Europe), 250 nM concentration of each primer, 10 µl of template DNA and 2.5 U of *Ex Taq*™ polymerase (TaKaRa, BioWhittaker Europe). Amplifications were carried out in a 9600 GeneAmp PCR system (PE Biosystems). The following thermocycling program was used: 95°C for 3 min (1 cycle), 95°C for 1 min, 55°C for 1 min., 65°C for 1 min. (40 cycli) and 65°C for 8 min (1 cycle). The PCR reaction was controlled with agarose gel electrophoresis at 85 V. The 16S rDNA amplicons were separated by DGGE (Muyzer et al, 1993) with a 35 to 65% denaturant gradient on a 8% acrylamide gel in 1 x TAE (40 mM Tris-acetate [pH 7.4], 20 mM sodium acetate, 1 mM EDTA). DGGE gels were run in an Ingeny phorU for 15 h at 120 V using a PS 304 Electrophoresis Power Supply from Life Technologies and then were stained with Sybr Green (Molecular Probes). Resolved PCR products were visualised by UV transillumination using an Image Master® VDS from Pharmacia Biotech. Pictures from the gel were taken with a Fuji Film Thermal Imaging System FTI-500. The digitised images were analysed with Bionumerics software. Cluster analysis was performed with Jaccard.

## 2.3 *Results and Discussion*

The number of samples analysed so far, is rather low. Therefore, the present analysis can only give an indication on the stability of the community. A thorough investigation of the community stability is foreseen in the MELiSSA contract 'MELiSSA - Engineering of the

Waste Compartment' EWC with contract number 15689/01/NL/ND in the Technical notes: 71.9.3, 71.9.4, 71.10.3 and 71.10.4.. Because the prototype Compartment I reactor is operational just now, this community information will become available in the coming months. As indicated in Figure 2.1, the microbial communities in both laboratory reactors change over time. However, some bands remain present throughout the experimental period and may relate to dominating populations.

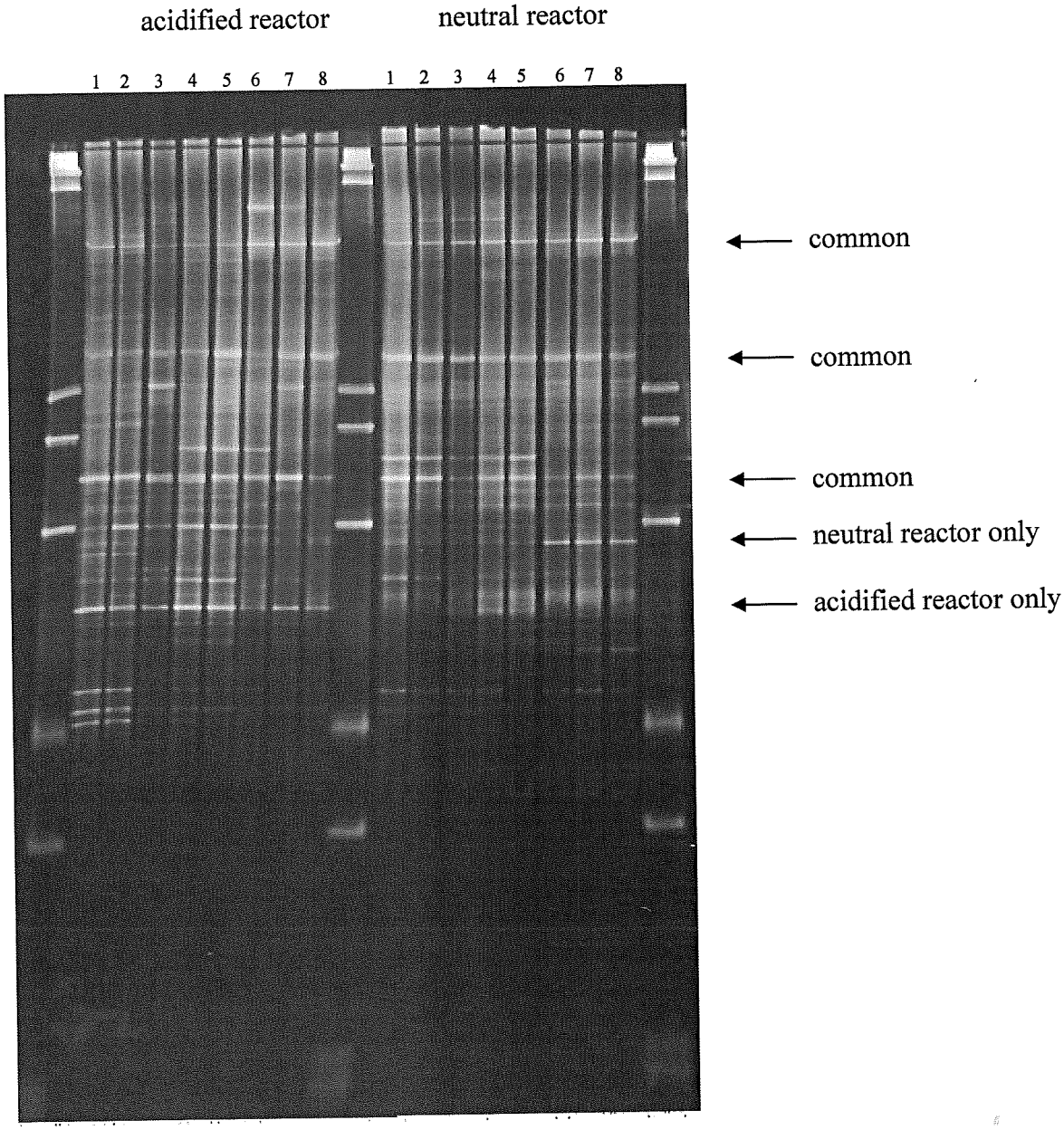



***Figure 2.1 DGGE fingerprints of sludge samples from the acidified reactor (numbered 1 to 8) and the neutral reactor (numbered 1 to 8). The samples are ordered according to sampling time.***

The banding pattern was generated with a universal primer. Hence, the fingerprints reflect the composition of the acidifying bacterial populations. Methanogenic bacteria are not visualized. This will be done at a later stage when methanogenic primers and the relevant reference strains are available.

The intensity of the bands does not correspond with the proportion of this particular species or group of species in the microbial community. This is due to the fact that preferential DNA amplification may occur for certain bacteria in the PCR step.

Some bands are present in both reactors, while others are specific for either the acidified reactor or the neutral one. This is indicated on *Figure 2.1*. Both the common and specific bands are worth further investigations. Identification of the underlying species will learn:

- which species appear when acid conditions prevail
- which ones will disappear in acidic conditions
- which ones are omnipresent independent of pH regulation.

Some bands which are only present in one of both reactors occur for a while and then disappear again. These are probably less crucial for reactor operation. Attempts to sequence and hence identify interesting bands have not yet been performed. This requires a larger set of fingerpints over longer periods of time.

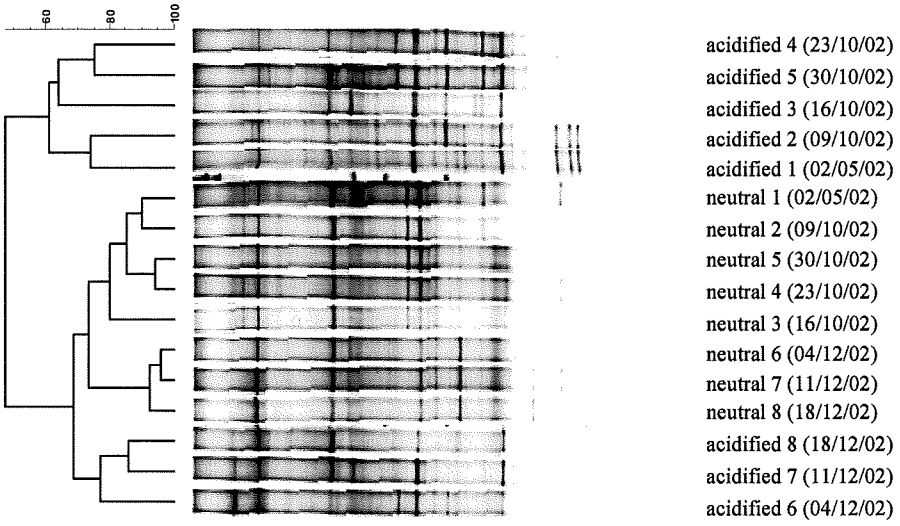



***Figure 2.2 Dendrogram of the reactor fingerprints, clustered with commercial software.***

When the fingerprints from both reactors are regrouped according to their degree of similarity (see Figure *2.2*.), the following observations can be made:

- there are 2 clusters: a top one consisting of 5 samples from the acidified reactor, and a bottom one with the other samples
- the first samples from the acidified and neutral reactor are expected to be more related to each other than samples taken at a later time point. However, the distance between these 2 samples and hence their degree of similarity is low. As a function of time, both reactor communities seem to converge to each other. E.g. samples taken in December show a higher similarity than the first sample. This is a surprising result. Although at that time, there was no full access to operational data of the reactors. Currently samples are evaluated over a much longer period of time and this tendency is no longer there. This will be reported in EWC and can be found in the undergraduate thesis which was sent to Christophe.
- The 3 sampling points in December are closely related per reactor. This seems to imply that a fairly stable community composition has been reached and few changes occur. However, the time between sampling is close to the hydraulic residence time (20 d) and this may explain why similarities are high. However, similarities between samples 1 and 2 are higher than the ones between samples 2 and 3 in spite of the fact that the time between sampling was much longer in the former case. We evidently need more long-term analyses to confirm if/when the community stabilized.
- Samples are generally grouped per reactor. The operating conditions or in this case the environmental conditions seem to determine the composition of the acidifying community (methanogens not detected with the primers used)

It is important to link these observations to the general reactor operating conditions (type of substrate, applied load, transformation efficiencies achieved, continuity in operation, .. ). At present, the number of samples analysed is too low to do that in a sound way. More samples have been taken and stored and will be analysed shortly.

Taking into account that the methanogenesis is the only difference in the process of those two reactors, it seems natural that this population needs to be investigated. However, methanogens are not detected because another primer and methodology was needed to do so. At the time of
reporting these were not yet available. The methodology is now in the process of testing. The actual tests will be performed soon in the framework of the EWC contract.

## 2.4 References

J.R. Marchesi, T. Sato, A.J. Weightman, T.A. Martin, J.C. Fry, S.J. Hiom, and W.G. Wade. 1998. Design and Evaluation of Useful Bacterium-Specific PCR Primers That Amplify Genes Coding for Bacterial 16S rRNA. Appl. Envir. Microbiol. 64: 795-799.

G. Muyzer, E.C. de Waal, and A.G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Envir. Microbiol. 59: 695-700.

A. Felske, A.D.L. Akkermans, and W.M. De Vos. 1998. Quantification of 16S rRNAs in Complex Bacterial Communities by Multiple Competitive Reverse Transcription-PCR in Temperature Gradient Gel Electrophoresis Fingerprints. Appl. Envir. Microbiol. 64: 4581-4587.

S. el Fantroussi, L. Verschuere, W. Verstraete, and E.M. Top. 1999. Effect of Phenylurea Herbicides on Soil Microbial Communities Estimated by Analysis of 16S rRNA Gene Fingerprints and Community-Level Physiological Profiles. Appl. Envir. Microbiol. 65: 982-988.

# 3 OPTIMISATION OF THE DETECTION OF PHYSIOLOGICAL CHANGES OF BACTERIA INDUCED BY STRESS CONDITIONS

## 3.1 Introduction

In the present chapter, we detail the different methods that are important for the analysis by flow cytometry of physiological changes induced by temperature and $H_2O_2$ stress. The following protocols concern the estimation of the DNA content, the membrane permeability and potential, the esterase activity, the intracellular pH and the superoxide anion production in the MELISSA bacterial strains.

## 3.2 Chemicals

Propidium iodide (PI), fluorescein diacetate (FDA), carboxy-fluorescein diacetate (CFDA) and rhodamine-123 (Rho-123) were purchased from Sigma, 3,3'-dihexyloxacarbocyanine iodide (DioC(6)3) and hydroethidine (HE) from Molecular Probes. Stock solutions of each dye were prepared as follows : PI was made up at 50 µg/ml in filtered isoton II (Beckman-Coulter) and conserved at 4°C in the dark. FDA was dissolved to 2 mg/ml in acetone and maintained at -20°C. CFDA was dissolved at 4 mg/ml in DMSO. Rho-123 was made up to 1 mg/ml in ethanol and maintained at -20°C. DioC(6)3 stock solution was 400 µM in ethanol and stored at -20°C. HE stock solution was 10 mM in DMSO and stored at -20°C. The working concentrations of PI, FDA, CFDA, Rho-123, DioC(6)3 and HE were 40 µg/ml, 10 µg/ml, 0.28 µg/ml, 10 µg/ml, $8.10^{-6}$ µmol/ml and 0,01 µmol/ml, respectively. All the working solutions were prepared freshly on the day of the experiments.

## 3.3 Stress conditions

*Induction of temperature stress* : A 100 µl aliquot of bacterial culture was transferred to a 5 ml polypropylene tube (Becton Dickinson) used for flow cytometrical purposes. For temperature stress, samples were incubated at one of those temperatures for one hour. The temperature range was -170°C (liquid nitrogen), -80°C, -20°C, 4°C (ice bath), 15°C, 28°C, 37°C, 45°C, 50°C, 60°C and 70°C. A wide temperature range was chosen in order to find temperatures that could cause physiological damage. Controls were obtained by incubating bacteria at 28°C. After the temperature stress, a volume of 400 µl of one of the stains (PI, FDA, CFDA, Rho-123, DioC(6)3 or HE) at optimal concentrations described above was added at room temperature in the dark. Measurements with the flow cytometer were performed within an hour.

*Induction to oxidative stress* : A 90 µl aliquot of bacterial culture was transferred to a 5 ml polypropylene tube (Becton Dickinson) used for flow cytometrical purposes. Oxidative

stress was generated by addition of hydrogen peroxide ($H_2O_2$ 30%, Merck) for an hour at the following final concentrations : 13,25; 27,5; 55; 110; 220; 440 and 880 mM. 10 µl of hydrogen peroxide (10 times more concentrated than the different concentrations tested) was added to the bacterial suspension and samples were incubated at room temperature for an hour. Controls were obtained by incubating bacteria without any hydrogen peroxide. After the one-hour incubation in the presence of hydrogen peroxide, a volume of 400 µl with one of the stains (PI, FDA, CFDA, Rho-123, DioC(6)3 or HE) at optimal concentrations described above was added at room temperature in the dark. Measurements with the flow cytometer were performed within an hour.

## *3.4 Flow cytometric measurements*

Flow cytometry was carried out using a Coulter Epics XL flow cytometer with an air-cooled argon ion laser, 15 mW output and a fixed wavelength emission of 488 nm. It has four fluorescence detection channels which simultaneously detect green, yellow-orange, and red light emission. This standard instrument is equipped with two light scatter detectors that measure forward (FSC) and side scatter (SSC) and four fluorescence detectors that detect appropriately filtered light at green (FL1, 525 nm; for FITC, DioC(6)3, Rho-123), orange (FL3, 620 nm; for PI), and red (FL4, 675 nm; for red fluorescence of HE) wavelengths. In addition to the log fluorescence parameter, log forward and side scatters were monitored, which are indicators of cell size and granularity, respectively. Software discriminators were set on forward side scatter (FS) signals to eliminate electronic and small particle (originating from the media, buffers or sheath fluid) noise. A total of 10,000 bacteria was recorded for each sample and all experiments were conducted in triplicate. Data were stored as listmode files and analysed off-line using the System II software (Beckman-Coulter).

For temperature stress, bacteria incubated at 28°C were used as negative controls and cells treated at -170°C and 70°C were used as positive controls for voltage and gain adjustments and assignment of histogram regions.

For oxidative stress, bacteria incubated in the absence of hydrogen peroxide were used as negative controls and cells treated with the highest concentration of hydrogen peroxide (880 mM) were used as positive controls for voltage and gain adjustments and assignment of histogram regions.

For analyses of the collected data, the discrimination between the bacterial population and the background was set with the help of the positive and negative controls and kept constant for all samples. Fluorescence data for cell suspensions were normalized and expressed as percentages of the value obtained for cells incubated at 28°C.

## 3.5 *Use of flow cytometry to check DNA content in the MELiSSA strains*

Flow cytometry is suited for cell cycle analysis, since information about different cell cycle stages can be obtained without the need for synchronization of the cell culture.

### 3.5.1 PREPARATION OF SAMPLES TO ANALYSE DNA CONTENT FROM BACTERIAL CULTURES

*Retention of fully replicated chromosomes.* To obtain bacteria that contain only fully replicated chromosomes, we added rifampicin (150 μg/ml) to the culture which is grown to an optical density of 0.1 to 0.3. The molecular mechanism of rifampicin is an inhibition of bacterial RNA polymerase that forms a complex with rifampicin. Rifampicin prevents the ribosomes from binding to messenger RNA, and thus proteins are not produced.

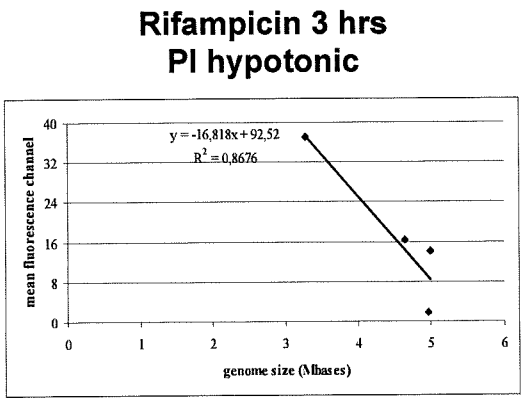



**Fig. 3.1. Correlation between the genome size and the mean fluorescence of propidium iodide (PI). Bacteria were treated for 3 hours with rifampicin and then stained with hypotonic PI.**

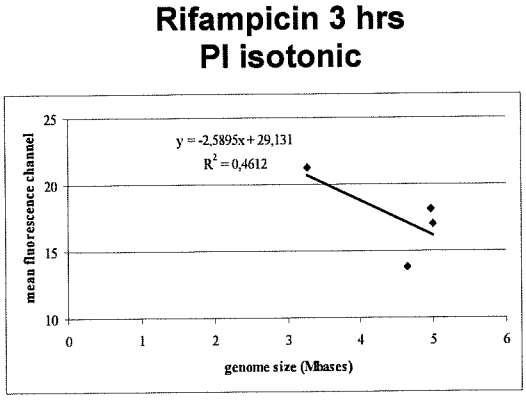



**Fig. 3.2. Correlation between the genome size and the mean fluorescence of propidium iodide (PI). Bacteria were treated for 3 hours with rifampicin and then stained with isotonic PI.**

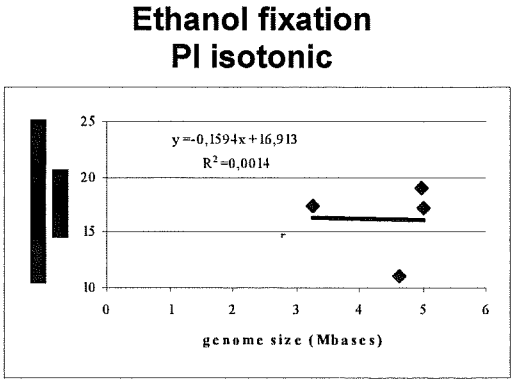



**Fig. 3.3. Correlation between the the genome size and the mean fluorescence of propidium iodide (PI). Bacteria were treated for 3 hours with rifampicin, fixed with 80% ethanol and then stained with isotonic PI.**

In our study, cultures were preincubated for 3 hours in the presence of rifampicin in order to allow complete runout of replication, such that all cells contain only fully replicated chromosomes.

*DNA staining.* Samples were chilled in 1 ml of chilled PBS buffer pelleted and washed in Tris buffer, pelleted again, and resuspended in the same buffer before an eventual staining with isotonic or hypotonic propidium iodide at 50 µg/ml. The effect of a preliminary fixation in 80% cold ethanol during vortexing was also tested. Different bacterial strains (*Escherichia coli 4.64 Mbases, Ralstonia metallidurans : 5 Mbases, S. oneidensis : 4.97 Mbases* and *Deionococcus radiodurans 3.28 Mbases*) were compared.

*RNAse treatment.* Propidium iodide binds double strain DNA and RNA. Therefore, some authors sometimes treat the samples with RNase. However, some authors do not include any RNAse treatment. In this study, we did not use the RNAse treatment.

As seen in Figures 3.1., 3.2., 3.3., the best results were obtained with a preliminary treatment of 3 hours of rifampicin in hypotonic PI as shown by the best correlation value (r value).

These tests consist of a first approach in order to measure the DNA content. A complete analysis in triplicates comparing fixation method and time of incubation with rifampicin is still needed. Indeed, the time of incubation of 3 hours with rifampicin and its concentration were chosen arbitrarily from the literature on experiments performed on *E. coli*. However, those parameters have not yet been optimised for the three other strains (*R. metallidurans, S. oneidensis and D. radiodurans*). Therefore, in a first step, rifampicin incubation time and concentration will be adapted for the three cell lines. Furthermore, a comparison will have to be made on various DNA stains. Propidium iodide stains well DNA of all bacterial cells but the fluorescence signals were significantly wider (spread on two fluorescence decades) than the ones observed in the literature with the resulting disadvantage that there was an overlap of fluorescence distribution with no sharp

fluorescence distribution. In the future, a special effort will be made on optimizing experimental conditions of rifampicin and propidium iodide (or another DNA stain).

## *3.6 Membrane permeability estimated by propidium iodide*

The loss of membrane integrity represents significant damage for cells due to multiple functions linked to the plasma membrane (permeability barrier, transport, respiratory activity, etc...). The maintenance of membrane integrity is commonly measured in eukaryotic cells as an indicator of cell damage or cell death. All microbial cells are bounded by the cytoplasmic membrane, which allows the cell to communicate selectively with its immediate environment. Passive and active transport systems across the membrane generate an electrochemical gradient. Outside of the cytoplasmic membrane, gram-negative bacteria have an additional outer membrane. Its function is mainly protective because it prevents the entry of toxic substances into the cell. It is less selective and more permeable than the cytoplasmic membrane. An intact polarized cytoplasmic membrane and active transport systems are essential for a fully functional, healthy cell. Few fluorescent dyes used in flow cytometry can cross intact cell membranes. Most of the membrane integrity assays use nucleic acid stains (for example propidium iodide or PI in this work), due to the high concentrations of nucleic acids within the cells and the large fluorescence enhancement exhibited by nucleic acid stains upon binding, leading to a clear separation between intact and dead cells. PI is a strongly hydrophilic, small molecule ($M_r$, 660). Only cells that have lost selective permeability due to the temperature treatment take up the dye, which stains nucleic acids. PI-positive cells fluoresce and can be quantified using flow cytometry.

Propidium iodide is an impermeant compound that markedly increases its fluorescence on binding to double-stranded nucleic acid. In this study, different concentrations (referred in the literature) of propidium iodide (PI) were compared : 0; 5; 10; 20; 40; 50; 60; 75; 92 µg/ml. *R. metallidurans* submitted to an hour stress at 70°C was used as a reference in order to test the different PI concentrations. As shown in figure 3.4., the staining was optimal from 40 µg/ml PI onwards. Indeed, from 40 µg/ml onwards, a plateau was reached which means that an equilibrium was reached between the dye and the amount of DNA. Therefore, the concentration of 40 µg/ml was chosen for all the experiments performed in order to study the effect of temperature, oxidative and pH stress.

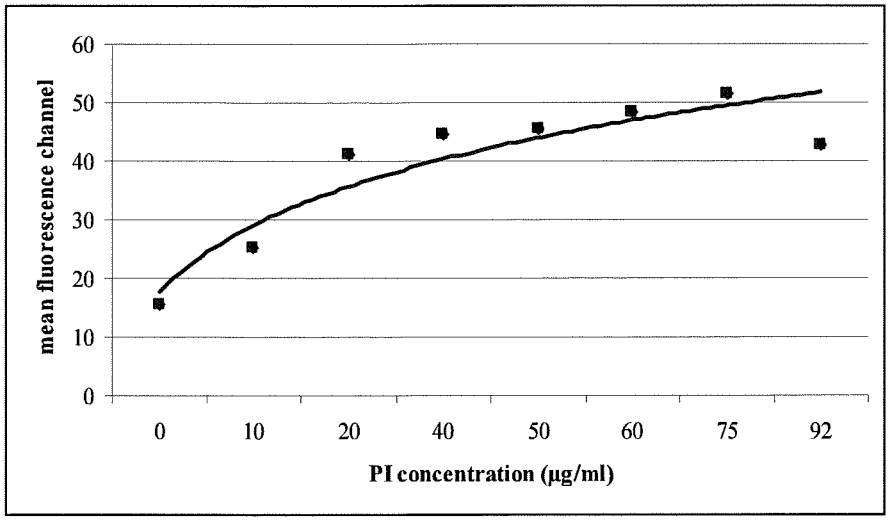



**Fig. 3.4. Optimisation of PI concentration : correlation between the PI concentration and the mean fluorescence channel. A plateau is obtained from 40 µg/ml.**

## *3.7 Membrane potential estimated by DioC(6)3 and Rhodamine-123*

The electrochemical potential occurring through the plasma membrane of metabolising bacteria is generated by respiration or by ATP hydrolysis. It results from the selective permeability of biological membranes to a variety of cations and anions leading to a difference of electric potential across the membrane. Inside, the cell is negatively charged compared with the outside of the cell, and membrane potential plays a central role in different cell-life processes (ATP synthesis, active transport, mobility, regulation of intracellular pH, etc.). Voltage-sensitive dyes have been developed to measure membrane potential in bacteria. Depending on the charge of the dye, they are accumulated in polarised (cationic dyes) or depolarised (anionic dyes) cells. In appropriate conditions, the amount of dye taken up can be directly related to the level of energy metabolism in the cell. DioC(6)3 and Rhodamine 123 (Rho-123) are lipophilic, cationic dyes commonly used to detect membrane potential. They can pass cell membranes, but are only retained in cells with a membrane potential (membrane potential negative inside), the fluorescence level of which is determined by the magnitude of the membrane potential (Diaper et al, 1992). Accumulation inside bacterial cells is favoured by a reduction in the magnitude of the membrane potential, allowing dye molecules to concentrate within the cell, and bind to lipid-rich components. Rho-123 and DioC(6)3 have been reported to be useful to detect depolarised cells of numerous Gram-positive and Gram-negative bacterial species. DioC(6)3 is widely used as a viability stain. Bacteria with a membrane potential exclude the dye, but non-viable bacteria with depolarized membranes allow it to enter the cell. Rho-123 is a voltage sensitive cationic dye that is electrophoretically taken up into energized bacteria by virtue of the trans-membrane electrochemical potential (negative inside) of the plasma membrane.

Different concentrations referred in the literature (1; 2; 5; 10; 20 and 50 µg/ml) of Rho-123 were tested on *R. metallidurans* submitted to an hour stress at 50°C or in the presence of 880 mM H2O2. These results were compared to control conditions (*R. metallidurans*) without being submitted to a stress. In our conditions, 10 µg/ml was the optimal concentration since from that concentration, the plateau was reached (figure 3.5.).

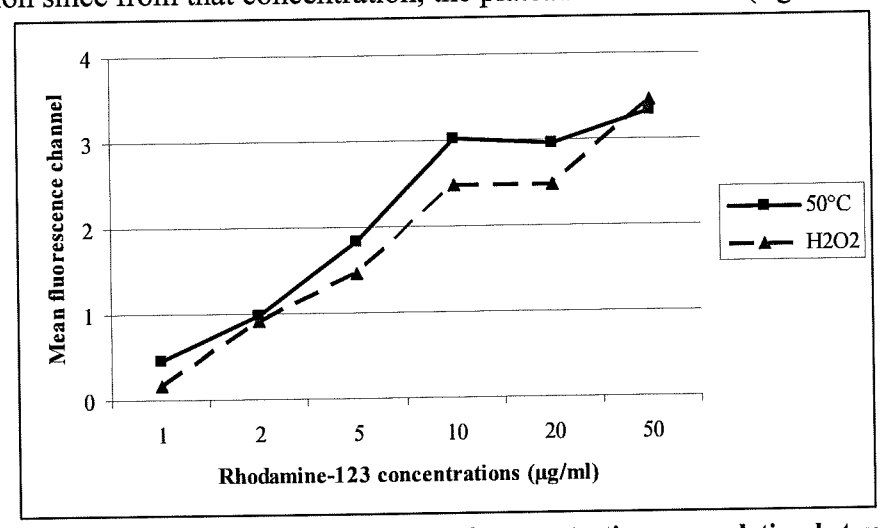



**Fig. 3.5. Optimisation of Rhodamine-123 concentration : correlation between the Rho-123 concentration and the mean fluorescence channel. A plateau is reached from 10 µg/ml.**

Different concentrations referred in the literature (8; 16; 40; 164 10-6 µmol/ml) of DioC(6)3 were tested on *R. metallidurans* submitted to an hour stress at 50°C or in the presence of 880 mM H2O2. In our conditions, 8x10-6 µmol/ml was the optimal concentration since in this case, the difference between control and stressed conditions was the most visible at that concentration (figure 3.6.).

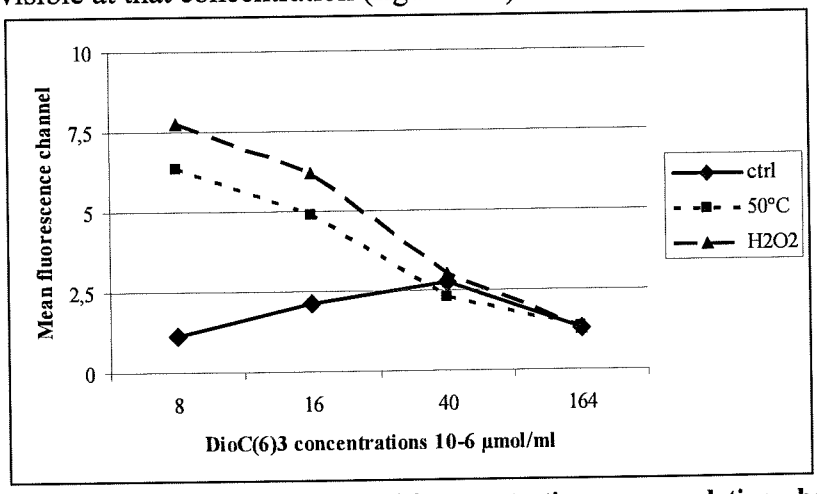



**Fig. 3.6. Optimisation of DioC(6)3 concentration : correlation between the DioC(6)3 concentration and the mean fluorescence channel. A plateau is reached from 8x10-6µmol/ml.**

## 3.8 *Esterase activity estimated by fluorescein diacetate*

Detection of esterase activity is measured using lipophilic, uncharged and non-fluorescent fluorogenic substrate (here fluorescein diacetate). Once within active cells, the substrate is cleaved by non-specific esterases (lipases, acylases but not acetylcholinesterases) releasing two acetates as well as a polar fluorescent product (fluorescein) that is known to be well-retained inside cells having an intact membrane. The dye rapidly leaks from cells with compromised membranes. Esterases are present in all living organisms, and these enzymes can be used to provide information on the metabolic state of bacterial cells. Although enzyme synthesis requires energy, the enzyme-substrate reaction does not, and this assay alone should be considered energy independent. However, dead or dying cells with damaged membranes rapidly leak the dye, even if they retain some residual esterase activity. Consequently, fluorogenic substrates for esterases often serve as activity and cell integrity probes that measure both enzymatic activity, which is required to activate their fluorescence, and cell-membrane integrity, which is required for intracellular retention of their fluorescent products and can therefore, help differentiating between live and dead/unhealthy cells.

Different concentrations referred in the literature (5, 10, 25, 50 and 100 µg/ml) of fluorescein diacetate were tested on *R. metallidurans* submitted to an hour stress at 50°C or in the presence of 880 mM H2O2. In our conditions, the pattern of fluorescence showed a kind of plateau from 10 µg/ml onwards even if the fluorescence intensity varied in function of the stress applied to *R. metallidurans*. In order to make a choice on the best concentrations of FDA, the three conditions were compared and it appeared that 10 µg/ml showed the best fluorescence intensity and the lowest fluorescence variation. Therefore, that concentration was chosen for the next experiments.

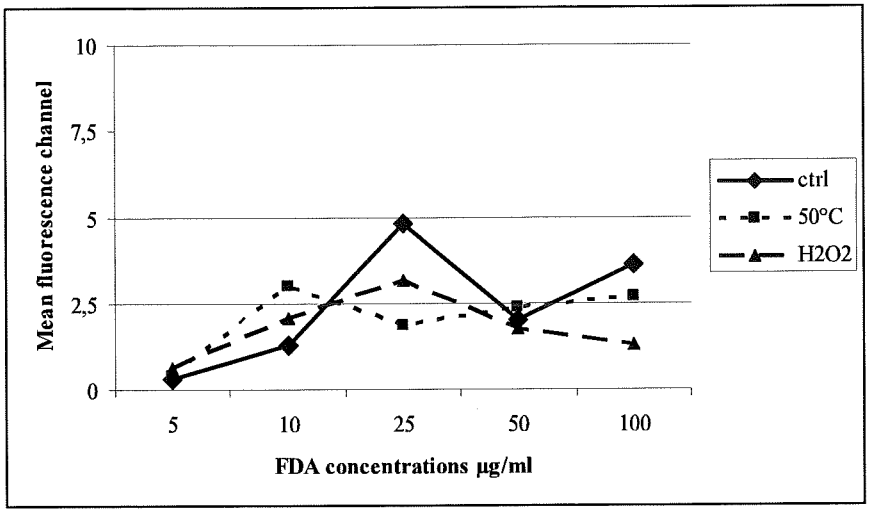



**Fig. 3.7. Optimisation of FDA concentration : correlation between the FDA concentration and the mean fluorescence channel. A plateau is reached from 10 µg/ml.**

## 3.9 *Intracellular pH estimated by carboxyfluorescein diacetate*

The intracellular pH ($pH_{in}$) of bacteria is critical for the control of many cellular processes, such as DNA transcription, protein synthesis, and enzyme activities. To study regulation and maintenance of the $pH_{in}$ a reliable method to measure the $pH_{in}$ is of outmost importance. Currently, one of the most commonly used fluorescent probes for $pH_{in}$ measurements is carboxy-fluorescein diacetate, succinimidyl ester (5(6)-CFDAse) which is well suited as it is highly pH dependent and has a fast response time (Haugland, 1992). Viable and culturable bacteria are able to keep their pH constant when they are incubated at pH = 3 whereas non-culturable cells can not. CFDAse couples irreversibly to both intracellular and cell-surface proteins by reaction with lysine side-chains and other available amine groups (Alvarez-Barrientos et al, 2000; Sincock et al, 2001).

Different concentrations referred in the literature (0.05; 0.1; 0.28; 0.56 and 1.12 µg/ml) of carboxyfluorescein diacetate were tested on *R. metallidurans* submitted to an hour stress in the presence of 880 mM H2O2. A comparison with temperature stress (after incubation at 50°C for an hour) was also performed. In our conditions, 0.28 µg/ml was the optimal concentration since from that concentration, a plateau was reached. Above 0.28 µg/ml in stress conditions, the fluorescence was not anymore optimal.

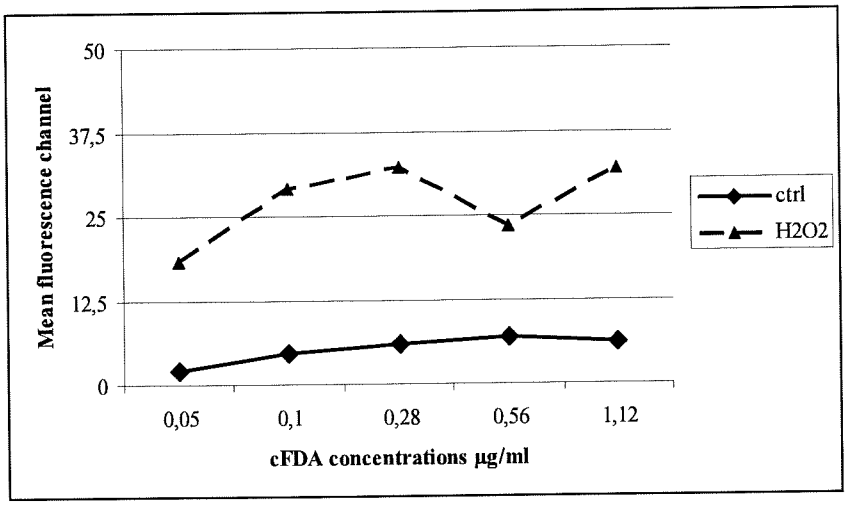



**Fig. 3.8. Optimisation of cFDA concentration : correlation between the cFDA concentration and the mean fluorescence channel. A plateau is reached from 0.28 µg/ml.**

## 3.10 *Superoxide anion production estimated by hydroethidine*

Reactive oxygen species (ROS/RNS), including a series of species including superoxide anion ($O_2^-$), hydrogen peroxide ($H_2O_2$), hydroxyl radical, nitric oxide, and peroxynitrite, have been implicated in the etiology of cell death. $O_2$- is usually considered a precursor of more reactive species to promote the production of secondary derivatives such as hydrogen peroxide and hydroxyl radical in intracellular oxidative chain events. Thus it is a

reasonable explanation that $O_2^-$, as a primary signaling molecule, promotes the production of ROS downstream, which act as signaling molecules further to cause cell death through a redox sensitive pathway. Furthermore, the redox sensitive signaling may interact with more classical signaling pathways such as transcription factor activation, gene expression, and cell proliferation.

Different concentrations referred in the literature (0.0016; 0.0025; 0.005; 0.01 µmol/ml) of hydroethidine were tested on *R. metallidurans* submitted to an hour stress in the presence of 880 mM H2O2. In our conditions, 0.01 µmol/ml was the optimal concentration since from that concentration, a plateau was reached (figure 3.9.). A comparison with temperature stress (after incubation at 50°C for an hour) was also performed. Even at the highest concentration taken from the literature (0.01 µmol/ml), a plateau was not really reached in control as well as in stress conditions. However, 0.01 µmol/ml was chosen as the optimal concentration since it showed the highest fluorescence intensity.

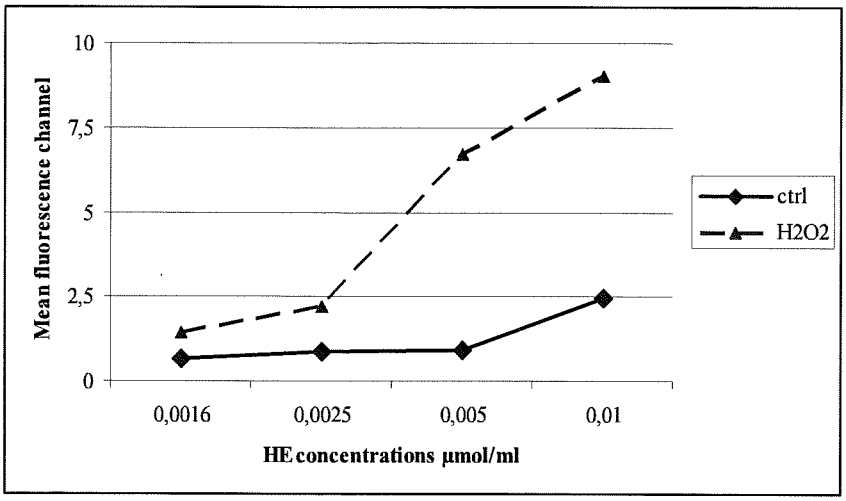



**Fig. 3.9. Optimisation of hydroethidine concentration : correlation between the hydroethidine concentration and the mean fluorescence channel. The concentration chosen is 0.01 µmol/ml.**

## *3.11 Conclusion*

*R. metallidurans* was our reference strain as a general comment in the MELiSSA study. Therefore, all our staining procedure and tests were optimized using that particular strain.

The concentrations of stains given in the following table were used to obtain optimal fluorescence intensities by flow cytometry. Those stains allowed us to precisely monitor variations in membrane permeability and potential, esterase activity, intracellular pH as well as the reactive oxygen species production.

The following concentrations need to be used in order to investigate the effect of stress using flow cytometry:

| Stain | Parameter | Concentration |
|---|---|---|
| Propidium iodide | Membrane permeability | 40 µg/ml |
| Fluorescein diacetate | Esterase activity | 10 µg/ml |
| Carboxy-fluorescein diacetate | Intracellular pH | 0.28 µg/ml |
| Rhodamine-123 | Membrane potential | 10 µg/ml |
| 3,3'-dihexyloxacarbocyanine iodide | Membrane potential | $8*10^{-6}$ µmol/ml |
| hydroethidine | Reactive oxygen species | 0.01 µmol/ml |

**Table 3.1. Recapitulative table taking into account the stains, the parameters measured and the chosen concentrations.**

# 4 WHOLE PROTEOMIC ANALYSIS FOR THE DETECTION OF STRESS: OPTIMISATION OF THE PROTOCOLS AND VALIDATION OF THE TOOLS USED

## *4.1 Introduction*

Wole-proteome analysis (a proteome being defined as the expressed protein complement of an entire genome), better known as 'proteomics', holds a key position in modern biology. This discipline recently emerged from the decades-long work on comprehensive protein visualization on two-dimensional gels, which was revitalized by developments in biological mass spectrometry and the growth in searchable sequences databases. Proteomics adds value to these databases by providing tools for the parallel separation and large-scale identification of proteins. Over the past ten years, mass spectrometry has become the technique of choice for protein or molecules characterization. The reason is the development of new methods for the ionization of proteins and peptides, especially matrix-assisted laser desorption-ionization and electrospray ionization. In the field of proteomics, the technique of MALDI-TOF MS is particularly suited to high throughput identifications of low femtomole level protein digest samples.

In this projet, a proteomic approach (based on two-dimensional gel electrophoresis separation and protein identification by mass spectrometry (mainly by MALDI-TOF) will be used to detect protein modifications during different stress conditions as temperature or light variation or oxidative stress and space conditions. It will also be used to identify surface membrane proteins mainly for the MELiSSA strains as *Arthrospira* sp. PCC8005 and *Rhodospirillum rubrum* ATCC25903.

## *4.2 Validation of the MALDI-TOF*

First, the MALDI-TOF has been tested to identify proteins separated by two-dimensional gel electrophoresis. This validation has been realized on *Ralstonia metallidurans*. In this context, the first two-dimensional gel electrophoresis database from narrow pH gradient containing about 1400 identified proteins has been built.

## *4.3 Materials and Methods*

### 4.3.1 BACTERIAL CULTURES

Mineral salts liquid medium, described in reference TN 70.9, supplemented with 0.2% (w/v) gluconate, was used to grow *R. metallidurans* CH34. Liquid precultures were grown

overnight in Wiame flasks at 30°C and shaken at 120 rpm on an orbital shaker. 0.5 ml of the precultures were used to inoculate a second set of flasks containing 25 ml of Schlegel medium supplemented with zinc sulfate (final concentration: 2mM), copper nitrate (final concentration: 0.85mM) or nickel chloride (final concentration: 2mM). Aliquots at log phase (OD600nm=0.3) and late plateau phase (OD600nm=0.7) were collected and centrifuged at room temperature at 6000 rpm for 10 min. Pellets were stored frozen at –80°C immediately after centrifugation.

### 4.3.2 SAMPLE PREPARATION

The bacterial pellets were pipette-suspended in lysis buffer (8M Urea, 4% (w/v) CHAPS, 40mM Tris, 0.2% (v/v) Pharmalyte 3-10, 2mM Tributyl Phosphine (Bio-Rad) , ¼ tablet/ml Complete Mini EDTA Free Protease Inhibitor Cocktail (Roche)) and further lysed for 1 minute in an Elma Transsonic 450/H sonicator at 4°C. The samples were treated at room temperature with DNase and RNase until they were no longer viscous, and then were centrifuged at 45000 rpm, 4°C for 1 hour. Protein concentration of the supernatants was measured by the Bradford method, according to the Bio-Rad Protein Assay kit, with Bovine gamma globulin as a protein standard. Supernatants were stored at –80°C.

### 4.3.3 2-D ELECTROPHORESIS

In the first dimension, each sample (200 µg for analytical gels and 1 mg for micropreparative gels, in a total volume of 150 µl) was subjected to isoelectric focusing in Immobiline DryStrips pH 4-7, 6-11, 3-10, 3.5-4.5, 4.5-5.5 and 5-6 (Amersham Pharmacia Biotech). The gels were rehydrated overnight in rehydration solution (2% (w/v) CHAPS, 8M Urea, 0.5% (v/v) Pharmalyte 3-10, 13mM DTE). Isoelectric focusing was performed on a Pharmacia Biotech Multiphor II system equipped with a Pharmacia Biotech EPS3500 XL power supply using a 3-phase program: the first phase was set at 500V, the second was a linear gradient spanning from 500 V to 3500 V and the final phase was set at 3500 V. Duration of each of these phases depends on the pI range of the strips used and was set according to the manufacturer's recommendations.

After isoelectric focusing, the gels were equilibrated two times, 20 minutes each, first in equilibration solution (6M Urea, 30% (v/v) glycerol, 2% (w/v) SDS, 50mM Tris-HCl pH 8.8) containing 65mM dithioerythritol, and second, in equilibration solution containing 135mM iodoacetamide. The strips were then placed on top of 12.5% SDS-PAGE gels in a 0.4% (w/v) agarose gel solution made with SDS-PAGE running buffer (192mM Glycine, 0.1% (w/v) SDS, 25mM Tris-HCl pH 8.3). The second dimension was run for approximately four hours at 500 V, 40 mA per gel, in a Bio-Rad Protean II Multicell gel system. Visualization of the protein spots was obtained by staining the gels with Coomassie Brilliant Blue G-250 or R-250, depending on the required sensitivity and the subsequent identification method chosen. Protein patterns within the gels were analysed as digitalized images using a high-resolution scanner in combination with the molecular analysis software PDQuest (Bio-Rad). Protein pIs and Mrs were approximated by

calibration of 2D gels using carbamylyte pI calibration markers and molecular weight standards (Amersham Pharmacia Biotech), respectively.

### 4.3.4 SAMPLE PREPARATION FOR MASS SPECTROMETRY ANALYSIS

Spots on the gel were excised using a 1mm sample corer (Fine Science Tools Inc.). Multiple cores were taken from weakly stained spots and, in some cases, identical protein spots were pooled from several gels. Excised gel pieces were placed in 1.5ml polypropylene Eppendorf tubes and washed twice in 50µl of 50mM NH4HCO3. The gel pieces were destained and dehydrated with 50µl of 50mM NH4HCO3, 50% (v/v) CH3CN and then dried in a centrifugal evaporator. Enzymatic digestion was performed with the addition of 10µl of 0.02µg/µl trypsin (Promega Madison, WI, USA) in 25mM NH4HCO3 to each gel piece, followed by incubation at 37°C overnight. For MALDI-MS, a 0.5µl aliquot of the digestion supernatant was spotted onto a sample plate with 0.5µl of matrix (5 mg/ml □-cyano-4-hydroxycinnamic acid and 0.5pmol/µl renin as internal standard in 25% (v/v) ethanol, 25% (v/v) acetonitrile, 0.05% (v/v) TFA) and allowed to air dry. For ESI-MS/MS, tryptic products were recovered by sequential extractions with 25mM NH4HCO3, 5% (v/v) formic acid and acetonitrile. Extracts were lyophilised in a vacuum concentrator and stored frozen until use.

### 4.3.5 MALDI-MS

MALDI-MS was performed using a Micromass M@ldi$^{TM}$ spectrometer (Manchester, UK) equipped with a 337nm nitrogen laser. The instrument was operated in the positive reflectron mode at 20kV accelerating voltage with time-lag focusing. Spectra were internally calibrated using the renin peptide, and the resulting peptide masses were automatically searched for in a local copy of the SWISS-PROT, *Ralstonia* databases using the ProteinLynx global server and the Protein Probe search engine (Micromass). An initial mass tolerance of 50ppm was used in all searches. Peptide modifications allowed during the search were carbamidomethylation of cysteines and oxidation of methionines. The maximum number of missed cleavages was set to 1.

### 4.3.6 ESI-MS/MS

Electro-Spray Ionization Mass Spectrometry (ESI-MS) and collision-induced dissociation (MS/MS) were performed with a Q-TOF 2 mass spectrometer (Micromass, Manchester) equipped with a Z-spray nano flow electrospray ion (nanoESI) source and a high-pressure collision cell. Samples were dissolved in 50% (v/v) acetonitrile, 5% (v/v) formic acid, and loaded into borosilicate nanoflow tips (Protana, Denmark). For MS/MS studies, the quadrupole was used to select the parent ions, which were subsequently fragmented in a hexapole collision cell using argon as collision gas and an appropriate collision energy (typically 20-35 eV). Data acquisition was performed with a MassLynx system based on Windows NT. MS/MS data were processed by a maximum entropy data enhancement

program, MaxEnt 3TM (Micromass, Manchester). Amino acid sequences were manually deduced with the assistance of Micromass' peptide sequencing program PepSeq.

### 4.3.7 ALGORITHMS FOR SEQUENCE SEARCHES AND ANALYSES

For some micro sequencing data the blastP algorithm was used to research against a protein database as SWISSPROT but the complete genome of *Ralsonia Metallidurans* was not include in this database. Thus, TblastX was used to research against a local genomic database of *R.Metallidurans*. The TblastX algorithm transforms the genomic database to protein database.
Therefore, searches for identity between microsequencing data and protein sequences were performed using the BLASTP algorithm on the 6-frame translated genomic sequences of *R. metallidurans*. Peptide sequences obtained via N-terminal sequencing or ESI-MS/MS were searched for in the *R. metallidurans* CH34 genomic database using TBLASTX. Prediction of signal peptide cleavage sites and subcellular localization was performed using the PSORT algorithm (http://psort.ims.u-tokyo.ac.jp/form.html;). Prediction of the role category of proteins was done according to TIGR (http://www.tigr.org/TIGRFAMs/;), matching annotations made on the genome of *R. solanacearum* pI and MW predictions were obtained using the "Compute pI/MW" tool from EXPASY (http://us.expasy.org/tools/pi_tool.html;).

## *4.4 Results*

### 4.4.1 DATABASE OF *R. METALLIDURANS* TOTAL CELLULAR PROTEINS

Total cellular proteins extracted from *R. metallidurans* CH34 grown in Schlegel medium were separated on 2-D gels covering different pI ranges. Using IPG strips of pH 3-10NL, 625 protein spots could be detected using colloidal Coomassie Blue. To increase the number of observed proteins, narrower pH gradients were used to separate proteins according to charge in the first dimension. In this condition, 531 Coomassie stained protein spots were detected from pH 4 to 7 and 568 from pH 6 to 11 (Figure 4.1). Protein identification was carried out on 352 different spots, chosen at random on your 2DE reference map, using a combination of automated Edman N-terminal microsequence analysis, peptide mass fingerprinting and tandem mass spectrometry. Peptide mass fingerprints were automatically searched against a local copy of the SWISS-PROT and *R. metallidurans* databases. Amino acid sequences obtained by N-terminal microsequencing or mass spectrometry were used to search for the draft genome database of *R. metallidurans* (on NCBI and PEDANT WWW sites) or non-redundant databases when no information could be obtained from the provisional genome of CH34. To date, the proteins from the 352 spots could be successfully identified and assigned to 224 genes of *R. metallidurans* spanning 126 of the total 712 contigs. These proteins are listed in Table 4.1 according to their corresponding spot number.

| Spot N° | Database Accession Number | Contig N° | Sequence Tag (E, T) / Sequence Coverage (M) | Identity of the closest homologue | Role Category | MW (KDa) P | MW (KDa) E | pI P | pI E |
|---|---|---|---|---|---|---|---|---|---|
| S0001 | ZP_00024248 | 662 | GIVDTLPAGILAGFPVVDVK, AEVPLSEMFGYSTALR | Elongation Factor G (EF-G) | Protein Synthesis | 77.5 | 77.4 | 5.24 | 5.27 |
| S0002 | ZP_00026369 | 710 | GLTFEVQQQLGDGVV, FDELSPSVDLLETGIK, YTLAGTEVSALLGR | ATP Synthase Beta Chain | Energy Metabolism | 51.0 | 47.3 | 5.25 | 5.24 |
| S0003 | ZP_00023573 | 630 | LNALSAGAEIIR | 5-Methyltetrahydropteroyltriglutamate-Homocysteine Methyltransferase | Amino Acid Biosynthesis | 38.5 | 37.4 | 5.3 | 5.35 |
| S0004 | ZP_00024868 | 682 | ILTTAYSELDQTVEAVNQGQIH, LDYADLVGAESQR, GAGDATAVLAEVLGAAVR | Two Component Response Regulator Homolog | Regulatory Functions | 36.1 | 37.5 | 5.32 | 5.37 |
| S0005 | ZP_00021917 | 497 | GVPEGEMQFSFDMPLR | Two Component Response Regulator Transcription Regulator | Regulatory Functions | 25.8 | 25.8 | 4.88 | 4.88 |
| S0006 | ZP_00022189 | 524 | VAFLVEAHQAGIFDIR | Protein-Export Protein SecB | Protein Fate | 18.9 | 20.7 | 4.36 | 4.36 |
| S0007 | ZP_00022189 | 524 | VAFLVEAHQAGIFDIR | Protein-Export Protein SecB | Protein Fate | 18.9 | 20.8 | 4.36 | 4.45 |
| S0008 | ZP_00026716 | 58 | LSAAQVIDPTQLDTDGR, FEGDVAVVLAPGGER | Transcription Elongation Factor (Transcript Cleavage Factor) | Transcription | 17.1 | 20.2 | 4.43 | 4.51 |
| S0009 | ZP_00022841 | - | IATMIYEAPLGEYIGR, L/RLKGDEFLYR, GDLVGELGFIDQTPHSR | ORF2 protein, similar to ORF7 from Rhodococcus opacus and other cAMP- and cGMP-dependent protein kinases | Unknown Function | 18.6 | 17.9 | 5.65 | 5.27 |
| S0010 | ZP_00022841 | - | IATMIYEAPLGEYIGR, L/RLKGDEFLYR, GDLVGELGFIDQTPHSR | ORF2 protein, similar to ORF7 from Rhodococcus opacus and other cAMP- and cGMP-dependent protein kinases | Unknown Function | 18.6 | 17.5 | 5.65 | 5.26 |
| S0011 | ZP_00026366 | 710 | AETATIARPYAEALE, VAVDPSLIGGVSV, VGDEVLDTSVR | ATP Synthase Delta Chain | Energy Metabolism | 19.1 | 20.1 | 4.83 | 4.83 |
| S0012 | ZP_00022818 | 579 | 42.22% | Peptidyl-Prolyl Cis-Trans Isomerase (Trigger Factor) | Cellular Processes | 50.2 | 47.0 | 5.19 | 5.00 |
| S0013 | ZP_00022358 | 540 | AAITASMVAELRAK, EVSLLNQSFVK, GFTLYVVGEGIEK | Elongation Factor TS | Protein Synthesis | 30.9 | 33.9 | 5.36 | 5.31 |
| S0014 | ZP_00024747 | 679 | IVLDASFSPV, LVMNIETDGVISPEEAIR, APQIDPILLRPVDDLELTVR | DNA-Directed RNA Polymerase (Alpha Chain) | Transcription | 35.6 | 38.2 | 5.49 | 5.51 |
| S0015 | ZP_00024747 | 679 | IVLDASFSPV, LVMNIETDGVISPEEAIR, APQIDPILLRPVDDLELTVR | DNA-Directed RNA Polymerase (Alpha Chain) | Transcription | 35.6 | 37.7 | 5.49 | 5.44 |
| S0016 | ZP_00022060 | 511 | VQLVGDDLFVTNTK | Enolase (2-Phosphoglycerate Dehydratase) | Energy Metabolism | 45.9 | 44.5 | 4.76 | 4.76 |
| S0017 | ZP_00021732 | 475 | EALMEPIDIQDLLAR, VNALGIGAQGLGGLATVLDVK | Fumarate Hydratase | Energy Metabolism | 54.3 | 50.1 | 5.67 | 5.66 |
| S0018 | ZP_00022908 | 587 | QLQAVLDALSIEQLG, IVLAYEPVW | Triosephosphate Isomerase | Energy Metabolism | 25.4 | 23.8 | 5.61 | 5.0 |
| S0019 | ZP_00022037 | 509 | AFADNAPIPGEFAFCVP, GKPGPEALVGTATAGGLR, TGTYTL | Unknown Protein | Unknown Function | 23.3 | 23.0 | 5.33 | 5.48 |
| S0020 | ZP_00023642 | 634 | PIHASPLLERLTN, TPVALQAMEAVAK, DAGAQFAVSPGLTQK | Bifunctional ALKH (KHG/KDGP Aldolase) : 4-Hydroxy-2-Oxoglutarate Aldolase and 2-Dehydro-3-Deoxyphosphogluconate Aldolase | Energy Metabolism | 22.1 | 21.1 | 5.54 | 5.50 |
| S0021 | ZP_00022056 | 511 | ALYQLGEIAPTIHDD, ECLVGAGAVVTENK, ELTDADVANLYR | Unknown Protein | Unknown Function | 18.2 | 20.9 | 5.54 | 5.52 |
| S0022 | ZP_00026790 | 124 | MEHKLPSLPYAHD, GTEFESASLEEIVK, VAVGTFGSGWAWLVK | Iron Superoxide Dismutase | Cellular Processes | | 20.9 | | 5.42 |
| S0023 | ZP_00025066 | 688 | GFFFESFNQK | dTDP-4-Dehydrorhamnose 3,5-Epimerase | Cell Envelope | 20.7 | 20.7 | 5.4 | 5.40 |
| S0024 | ZP_00022818 | 579 | AEDFVYVLGEGR | Peptidyl-Prolyl Cis-Trans Isomerase (Trigger Factor) | Cellular Processes | 50.2 | 51.4 | 5.19 | 5.00 |
| S0025 | ZP_00022910 | 587 | AQIAASDLDLVVAGTER | Polyribonucleotide Nucleotidyltransferase | Transcription | 77.6 | 77.9 | 5.67 | 5.53 |
| S0026 | - | 463 | PVPFVFVSVVASGNAGFALAEK | Unknown Protein | Unknown Function | 21.5 | 12.9 | 9.27 | 5.30 |
| S0027 | ZP_00022097 | 515 | EAMADVGVTGLTVTEVK, GAEYVVDFLPK | Nitrogen Regulatory P-II Transcription Regulator | Regulatory Functions | 12.3 | 12.9 | 5.7 | 5.4 |
| S0028 | ZP_00021631 | 461 | YDINNEGLLFHAIVDGEQVTCVVTR, AALVLIEQGVPQPIVVK | Unknown Protein | Unknown Function | 9.9 | 9.9 | 4.65 | 4.65 |
| S0030 | ZP_00022691 | 569 | IVVAVDGSSTSDLALQEAIR | Unknown Protein | Unknown Function | 16.1 | 13.7 | 4.95 | 5.00 |
| S0031 | ZP_00022817 | 579 | GLGLVPMVVEQSGR, LNSILSEVTGQPVEK | ATP-Dependent Protease (Proteolytic Subunit) Transmembrane | Protein Fate | 23.6 | 20.9 | 5.15 | 5.24 |
| S0032 | ZP_00021742 | 476 | TATPSPTPHRGETVE, MPGMSGPELQER, AIDFIEKPFDESELR, LNVNTVADLLR | Response Regulator Transcription Regulator | Regulatory Functions | 23.3 | 22.6 | 5.31 | 5.39 |
| S0033 | ZP_00022037 | 509 | AFADNAPIPGEFAFCVP, GKPGPEALVGTATAGGLR, TGTYTL | Unknown Protein | Unknown Function | 23.3 | 22.8 | 5.33 | 5.38 |
| S0034 | ZP_00025158 | 690 | LVMMTNIPGVMDK, ISSALDAAKSGVHSVHIIIDGRIEHSLLLEILTEQAFGTMIRSH | Acetylglutamate Kinase | Amino Acid Biosynthesis | 33.1 | 33.8 | 5.34 | 5.27 |
| S0035 | ZP_00022800 | 578 | NQEVAVIGGGNTAVEEALYLSHIASK | Thioredoxin Reductase Oxidoreductase | Energy Metabolism | | 34.3 | | 5.28 |
| S0036 | ZP_00022040 | 509 | KPQAKTPAPI | Unknown Protein | Unknown Function | 18.1 | 18.1 | 9.43 | 9.47 |
| S0037 | ZP_00026162 | 708 | LSAFDVILGEPIPAK | Phosphoribosylaminoimidazole-Succinocarboxamide Synthase | Purines, Pyrimidines, Nucleosides and Nucleotides | | 35.0 | | 5.35 |
| S0038 | ZP_00022256 | 531 | SNVTLGGNAIEVGGKFPHSGDKA, VLNIVPSLDTPVCQASA, AVVVVDENDTVK | Thiol Peroxidase-Scavengase P20 Oxidoreductase | Cellular Processes | 17.3 | 18.4 | 5.83 | 5.7 |
| S0039 | ZP_00022213 | 527 | TIAEAVDLPVVLYNVPG, TVADMNNDTILR | Dihydrodipicolinate Synthase | Amino Acid Biosynthesis | 31.1 | 31.3 | 5.49 | 5.44 |
| S0040 | ZP_00021208 | 371 | TDALAVEGLDSAIER | Methylisocitric Acid Lyase | Energy Metabolism | 32.8 | 33.5 | 5.52 | 5.45 |
| S0041 | ZP_00022358 | 540 | 57.34% | Elongation Factor TS | Protein Synthesis | 30.9 | 33.8 | 5.36 | 5.43 |
| S0042 | ZP_00023023 | 594 | 50.32% | Aspartate-Semialdehyde Dehydrogenase (ASA Dehydrogenase) (ASADH) | Amino Acid Biosynthesis | 40.7 | 39.4 | 5.67 | 5.60 |
| S0043 | ZP_00022294 | 534 | IYAELIGFGMSGDAFHMTAPNTDGPR, NNFGFGGTNGTLVFR | 3-Oxoacyl-[Acyl-Carrier-Protein] Synthase II | Fatty Acid and Phospholipid Metabolism | 42.9 | 43.3 | 5.78 | 5.81 |
| S0044 | ZP_00021932 | 499 | AEDSLGDELVGAVITVPAYFDDAQR, LTLDADTFIQITANLVQK | Chaperone | Protein Fate | 66.2 | 43.7 | 4.96 | 6.04 |
| S0045 | ZP_00026423 | 711 | VPNTDNLGLDAVGLAADAR | Dihydrolipoamide Dehydrogenase (Component of Pyruvate and 2-Oxoglutarate Dehydrogenases Complexes) Oxidoreductase | Energy Metabolism | 50.5 | 47.3 | 6.09 | 6.23 |
| S0046 | ZP_00024195 | 660 | MTVIAGDAFGAESPVR, GIYPVDGNLEVDGETLPAEHLVVLVPGEPVTIK | Pirin-Like Protein | Unknown Function | 31.9 | 34.2 | 5.7 | 5.59 |
| S0047 | ZP_00024369 | 666 | FEGSAQYVATDDLK, ELVAAAMESFYEMR | Unknown Protein | Unknown Function | 33.3 | 33.2 | 5.6 | 5.60 |
| S0048 | ZP_00023736 | 639 | GKIIGIDLGTTNSCVS/49.69% | Chaperone protein DnaK (Heat Shock Protein 70) | Protein Fate | 69.2 | 71.4 | 4.86 | 4.90 |
| S0049 | ZP_00026366 | 710 | 42.54% | ATP Synthase Delta Chain | Energy Metabolism | 19.1 | 20.2 | 4.83 | 4.83 |
| S0050 | ZP_00022824 | 580 | TTLTAAIATVLAAK, AVDGTFLMPVEDVFSISGR, LIAPIAMEEGLR | Elongation Factor TU | Protein Synthesis | 43.1 | 43.2 | 5.41 | 5.38 |
| S0051 | ZP_00025772 | 701 | MNLRPLHDRVIVKR | 10kDa Chaperonin (Protein CPN10; GroES) | Protein Fate | 6.6 | 12.9 | 8.12 | 5.57 |
| S0052 | ZP_00024238 | 662 | 60.00% | 50S Ribosomal Subunit Protein L7/L12 | Protein Synthesis | 12.6 | 12.6 | 4.83 | 4.83 |
| S0053 | ZP_00023150 | 603 | MSEQIKYVSDASF | Thioredoxin 1 (Redox Factor) | Energy Metabolism | 11.9 | 12.7 | 4.98 | 5.01 |
| S0054 | ZP_00025145 | 690 | SDKDYMNEAQLAFFKN | DnaK Suppressor | Cellular Processes | 14.5 | 14.5 | 5.21 | 5.21 |
| S0055 | ZP_00021462 | 433 | SKVQLQTTQGAIVIELDAEK | Peptidyl-Prolyl Cis-Trans Isomerase B | Cellular Processes | 18.2 | 18.3 | 5.36 | 5.35 |
| S0056 | ZP_00025770 | 701 | AAKDVVFGDAARAKM | Chaperonin (Protein CPN60, GroEL, AMS) | Protein Fate | 57.2 | 57.2 | 5.11 | 5.11 |

**Table 4.1: List of *R. metallidurans* CH34 proteins identified by N-terminal sequencing (E), peptide mass fingerprinting (M) or tandem MS sequencing (T). Protein database accession numbers and identities of the closest homologue were obtained via BLASTP on NCBI database. Contig position of the open reading frames was determined by BlastP on PEDANT database. Sequence informations generated by Edman degradation (E, underlined) or tandem MS (T) are indicated. When MALDI was used for protein identification, sequence coverage is indicated as calculated by the ProteinProbe search engine. TIGR role categories were attributed by identity to the closest homologue. Experimental (E) pI and Mr values were calculated from analysis of the gel images performed with the PDQuest Software. The respective theoretical isoelectric point and molecular weight values were calculated using the Predict pI/MW tool from EXPASY WWW server (see previous page).**

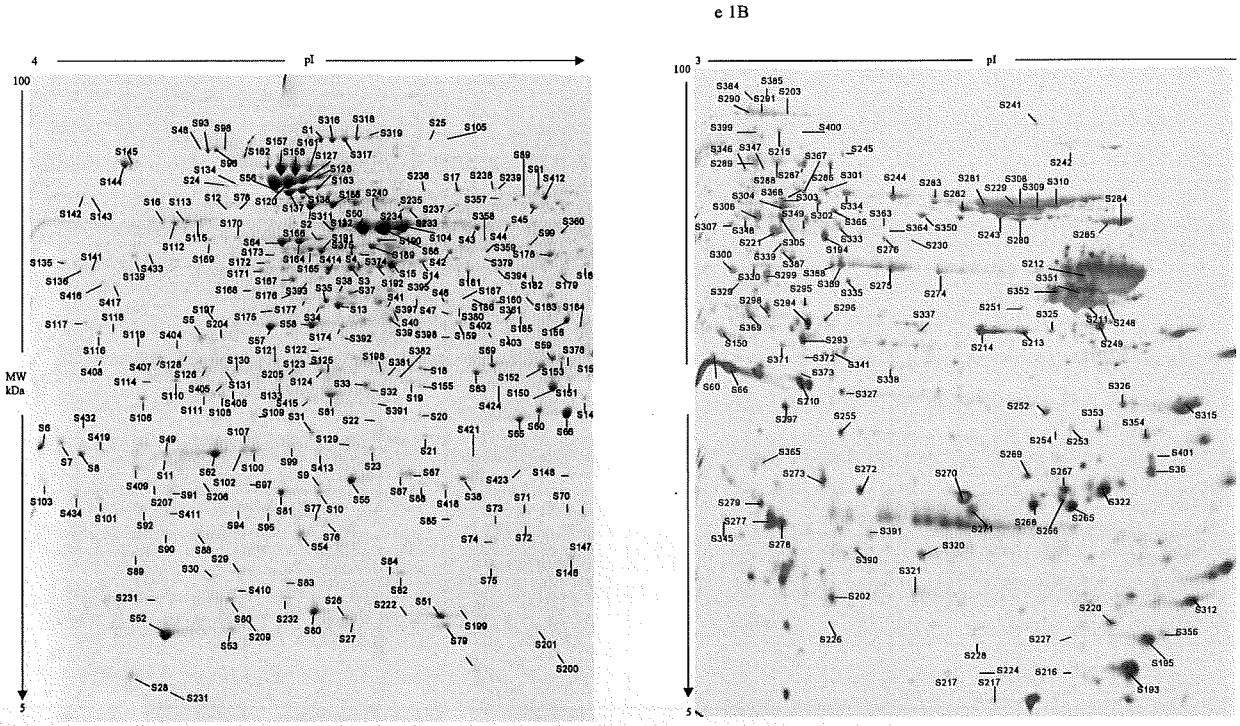



**Figure 4.1: Analytical 2D gel map of *R. metallidurans* total cellular proteins focused using a pH4-7 (A) and a pH6-11 (B) IPG strip and stained with colloidal Coomassie. Identified proteins are labeled by a spot number.**

While 142 different ORFs were assigned to the same number of protein spots, the other 82 ORFs were more than once, to the remaining 210 protein spots that correspond to different isoforms or fragments. Three spots (351, 352 and 394) are not pure and were found to contain two different protein species each. This is possibly due to contamination of the spot of interest by other surrounding spots of high intensity. Finally, 13 spots, corresponding to 11 different protein species, are located in the borders of a contig. As a

consequence, their complete corresponding gene sequence, their predicted isoelectric point and molecular weight are not available.

Figure 4.2 shows a comparison between the predicted subcellular localization of the 224 proteins identified in this work (A) and the predicted localization of the 8131 proteins putatively encoded by the *R. metallidurans* genome as displayed on the PEDANT WWW site (B). In both cases, the ratios of cytoplasmic/membrane proteins are similar: about one fourth of the proteins are located in the (inner or outer) bacterial membrane. Thus, we conclude that our experimental working conditions (protein extraction, solubilization, etc.) are suitable to to identify proteins whatever their subcellular localization is. Figure 4.3 displays the distribution of every ORF identified in this work between the TIGR functional categories. More than half of the proteins identified perform essential functions (biosyntheses, cellular processes, energy metabolism), consistent with the fact that most enzymes belonging to this group are generally expressed at high and constitutive levels. It should also be noticed that hypothetical proteins and proteins with unclassified or unknown function represent a quarter of the identified proteins, despite an elevated spot intensity, which may point out an important, though yet unknown, function accomplished by these proteins.

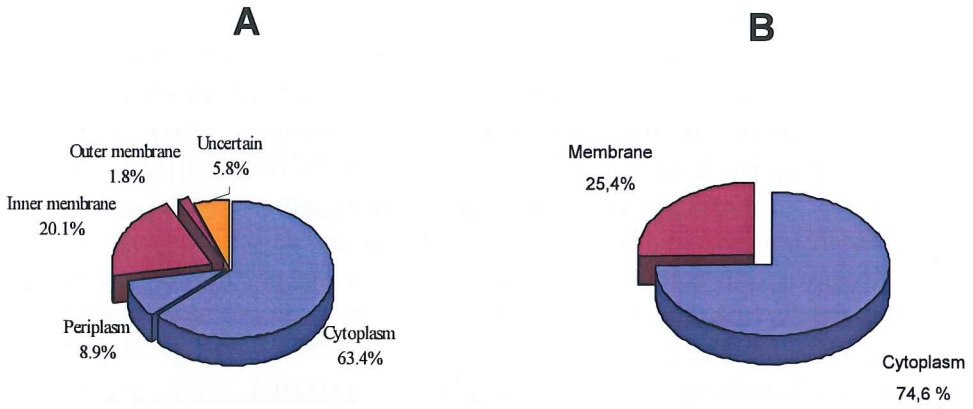



**Figure 4.2: Comparison of predicted subcellular localization of proteins identified in this work and of all proteins putatively encoded by the *R. metallidurans* genome.**

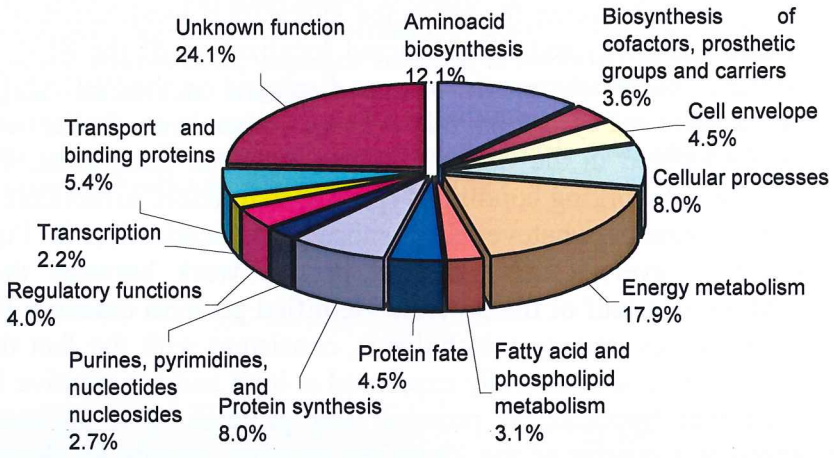



**Figure 4.3: Functional classification (TIGR role categories) of the proteins identified in this work.**

With complex samples such as total bacterial cell extracts, 2-DE electrophoresis on a single wide-range pH gradient reveals only a small percentage of the whole proteome because of insufficient spatial resolution and the difficulty to reveal low copy number proteins in presence of the most abundant proteins. One remedy is to use of multiple overlapping narrow IPGs in the first dimension (zoom-in gels) and/ or used prefractionation proteins by isolation of cell compartments (extracellular, cytoplasmic, periplasmic) and by sequential extraction procedures (membranes proteins) with increasingly powerful solubilizing buffer (usually aqueous buffers, organic solvents and detergent-based to extraction solutions). Zoom-in gels in the pH range between pH 3.5-4.5, 4.5-5.5, 5-6 were also used in this work. Extracellular proteins have been purified by filtration of the culture (in the log phase, D.O at 600 nm=0.5) and precipitation of the proteins by the TCA-acetone method. Periplasmic proteins were separated by a chloroform extraction where as the membrane proteins were extracted by a chloroform/methanol extraction.

By this approach, more 2000 protein spots were visible in the all different 2-DE gels and extracts. At this date, a dynamic database containing about 1400 identified proteins have been built (Figure 4.4, 4.5). In the future, this database will be used to monitoring proteome modifications during different growing stress conditions as spatial flight conditions (MESSAGE 1 and 2 experiments, MELGEN 2)

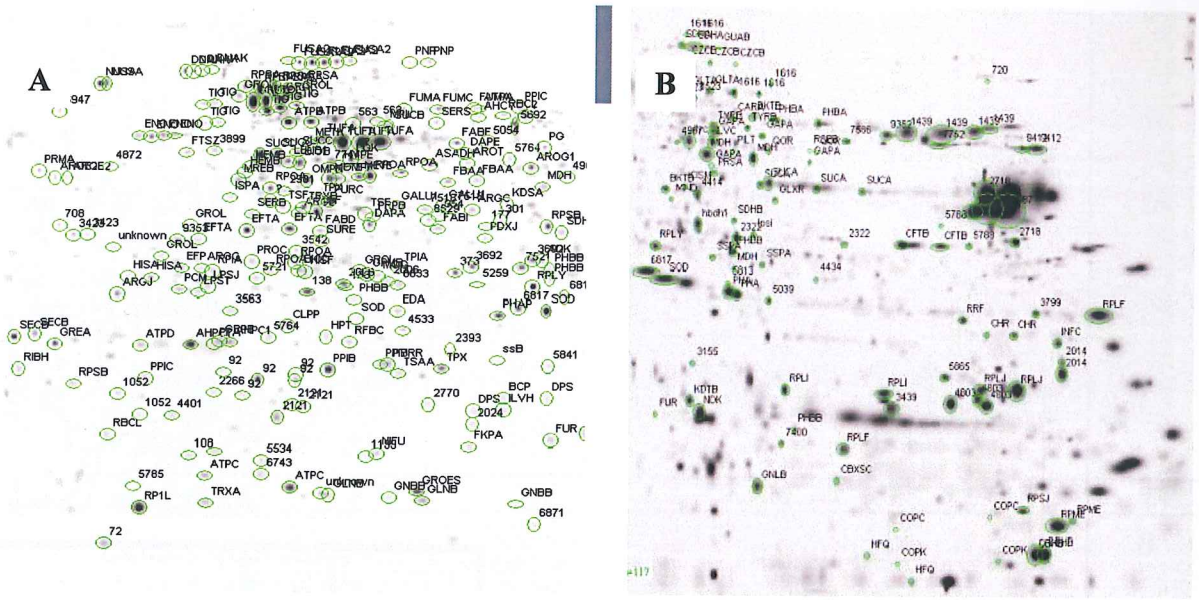
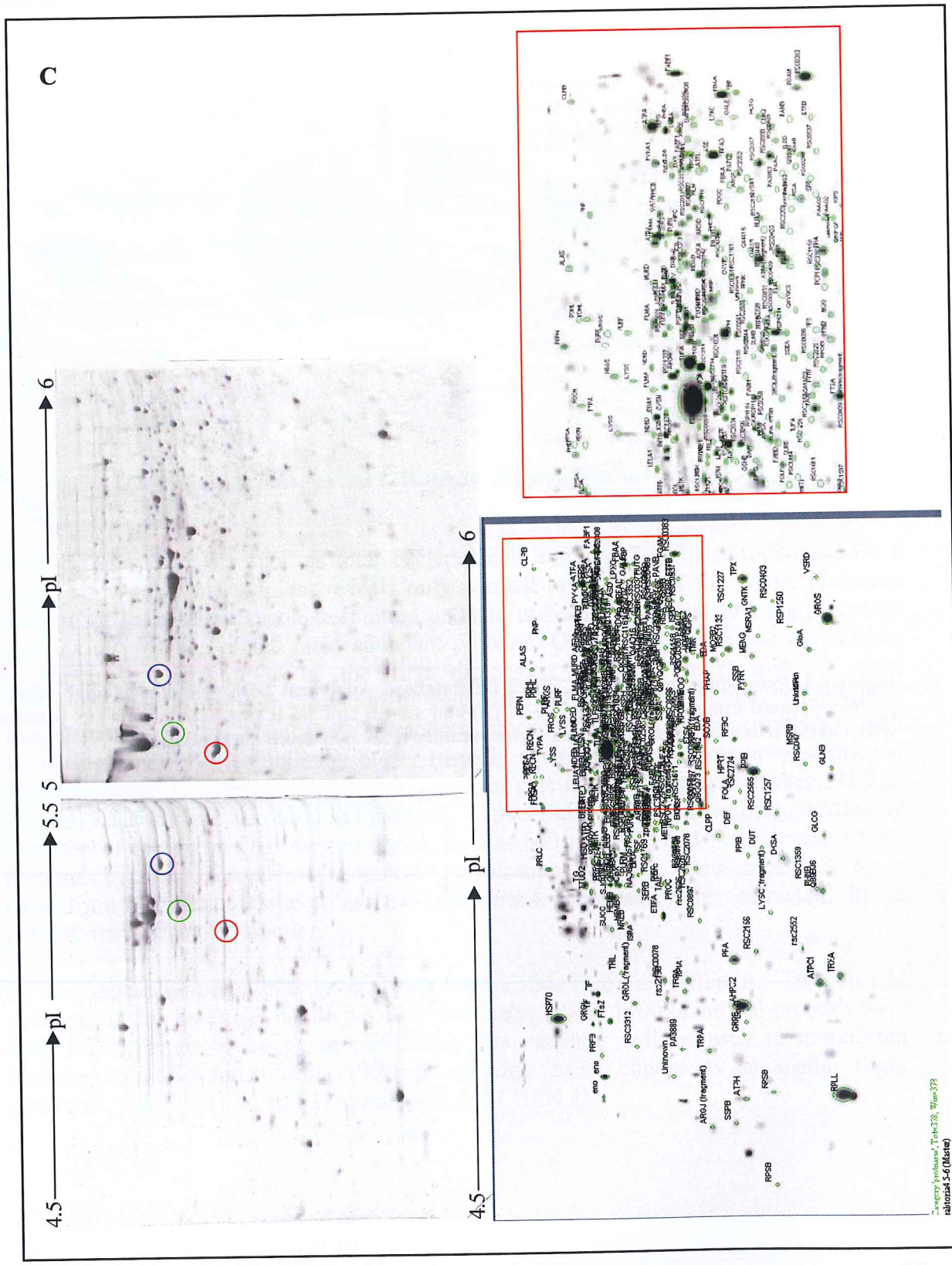



**Figure 4.4. *Ralstonia metallidurans* CH34 2-DE database. Analytical 2-DE gel maps of total cellular proteins focused using pH 4-7 (A), pH 6-11 (B) and 4.5-5.5 and 5-6 (C(see next page)). The proteins were characterised by a gene name and a gene number if the function is not established**


C
pI
6
5.5
5
4.5
pI
6
4.5

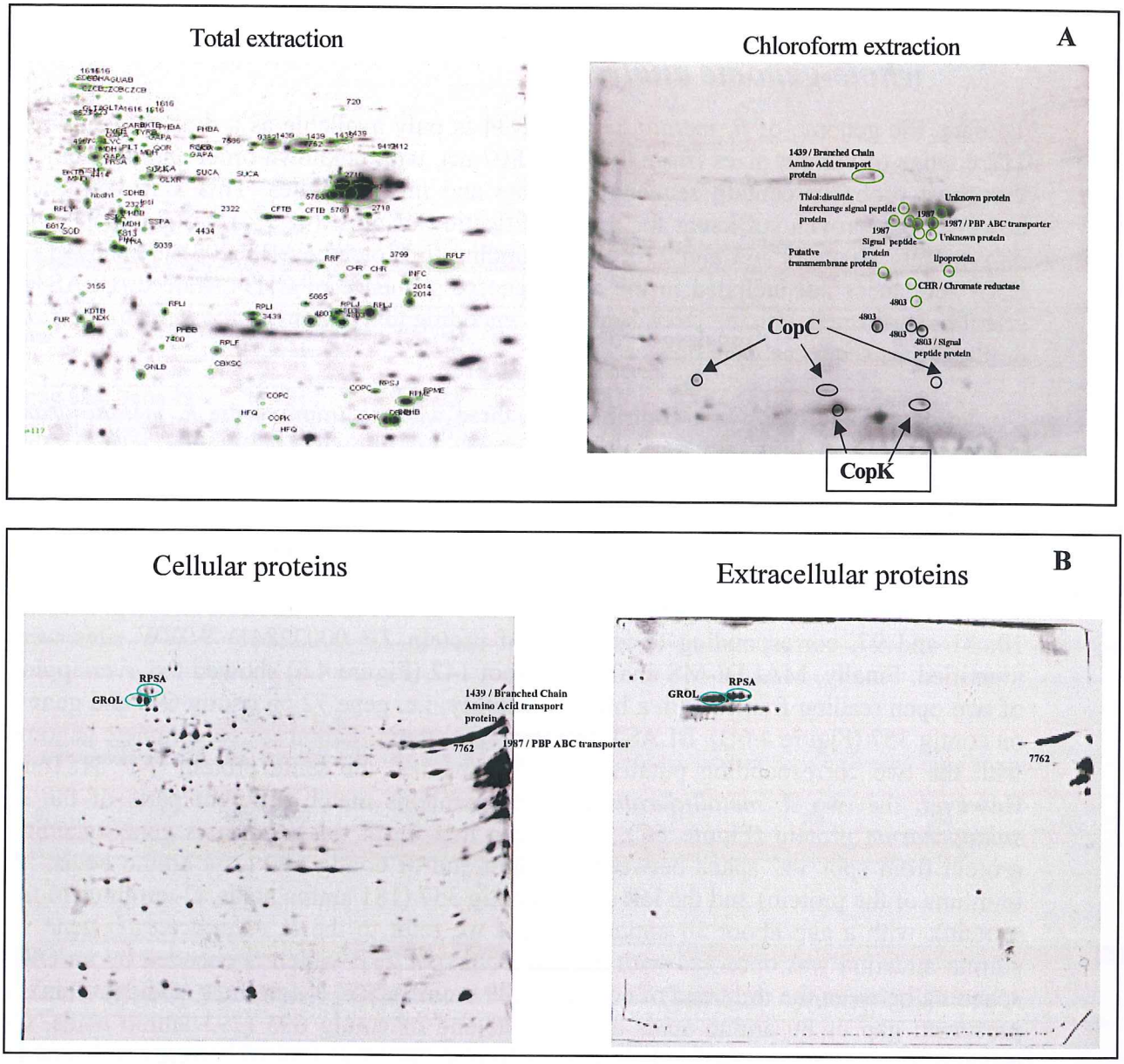
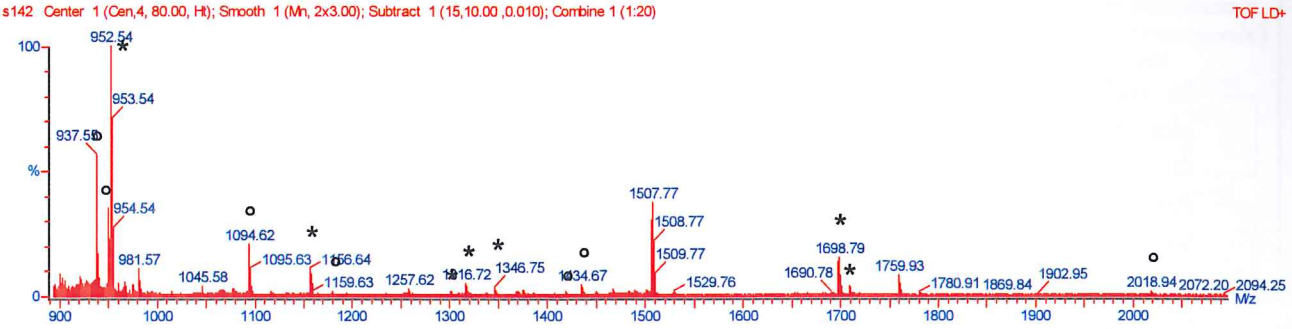



**Figure 4.5. Subcellular proteome of *Ralstonia metallidurans* CH34. Periplasmic proteins purified by a Chloroform extraction (as described in material and methods) were separated by 2-DE using a 6-11 pH gradient during the first dimension (A). Extracellular proteins were separated by filtration and submitted to 2-DE using a 3-10 Immobiline pH gradient.**

## 4.5 *Contribution of the R. metallidurans protein database to whole-genome analyses*

To date, the genome of *R. metallidurans* CH34 is only available as a draft collection of 712 contigs of various sizes (from 70 to 165807 nt), with unknown order and orientation, that most probably contain sequencing errors and misassemblies. This draft sequence, however, has proven sufficient for the identification of most (98.9%) protein spots. One only protein (spots N° 65 and 210, corresponding to protein AAB36203) is encoded by DNA sequences not included in the draft genome sequence of CH34: indeed, BLASTN searches performed with the DNA sequences encoding this protein on the *R. metallidurans* draft genome sequence were unsuccessful.

On the other hand, this work contributed in three ways to improve the *R. metallidurans* genome database. First, microsequence analyses by Edman degradation, even if this technique does not allow high throughput identification of proteins, led to the correction of the in silico predicted N-terminal sequences for 12 different proteins and the confirmation of N-terminal in silico prediction for 43 different protein spots identified by mass spectrometry. Second, during this work, proteins not included by ORF prediction softwares, either on NCBI (spots N° 26, 119, 220 and 360) or on PEDANT (spots N° 9, 10, 81 and 97, corresponding to isoforms of protein ZP_00022841) WWW sites were identified. Finally, MALDI-MS analysis of spot 142 (Figure 4.6) showed the overlapping of two open reading frames with a high probability: i.e. gene 72 on contig 690 and gene 1 on contig 357 (Figure 4.6B). BLAST search on *Ralstonia solanacearum* complete genome with the two corresponding putative proteins spanned the same protein (NP_518166). However, the two *R. metallidurans* predicted proteins match different parts of the *R. solanacearum* protein (Figure 6C), suggesting that the *R. metallidurans* gene encoding protein from spot 142 spans between the right end of contig 690 (168 amino acids, N-terminus of the protein) and the left end of contig 357 (181 amino acids, C-terminus of the protein), with a gap about 30 amino acids, if we refer to the *R. solanacearum* gene. A similar situation was observed with protein from spot 231, which is encoded by an ORF spanning between the right end of contig 94 (89 amino acids, N-terminus of the protein), a presumed gap of 80 amino acids and the left end of contig 693 (193 amino acids, C-terminus of the protein). Both contig connections, 690-357 and 94-693, were confirmed in the proper orientation by preliminary scaffold predictions performed in order to proceed to contig closure. In this context, we want to stress the fact that MALDI-MS, when associated to powerful analysis algorithms such as GlobalServer, is a very efficient tool to identify proteins, even when the target genome is only available as a draft sequence.




| Name | Score | Probability (%) | Match | Coverage (%) | MW | pI |
|---|---|---|---|---|---|---|
| **Search with all masses** | | | | | | |
| **Contig 690. Gene 72** | **10.76** | **56.33** | **7** | **42.61** | **18776** | **4.5** |
| Contig 357. Gene 1 | 2.55 | 43.67 | 7 | 24.06 | 20398 | 4.5 |
| Contig 630. Gene 24 | -3.34 | 0 | 7 | 29.30 | 28334 | 11.7 |
| **Search unmatched masses** | | | | | | |
| **Contig 357. Gene 1** | **2.73** | **98.87** | **7** | **24.06** | **20398** | **4.5** |
| Contig 630. Gene 24 | -3.40 | 0.22 | 7 | 29.30 | 28334 | 11.7 |
| Contig 711. Gene 238 | -3.63 | 0.17 | 9 | 23.08 | 32701 | 7.2 |

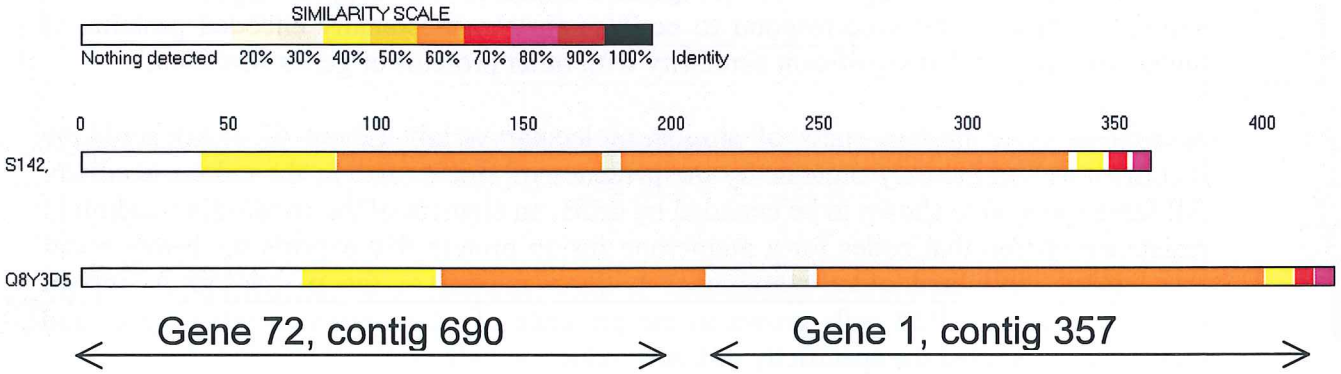



**Figure 4.6: Sequence analysis of protein spot S142. A. MALDI-MS mass spectrum of protein spot S142. Asterisks (*) indicate peptides corresponding to the protein encoded by gene 72 on contig number 690 and circles (o) locate peptides corresponding to the protein encoded by gene 1 on contig number 357. B. Peptide matching results for protein spot S142. Results were obtained from all masses and subsequently from unmatched masses. C. Tentative reconstitution of protein S142. Sequence alignment of protein Q8Y3D5 from *R. solanacearum* with *R. metallidurans* protein S142 reconstituted by juxtaposition of gene 72 (contig 690) and gene 1 (contig 357). Alignment was realized with the SIM program and viewed with the LALNVIEW software (http://us.expasy.org/tools/sim-prot.html).**

## 4.6 *Proteome analysis of the copper, zinc and nickel response in R. metallidurans*

In order to illustrate the use of our database in proteome studies of *Ralstonia metallidurans*, a preliminary comparative study was conducted with strain CH34 grown in the presence or in the absence of zinc, copper and nickel sulfate in the culture medium. Analysis of the protein expression pattern differences between the two culture conditions was performed, and differentially expressed proteins were identified using our *R. metallidurans* 2-DE database. The proteome maps of bacteria growing in gluconate were designated as our reference maps to detect differentially expressed proteins.

The results of the gel comparison (Figure 4.7) showed that more 10 proteins were newly expressed in presence of copper (0.8mM). In order to identify the differentially expressed proteins, spots were excised and analysed by MALDI-TOF. Among these proteins, we found more proteins from which the corresponding gene are localised on the plasmid pMOL30 and known to be involved in the copper resistance (CopR, CopC, CopB, CopA). Interestingly, a new protein (CopK), not observed by the genomic approach, also localised in the pMOL30 Copper cluster was expressed in presence of copper (Mergeay et al., 2003). This protein is a *Ralstonia metallidurans* specific protein (any similarity was found with a gene from an other genus, strains....). It is a periplasmic protein as shown by the analysis of the bacterial chloroform extract.

Two others proteins are up- and down regulated in presence of copper, respectively CopX and Cuf. These proteins correspond to periplasmic chromosomally encoded proteins of unknown function. No significant similarity with other proteins or genes was found.

A train of three protein spots of similar molecular weight (about 57 kDa) could be identified as significantly induced by the presence of zinc sulfate in the culture medium. All three spots were shown to be encoded by czcB, an element of the cobalt-zinc-cadmium resistance operon that codes for a membrane fusion protein that exports the heavy metal ions across inner and outer membranes without release in the periplasm. A control experiment using CH34 cells grown in the presence of other heavy metals (copper and nickel) demonstrated the specificity of CzcB induction by zinc.

Down-regulation of Cuf by copper and Zinc suggests a common regulation of the corresponding gene by divalent cations. Yet, it is unclear whether Cuf or CopX are involved in any form of heavy metal resistance mechanism in the periplasmic space or if this regulation is solely a consequence of the presence of heavy metal ions in the bacteria's environment. In parallel, we observed also an up-regulation of the different proteins involved on the copper resistance in presence of Zinc and not in presence of nickel. Theses results suggest a possible synergy between different molecular mechanisms of heavy metal resistance.

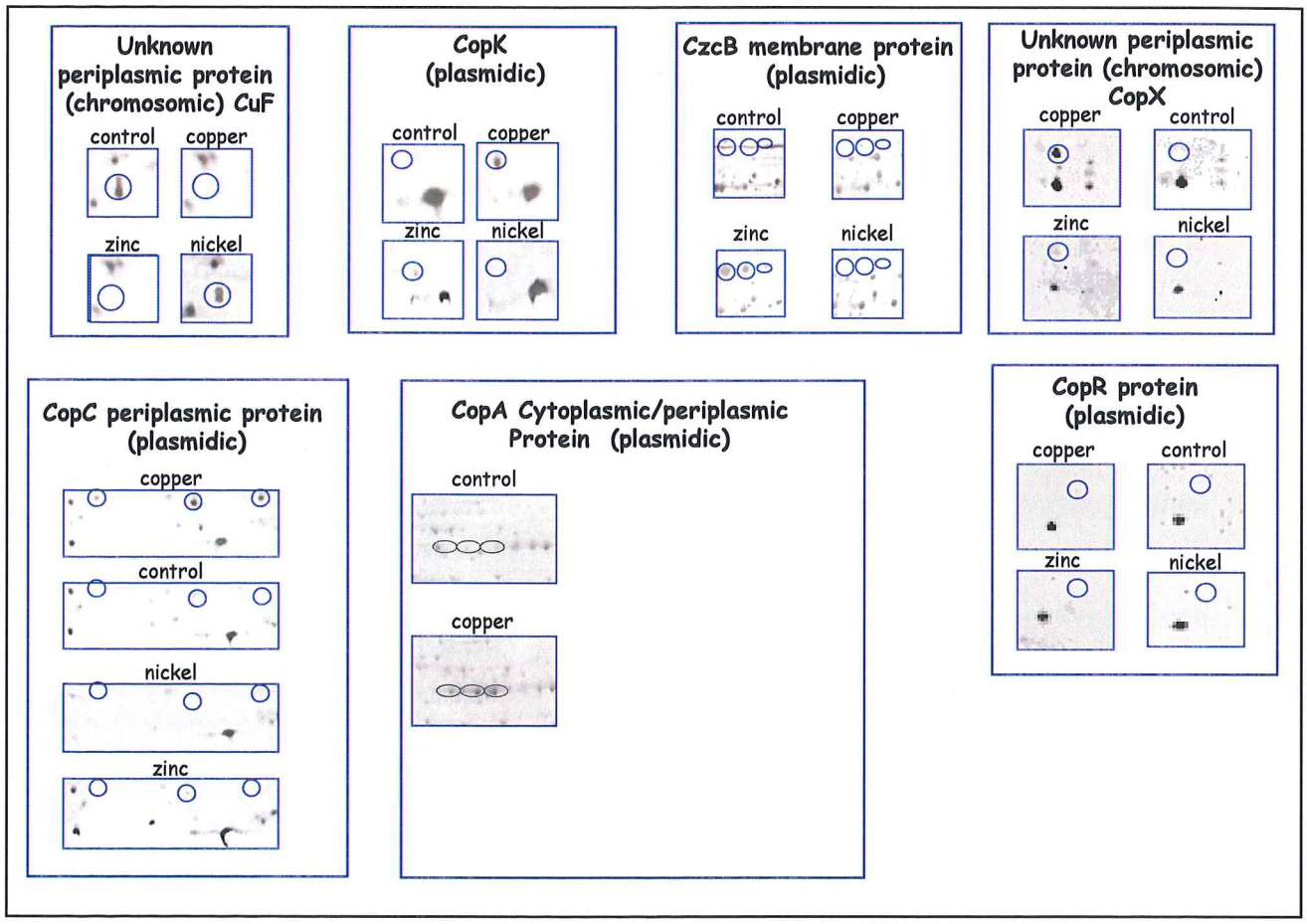



**Figure 4.7. Specific zinc, copper and nickel induction of protein expression in *R. metallidurans*. Bacterial cultures and protein extracts were prepared according to 'material and Methods'.**

## *4.7 Proteome analysis of the R. metallidurans growing in acetone medium*

An other example to illustrate the use of our database in of *Ralstonia metallidurans* corresponds to the proteome analysis of the strain CH34 grown in the presence or in the absence of acetone in the culture medium.

Diverse microbes (*Xanthobacter autotrophicus* and *Rhodobacter capsulatus*) are capable of grow using acetone as the primary source of carbon and energy. Two distinct transformations have been proposed as the first step of acetone metabolism in microorganisms: an oxygen dependent hydroxylation reaction producing acetol and a Co2 dependent carboxylation to produce acetoacetate. The aims of this proteomic approach are firstly to confirm that *R. metallidurans* is able to use the acetone as carbon source and secondly to characterize acetone degrading enzymes at the molecular level.

Preliminary results show more proteins up-regulated in presence of acetone. Among these proteins we found the acetone carboxylase alpha, beta and gamma subunit proteins (Figure 4.8). This enzyme is inducible and is expressed at high levels in cell grown with acetone or acetone +gluconate as the source of carbon. The enzyme is a multimeric protein that catalyses the mg ATP dependent carboxylation of acetone.

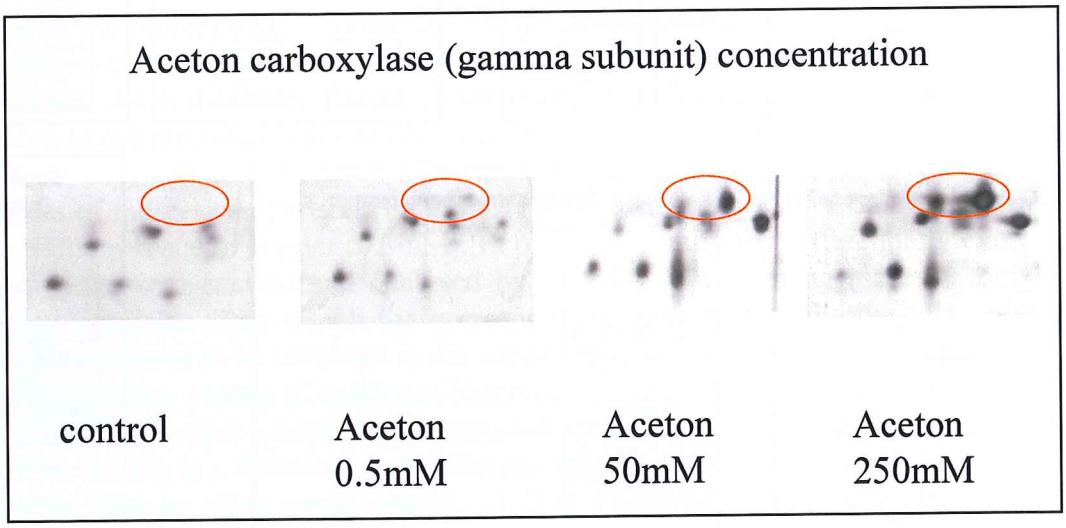



**Figure 4.8. Acetone carboxylase expression in presence of different acetone concentration as the sole source of carbon.**

## 4.8 *Conclusions*

In this study, we constructed a database of total cellular proteins from *Ralstonia metallidurans* CH34. This database will serve in the future as a reference to which proteins differentially expressed, consecutively to heavy metal contamination or to exposure to a variety of stress and environmental conditions (i.e. space conditions), will be compared. Protein identification was carried out via N-terminal amino acid sequencing, MALDI-TOF MS and tandem MS. So far 1400 different proteins were characterized out of 1600 protein spots. 99.9 % of proteins tested with the MALDI-TOF MS have been identified with success. Although the proteome map is still not complete, one could appraise the importance of proteomics for genome analyses (especially by the MALDI-TOF approach) through (1) the identification of previously undetected open reading frames, (2) the identification of proteins not encoded by the already sequenced genome fragments, (3) the characterization of protein-encoding genes spanning two different contigs, enabling their merging, and (4) the precise delineation of the amino-terminus of several proteins. Finally, this map will prove a useful tool, firstly, in the identification of proteins differentially expressed in the presence of different heavy metals and secondly, in the identification of the mechanism to use acetone as carbon source in *R. metallidurans*. These last results are the first demonstration of the ability of *R. metallidurans* to use the acetone as carbon source.

These results are published in an international Proteomic journal " Proteomics" (Global analysis of the *Ralstonia metallidurans* proteome: prelude for the large-scale study of heavy metal response. Isabelle Noël-Georis1, Tatiana Vallaeys2,3, Renaud Chauvaux1, Sébastien Monchy2, Paul Falmagne1, Max Mergeay2 and Ruddy Wattiez1. Proteomics. 2004). An other article will be submitted as soon as possible entitled "Proteomic analysis of copper response in *Ralstonia metallidurans* CH34 unravels new resistance mechanisms. Ruddy Wattiez , Isabelle Noël-Georis, Vanessa Auquier, Tatiana Vallaeys, Guy Vandenbussche, Jean-Marie Ruysschaert, Paul Falmagne, Max Mergeay." In preparation, 2004.

This approach and technology, particularly appropriate to begin a proteomic study from uncompleted bacterial genome will be used to detect and identify protein modifications in different stress conditions (temperature variation, oxidative stress, space conditions...) especially with the MELiSSA strains as *Arthrospira* sp. PCC8005 and *Rhodospirillum rubrum* ATCC25903.

Mergeay M, Monchy S, Vallaeys T, Auquier V, Benotmane A, Bertin P, Taghavi S, Dunn J, van der Lelie D, Wattiez R. *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes.
FEMS Microbiol Rev. 2003 Jun;27(2-3):385-410. Review.

# 5 HIGH RESOLUTION GENOTYPING OF MELISSA STRAINS

## 5.1 Introduction

The AFLP technology was originally developed for the detection of genetic markers in plants and animals but was introduced as a general genotyping method to the field of bacterial taxonomy because of its high resolution and flexibility[1]. The AFLP concept has been explained in great detail in TN70.2 (section 5.3.1).

## 5.2 Material and methods

Genomic DNAs of *R. metallidurans*, *R. rubrum*, and *A. platensis* were prepared as described in TN70.10. The design of adaptors and AFLP primers has been described in TN70.3 (section 5.4.2). All procedures for preparation of template DNA and setting up AFLP reactions are also described in TN70.3 (sections 5.4.3 & 5.4.4). The number of cycles after the touch-down was emperically set at 23 (e.g. a total of 35 cycles). Additional adaptors and primers that were used are listed in Table 1.

Here follows the description of AFLP protocol used on *R. metallidurans* and *R. rubrum.*

### 5.2.1 PREPARATION OF DNA

(Essentially the Pitcher method)

1) take cells directly from the petri plate (1 full ose)
2) resuspend into 100 µL TE (10 mM Tris, 1 mM EDTA pH 8.0)
3) centrifuge 2 min at 13K, take off supernatans
4) add 500 µL GES (see Pitcher protocol)
   ca. 5-10 min ice till clearance
5) add 250 µL 7M $NH_4Ac$
   ca. 10 min ice
6) add 500 µL ice-cold chloroform:iso-amylalcohol (24:1), mix well (not vortex)
7) centrifuge 20 min at 13K
8) take 700 µL supernatans into new eppendorf tube, add 380 µL iso-propanol
9) centrifuge 5 min at 13K (best at 4°C if possible)
10) take off supernatans, quick respin, take off second time
11) wash DNA pellet 2x with 100 µL 70% EtOH
12) air dry DNA pellet (ca. 10 min on RT)

[1] Janssen *et al.* (1996). Microbiology 142:1881-1893.

13) dissolve DNA into 100 µL MQ-$H_2O$
14) add 10 µL RNAse (500 µg/mL), incubate at 37°C for 90 min
15) re-precipitate DNA, dissolve into 100 µL $T_{0.1}E$ (10 mM Tris, 0.1 mM EDTA pH 8.0), store at –20°C

## 5.2.2 PREPARATION OF TEMPLATE

1) restriction of DNA: general conditions: 1x OPA buffer, 5 mM DTT, 100 µg/mL BSA, 10 units RE. Perform digests separated, highest incubation temperature first. Make sure that the above reaction conditions are always maintained. Digest always with excess enzymes and prolonged time (90 min for each digest).
2) Ligation of adapter: general conditions: 1x OPA buffer, 5 mM DTT, 100 µg/mL BSA, 0.2 mM ATP, 0.4 µM T(etra)-adaptor, 0.04 µM H(exa)-adaptor, 1 unit T4-ligase
3) Try to go for the following conditions: first restriction in 20 µL, second restriction in 25-30 µL (or simultaneous restriction at 25-30 µL), ligation in 40 µL.(consecutively, e.g. in the same tube)

   Example (for 10 DNAs) (*):

   MIX-I: 30 µL 10x OPA
   30 µL 50 mM DTT
   30 µL 1 mg/ml BSA
   90 µL MQ-$H_2O$
   10 µL HindIII (10 U/µL)
   10 µL MseI (10 U/µL)

   MIX-II: 10 µL 10x OPA
   10 µL 50 mM DTT
   10 µL 1 mg/mL BSA
   20 µL 4 mM ATP
   8 µL HindIII-adaptor (2 µM)
   8 µL MseI-adaptor (20 µM)
   34 µL MQ-$H_2O$

   10 µL DNA + 20 µL MIX-I (incubate 90-120 min 37°C)
   → 30 µL digest + 10 µL MIX-II (incubate 90 min 37°C)

4) Precipitation:
   40 µL template
   95 µL MQ-$H_2O$
   66 µL $NH_4Ac$ (7.5 M)
   200 µL iso-propanol (ice-cold)

   5 min at RT (or ice)
   centrifuge 15 min at 13K
   wash 1x with 100 µL 70% EtOH
   centrifuge 10 min at 13K
   dissolve pellet into 100 µL $T_{0.1}E$
   store at –20°C

## 5.2.3 AFLP REACTIONS

usually, 2 mixes are prepared

for 10 reactions, MIX-A (*):

58.5 µL MQ-$H_20$
20 µL 10x PCR buffer
1.5 µL Taq DNA PolymeraseI (5 U/µL)

for 10 reactions, MIX-B (*):

71 µL MQ-$H_20$
6 µL T-primer (50 ng/mL)
3 µL H-primer (20 ng/mL)
20 µL dNTP (2 mM)

set-up: 2 µL template + 8 µL MIX-A
+ 10 µL MIX-B

AFLP cycler program:

1 cycle 60s@94° / 30s@65° / 60s@72°
11 cycles 30s@94° / 30s@64.3° → 56° / 60s@72°
23 cycles 60s@94° / 30s@56° / 60s@72°
1 cycle 3 min @ 72°C

After the run, add 3 µL loading dye, incubate 3 min at 94°C, store if needed at –20°C; (*) when making mixes, always count some loss for pipetting, so for 12 DNAs make mixes for 13

| Enzyme | Sequence | Function | Notation | 5'-3'-sequence | Length |
|---|---|---|---|---|---|
| XhoI / SalI | C/TCGAG | X-adaptor | XhoAd01 | CTCgTAgACTgCgTACA | 18 |
| | G/TCGAC | | XhoAd02 | TCgATgTACgCAgTC | 15 |
| | | | | | |
| MseI | T/TAA | M-adaptor | MseAd01 | gACgATgAgTCCTgAg | 16 |
| | | | MseAd02 | TACTCAggACTCATC | 15 |
| | | | | | |
| | | M-primers | M01 | gATgAgTCCTgAgTAAA | 17 |
| | | | M02 | gATgAgTCCTgAgTAAC | 17 |
| | | | M03 | gATgAgTCCTgAgTAAg | 17 |
| | | | M04 | gATgAgTCCTgAgTAAT | 17 |
| | (IRDye800) | S-primer | S02 | gACTgCgTACATCgACC | 17 |
| | | | | | |
| | (IRDye800) | X-primer | X02 | gACTgCgTACATCgAgC | 17 |

**Table 5.1: additional AFLP primers and adapters for MELiSSA strains**

Visualisation of AFLP patterns was done on a Li-Cor Global $IR^2$ System essentially as described in TN70.3 (5.4.5). The length of the run was usually 2 hrs. Image files were collected and locally stored as TIF files.

The cycle program used for AFLP entails one cycle of 60 sec at 94C for full denaturation of template DNA, followed by 11 cycli with decreasing annealing temperature with 30 sec denaturation. These first 12 steps is also called the touch-down part of the PCR, and starts at a high annealing temperature (65C) and ends at the optimal annealing temperature (56C). This stringency ensure high specificity of primer to template hybridisation e.g. very few mismatched primers will eventually contribute to the PCR. Incubation of PCR product in the presence of formamide (in the loading dye) ensure complete denaturation of DNA. Keep in mind that separation of amplicons is done in denaturing conditions (i.e. 7 M urea gels) and that only one strand of DNA is labeled through the labeled primer. Electrophoresis of single standed DNA results in much better, sharper banding patterns.

## *5.3 Results and discussion*

We first set out to test restriction of the prepared gDNAs with endonucleases *Apa*I, *Hin*dIII, and *Taq*I, which we thought would be ideal for genomes with 50-60 mol% (G+C) (see also TN70.3, section 5.4.1). Note that the quality of the *A. platensis* DNA was not so good, and the yield was much lower than for the other gDNAs. We even noticed considerable degradation of this gDNA when stored at –20 C. As expected, *Taq*I digested the three gDNAs in very small fragments (less than 1 kb; Fig. 5.1) while *Apa*I and *Hin*dIII fragments were in the range of 2-9 kb and 8-20 kb, respectively.

For all three gDNAs, two templates (Apa-Taq and Hin-Taq) were used, each subjected to four different primer combinations. The layout of this experiment is given in Fig. 5.2. In total, eight different reaction mixes were prepared, and 24 reactions were loaded on a 5.5% polyacrylamide gel. The AFLP reactions for *A. platensis* resulted in very few or no bands, possibly because of the poor quality of the gDNA (it may have been degraded during the template preparation). Banding patterns contained between 10 and 60 bands but were generally rather small in size with most bands situated below 400 bp. Apparantly, restriction by *Taq*I was more frequent then anticipated (figure 5.4). No other method was hereafter tested for *A. platensis.*

Consequently, we decided to test out additional restriction enzymes, namely *Pst*I, *Sal*I, *Xho*I, and *Mse*I (results not shown). Remarkably, *Pst*I did not digest *A. platensis* nor *R. rubrum* gDNA, whereas also *Sal*I and *Xho*I could not digest *A. platensis* gDNA. These digest were repeated once but gave the same results. *Mse*I cuts all three gDNAs to small fragments, as would be expected of a frequent cutter. It is possible that *A. platensis* and *R. rubrum* DNA is heavily methylated preventing particular restriction endonucleases to bind to the corresponding recognition sites.

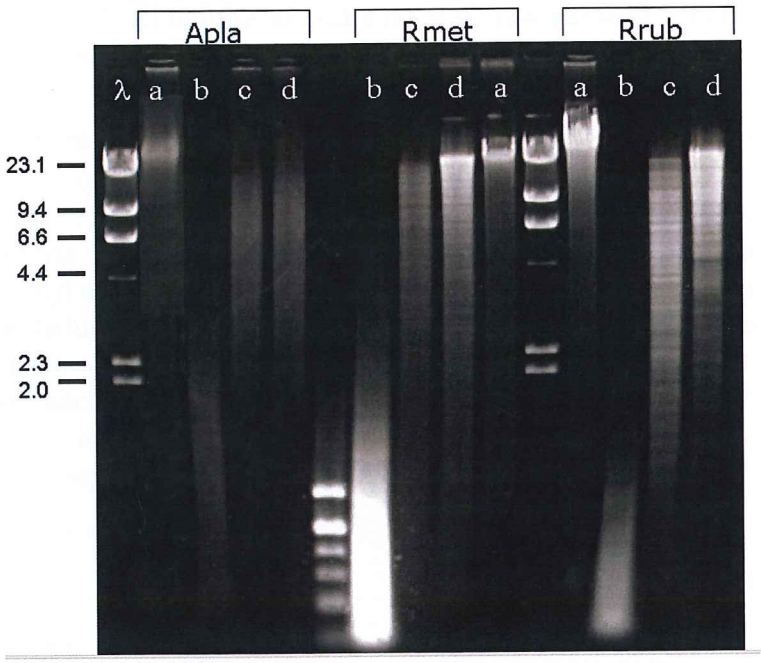



**Fig. 5.1: restriction digests of gDNAs from *A. platensis* (Apla), *R. metallidurans* (Rmet), and *R. rubrum* (Rrub); l: lambda DNA cut with HindIII; (a) undigested, (b) TaqI, (c) ApaI, and (d) HindIII.**

AFLP analysis was applied to all three gDNAs using new templates Apa-Mse, Hin-Mse, Sal-Mse, and Xho-Mse, and a total of 16 primer mixes was made (using Mse-primers M01 through M04). The layout of this experiment is displayed in the aflp-grid of Fig. 5.3) and the resulting image is given in Fig. 5.5. Clearly, the banding patterns obtained with MseI are much better in that the average size of the fragments is much larger, and bands are more evenly distributed over the length of the gel (figure 5.5).

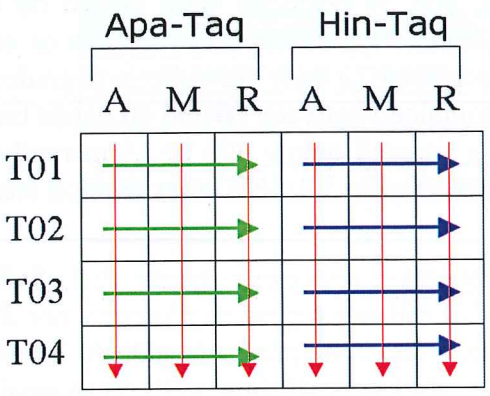



**Fig 5.2: aflp-grid to set up AFLP reactions; A, *A.platensis*; M, *R. metallidurans*; R, *R. rubrum*. T01-T04: Taq-primers. Arrows: red, template DNA; green & blue, primer mixes**

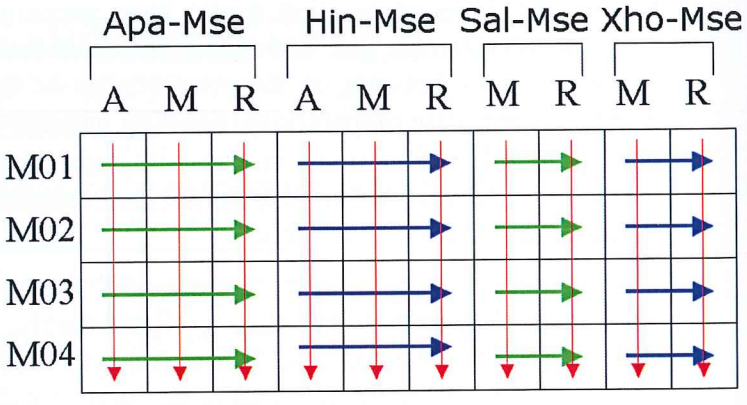



**Fig 5.3: aflp-grid to set up AFLP reactions; A, *A.platensis*; M, *R. metallidurans*; R, *R. rubrum*. T01-T04: Taq-primers. Arrows: red, template DNA; green & blue, primer mixes**

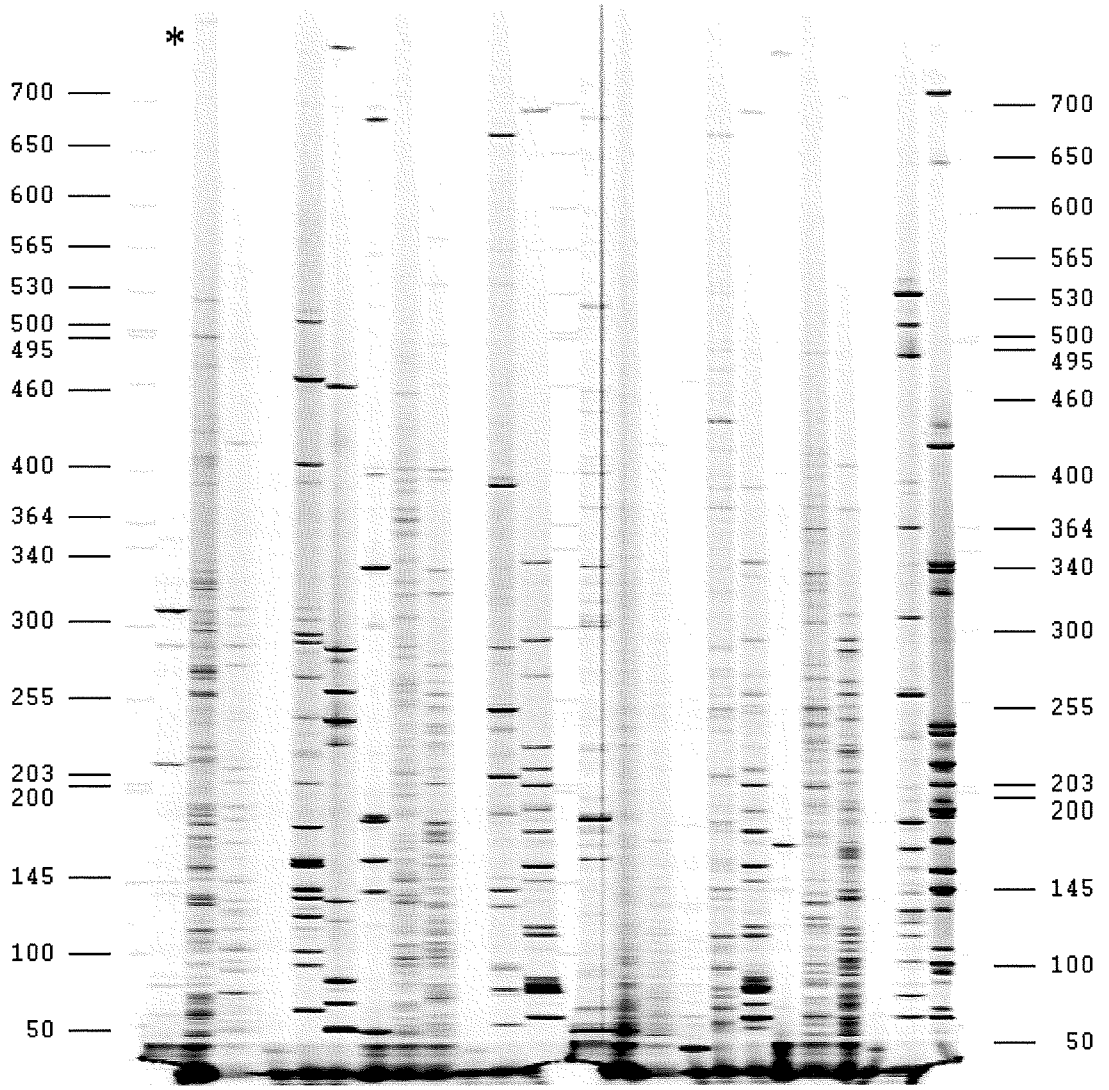



**Fig 5.4: AFLP reactions on the gDNA of *A. platensis*, *R. metallidurans*, and *R. rubrum* using *Apa-Taq* and *Hin-Taq* templates. The loading scheme follows the grid displayed in Fig. 2, i.e. first the top row T01 is loaded then the T02 row, and soforth, starting from the first reaction marked with an asterisk.**

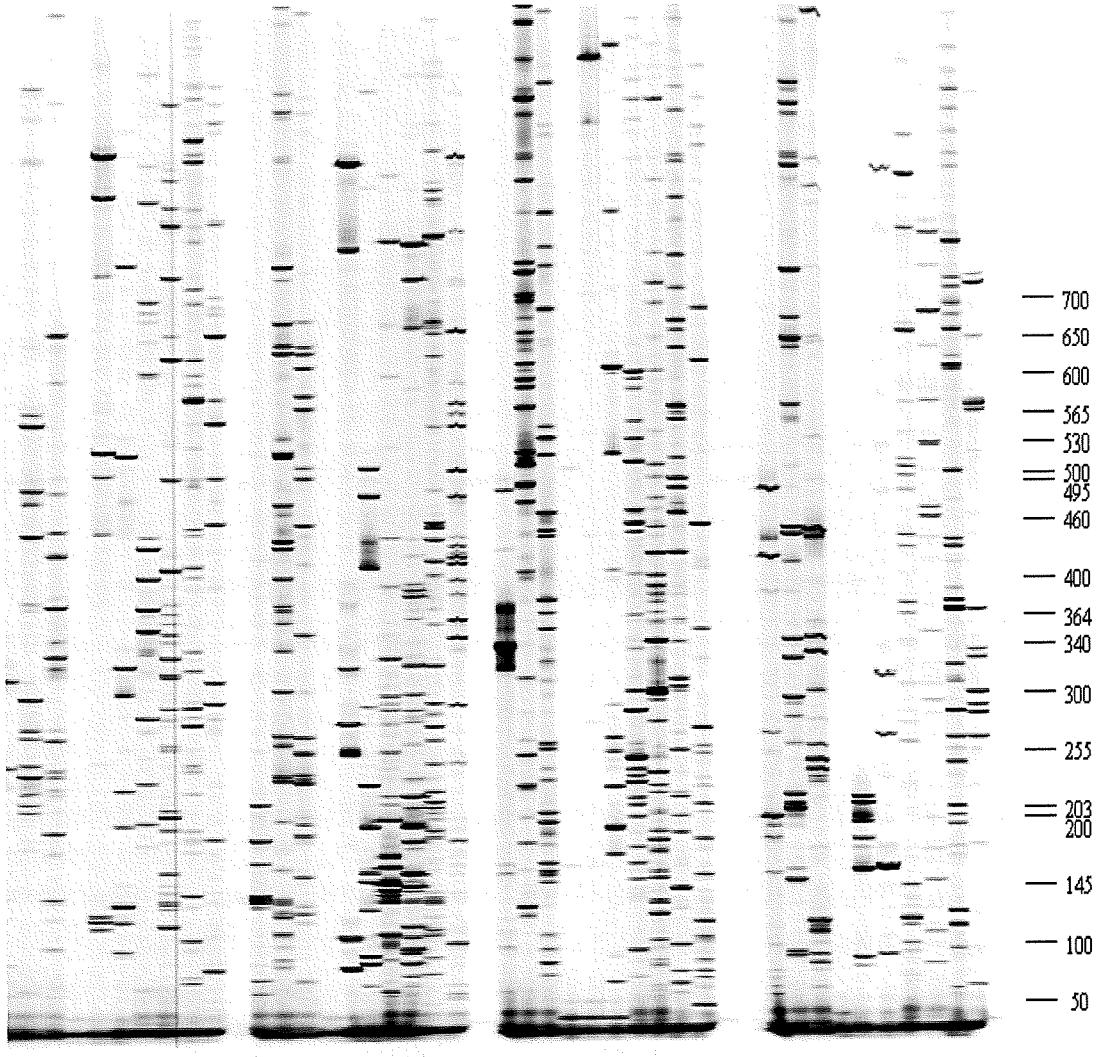



**Fig 5.5: AFLP reactions on the gDNA from *A. platensis*, *R. metallidurans*, and *R. rubrum* using *Apa-Mse*, *Hin-Mse*, *Sal-Mse*, and *Xho-Mse* templates. The loading scheme follows the grid displayed in Fig. 3, i.e. first the top row M01 is loaded then the M02 row, etc.**

Again, AFLP patterns obtained for *A. platensis* gDNA only contained 10 bands or less, all of small size (< 460 bp). Clearly, either this DNA is heavily degraded or methylated.

| Organism | EC | PC | nr. Bands | av. size |
|---|---|---|---|---|
| RMET | Apa-Mse | A01-M01 | 26 | 350 |
| | | A01-M02 | 40 | 460 |
| | | A01-M03 | 45 | 500 |
| | | A01-M04 | 37 | 500 |
| | Hin-Mse | H01-M01 | 11 | 420 |
| | | H01-M02 | 7 | 250 |
| | | H01-M03 | 2 | - |
| | | H01-M04 | 10 | 140 |
| | Sal-Mse | S02-M01 | 24 | 460 |
| | | S02-M02 | 32 | 280 |
| | | S02-M03 | 42 | 450 |
| | | S02-M04 | 30 | 460 |
| | Xho-Mse | X02-M01 | 52 | 500 |
| | | X02-M02 | 50 | 420 |
| | | X02-M03 | 46 | 500 |
| | | X02-M04 | 44 | 500 |
| RRUB | Apa-Mse | A01-M01 | 25 | 350 |
| | | A01-M02 | 31 | 380 |
| | | A01-M03 | 40 | 460 |
| | | A01-M04 | 32 | 420 |
| | Hin-Mse | H01-M01 | 9 | 280 |
| | | H01-M02 | 15 | 200 |
| | | H01-M03 | 16 | 340 |
| | | H01-M04 | 9 | 380 |
| | Sal-Mse | S02-M01 | 33 | 460 |
| | | S02-M02 | 38 | 300 |
| | | S02-M03 | 48 | 400 |
| | | S02-M04 | 19 | 400 |
| | Xho-Mse | X02-M01 | 22 | 460 |
| | | X02-M02 | 34 | 380 |
| | | X02-M03 | 34 | 380 |
| | | X02-M04 | 18 | 340 |

**Table 5.2: various enzyme combinations (ECs) and AFLP primer combinations (PCs) tested on genomic DNA of *R. metallidurans* (RMET) and *R. rubrum* (RRUB). Shaded lines depict optimal ECs and PCs for the used equipment. Average sizes of obtained fragments are in basepairs.**

## 5.4 Conclusion

AFLP looks to be promising for the analysis of mutations in *R. rubrum*. If contaminant DNA would be present in the sample, it could not lead to misinterpretations. Microbial contamination would result in a mixture of banding patterns (depending on the degree of contamination of course), while a mutation may result in only an electrophoretic shift of one particular band (amplicon). A mutation however not necessarily will show up in AFLP. Keep also in mind that there exist three forms of mutation: insertion, deletion, and nucleotide substitution (within the recognition site[s] for the restriction enzyme[s] used).

The use of *Arthrospira* sp. DNA is very problematic. It is easily degraded, even while stored at – 20°C, and may have undergone sequence-specific methylation, preventing proper digestion, or alternatively the preparation of this DNA has co-purified enzymatic inhibitors. It will thus be necessary to focus on alternative gDNA preparations from *A. platensis* before AFLP analysis on this species can be continued.

The *Mse*I enzyme is clearly a better choice above *Taq*I, and the enzyme and primer combinations that we suggest for future AFLP analyses on *R. metallidurans* (RMET) and *R. rubrum* (RRUB) are boxed in table 4.2.

# 6 DNA PREPARATION OF *ARTHROSPIRA* SP PCC8005

## *6.1 Introduction*

The procedure we used follows closely the proposed protocol of Wu *et al.* and Zhang *et al.*, and basically consists of a) cell lysis and b) DNA purification. Cells are first washed, then pretreated sequentially with sarkosyl, lysozyme, SDS, and finally proteinase K, and actual lysis is achieved by adding phenol. Purification of DNA was done by chloroform / isoamylalchol extraction and precipitation in isopropanol in the presence of a high molarity of sodiumchloride (Addendum A). To our surprise, a large amount of precipitate was obtained (Fig. 6.1.) and we suspected that this was not all genomic DNA (gDNA). The precipitate was dissolved in MQ-water of $T_{0.1}TE$ and the concentration of DNA and impurities was measured on an Eppendorf Photometer. A sample was taken for electrophoresis and the dissolved gDNA was again subjected to a phenol extraction, a chloroform / isoamylalcohol extraction, and a preicipation by isopropanol and ammonium acetate. From this, a second sample was taken.

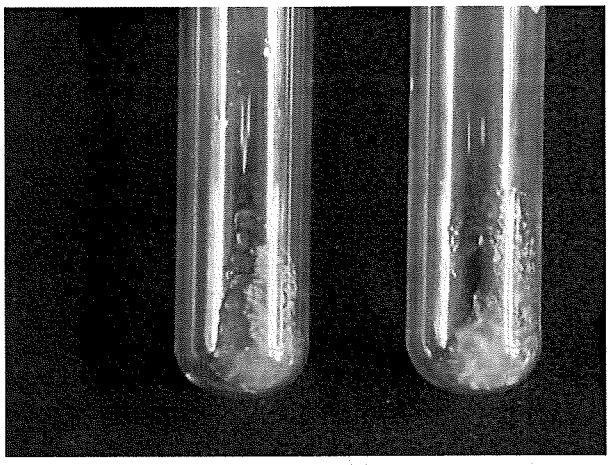


**Fig. 6.1: precipitate after the first series of phenol/$CHCl_3$:IA extractions**

Both DNA samples were electrophorised on a 1% agarose gel in the presence of etidiumbromide to check the integrity and purity of the DNA (Fig. 6.2). In both cases, DNA of high molecular weight could be observed. However, also a lot of impurities were clearly visible in the well. Co-purified RNA of low $M_r$ was also slightly visible.

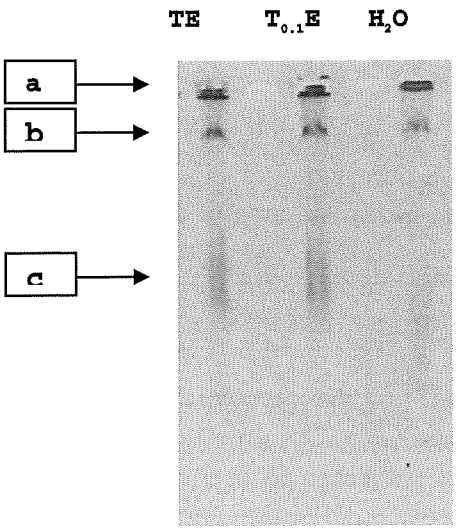



**Fig 6.2: DNA preparation of *Arthrospira* sp. PCC8005 (basic purification) (a) co-purified; (b) genomic DNA; (c) RNA**

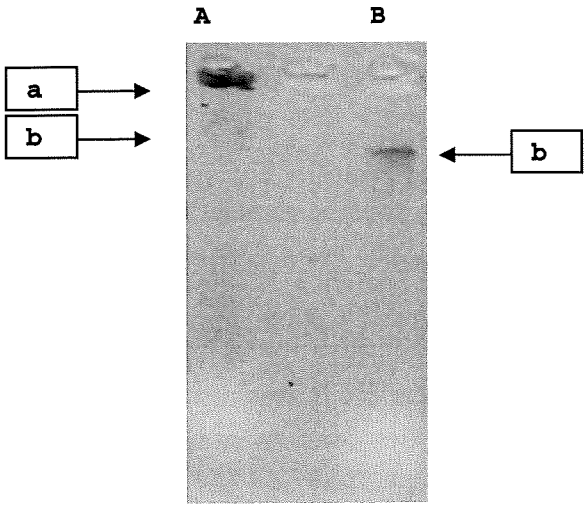



**Fig 6.3: DNA prepartion of A. platensis (additional purification) (a) co-purified impurities; (b) genomic DNA; (A) additional phenol extraction; (B) FastDNA purification**

Because of the poor quality of the gDNA of the basic purification, subsequent clean-up steps were tested using an additional phenol extraction and the FastDNA kit of Bio101. The latter purification is based on the adherence of nucleic acids to silica and is done in the presence of guanidine thiocyanate, a powerful protein denaturant. Purification was done in batch, e.g. by mixing the glass beads with the pre-purified gDNA. The resulting gDNA was also loaded on a 1% agarose gel (Fig. 6.3).

Note that with the additional phenol extraction, most of the gDNA seems to be trapped within the co-purified impurities (Fig.3A). Only by using the in batch purification of the FastDNA kit, higly pure gDNA could be obtained, completely free of impurities. However, the final yield is very low. In total, from 120 mL *A. platensis* cells, only ca. 5 microgram of gDNA was obtained, this is approximately 1000-fold less then one would expect for a full grown *E. coli* culture of the same volume.

## 6.2 Reagents

| | |
|---|---|
| Solution A: | 100 mM Tris pH 8.0, 50 mM EDTA, 100 mM NaCl |
| Solution B: | 50 mM Tris pH 8.0, 5 mM EDTA, 100 mM NaCl |
| Lysis Buffer: | 50 mM Tris pH 8.0, 1 mM EDTA, 25% sucrose |
| TE: | 10 mM Tris pH 8.0, 1 mM EDTA |
| $T_{0.1}E$: | 10 mM Tris pH 8.0, 0.1 mM EDTA |

## 6.3 DNA measurements

The quality of the DNA was assessed by the measurement of the UV absorbance at various wavelengths on an Eppendorf Photometer. In general terms, double stranded DNA is considered of high quality if the A260/A280 value is between 1.8 and 2.1. The A230 value reflects the presence of organic contaminants, while at 320 nm no nucleic acids can be detected and only background values are measured. Hence, both the A230 and A320 values should be as low as possible, but the A260/A280 and A260/A230 ratios should be accordingly high.

| DNA preparation | A230 | A260 | A280 | A320 | A260/A280 | A260/A230 |
|---|---|---|---|---|---|---|
| basic purification | 2.138 | 2.133 | 1.331 | 0.181 | 1.60 | 1.00 |
| additional phenol extraction | 1.135 | 1.632 | 0.929 | 0.053 | 1.76 | 1.44 |
| additional cleanup by FastDNA kit | 0.000 | 0.839 | .393 | 0.143 | 2.14 | NA* |

*not applicable

**Table 6.1: Ultraviolet absorbance measurements of nucleic acid preparations using three different purification methods (UV wavelengths expressed in nanometer)**

As can be seen, only after the second additional cleanup a satisfactory A260/A280 ratio is obtained. Moreover, no contaminants were detected at A230.

## 6.4 Protocol

1. spin down 30 mL (10 min 6,000 rpm) of *A. platensis* grown for 4-5 weeks at RT.
2. wash the cells in 10 mL of solution A
3. spin down 30 mL (10 min 6,000 rpm)
4. resuspend cells in 9 mL solution A, keep at RT for 10 min
5. add 1 mL 1% sarkosyl, keep at 4°C for 2-3 hours
6. spin down (10 min 8,000 rpm) at 4°C

   (note: supernatans should be colored – it was only faintly green)
7. wash cells with 10 mL solution B
8. spin down (10 min 8,000 rpm)
9. resuspend cells in 2 mL LYSIS buffer
10. add 0.2 mL of 50 mg/mL lysozyme; incubate for 30 min at 37C
11. add 0.3 mL of 10% SDS; incubate for 10 min at 37C
12. add 0.2 mL of Proteinase K (1 mg/mL); incubate for 60 min at 37C
13. add 3 mL buffer saturated phenol; mix – not vortex; leave O/N at 4C
14. spin down (30 min 18,000 rpm), take supernatans and repeat phenol step

15. extract the non-organic phase 2x with chloroform:isoamylalcohol (24:1)

16. add to the final supernatans 2/3 volume 5 M NaCl and 1 volume isopropanol

(e.g.: 2.5 mL supernatans + 1.7 mL NaCl + 4.2 mL isopropanol)

17. spin down (10 min 9,500 rpm) in glass pyrex tubes

18. wash the DNA pellet with 70% ethanol

19. dissolve DNA pellet in 0.1 – 0.5 mL $T_{0.1}E$ (depending on the yield)

20. measure DNA concentrations and purity ($A_{230}$ / $A_{260}$ / $A_{280}$)

(21). Because of the impurities seen from the spectrometric values, an additional phenol/chloroform extraction was performed on the dissolved DNA

(22). Additional cleaning step (see text): use FastDNA Kit of BIO101 according to the instructions of the manufacturer.

## 6.5 Conclusion

Although the procedures outlined by Wu and Zhang were meticulously followed, the preparation of gDNA from *A. platensis* appears to be very difficult, most likely owing to problematic cell lysis. It is recommended to try out other methods for cell lysis, for instance in the presence of chaothropic agents such as sodiumiodide, or physical breakage of the cells by French Press or by freeze-thawing with liquid nitrogen (the latter is succesfully used for plant material).

## 6.6 References

1. Wu *et al.* (2000). Plant Mol. Biol. Rep. 18:385.
2. Zhang *et al.* (2003) personal communication

# 7 *RHODOSPIRILLUM RUBRUM* CULTIVATION AND MOLECULAR METHODS

## 7.1 Introduction

*R. rubrum* is a gram-negative motile bacterium, which grows preferably photoheterotrophically under anaerobic conditions in light (Lüderitz and Klemme, 1977). It also grows under microaerobic to aerobic conditions in the dark (Anderson and Fuller, 1967a). Molecular hydrogen may be used as a photosynthetic electron donor during photoautotrophic growth (Anderson and Fuller, 1967b). Autotrophic $CO_2$ fixation is well documented and occurs via ribulose-1,5-bisphoshate carboxylase under autotrophic growth conditions (Anderson and Fuller, 1967a). Under anaerobic dark conditions, fermentation of sugars is possible (Gorrel and Uffen, 1977). *R. rubrum* is also able to perform aerobic dark metabolism with DMSO and trimethylamine N-oxide as electron acceptors (Schultz and Weaver, 1982).

In this chapter growth conditions of *R. rubrum* is being tested. Likewise DNA and RNA extraction methods will be evaluated.

## 7.2 Growth tests of *R. rubrum*

| Medium | Aerobically+dark | Anaerobically+light | Incubation period (days) |
|---|---|---|---|
| **plates** | | | |
| 1408 (100NaCl) | - | - | 15 |
| 1408 (10NaCl) | +/-- | ++ | 15 |
| 1308 | +/-- | ++ | 15 |
| melissa | - | ++ | 15 |
| | | | |
| M9 | - | - | 15 |
| M9 acetate | - | - | 15 |
| 284 | - | - | 15 |
| 284 acetate | - | - | 15 |
| 284 gluconate | +/- | | 13 |
| Sz | - | - | 15 |
| Sz acetate | - | - | 15 |
| Sz lactaat | +/-- | - | 13 |

**Table 7.1. Growth test of *R. rubrum* ATCC25903 on media grown in anaerobic dark and aerobic light conditions. The description of the media used is given in TN70.1.**

| Medium | Aerobically+dark | Anaerobically+light | Incubation period (days) |
|---|---|---|---|
| | | | |
| 869 | + | - | 7 |
| 869 tet (20) | - | | 15 |
| 869 spect (20) | - | | 15 |
| 869 kan (50) | - | | 15 |
| 869 chlora (25) | - | | 15 |
| 869 amp (50) | - | | **15!** |
| 869 carb (100) | - | | **15!** |
| | | | |
| 869 Ni (1µg/ml) | - | | 15 |
| 869 Zi (3µg/ml) | - | | 15 |
| 869 Tell (2mM) | + | + | 7 |
| 869 Tell (5mM) | +/- | +/- | |
| 869 Tell (10mM) | - | - | |
| | | | |
| Sistrom | ++ | ++ | 7 |
| Sistrom –amino acids | ++ | ++ | 7 |
| Sistrom tet (20) | - | | 7 |
| | | | |
| **Liquid media** | | | |
| Melissa LIGHT | + | ++ | 15 |
| Melissa DARK | + | - | 15 |
| 869 LIGHT | + | | 8 |
| 869 DARK | + | | 8 |
| LB DARK | + | | 16 |

**Table 7.1. (continued) Growth test of *R. rubrum* ATCC25903 on media grown in anaerobic dark and aerobic light conditions. The description of the media used is given in TN70.1.**

As indicated in the literature *R. rubrum* grow preferably photoheterotrophically. Rich medium 869, however showed no growth of *R. rubrum* anaerobically or aerobically in the light. Standard minimal medium used in our lab cannot be used for *R. rubrum*. Probably *R. rubrum* needs essential vitamins and growth factors not included in the minimal media (Table 7.1).

*R. rubrum* was sensitive to the antibiotics tested. However, it was observed that *R. rubrum* is able to produce spontaneous rifampicin mutants (100µg/ml) at low frequencies ($10^{-8}$

mutant cfu/total cfu). *R. rubrum* was sensitive to heavy metals, but resistant to levels of tellurium up to 5mM (Table 7.1).

## 7.3 *Effect of temperature and vacuum on R. rubrum growth*

*R. rubrum* was grown during 1 month on MELiSSA medium, under anaerobic light conditions at 30°C. Thereafter *R. rubrum* growth was tested on Sistrom/869 plates under vacuum/aerobic, light/dark conditions at 30/22°C. Vacuum was obtained using a HFE Vacuum Systems vacuum machine Compact Line (C-Tech, The Netherlands).

| Growth on 869 medium | | |
|---|---|---|
| Temperature | 22°C | 30°C |
| Vacuum; Dark | ++, red | ++, red |
| Open; Dark | ++, red | ++, red |
| Vacuum; Light | - | - |
| Open; Light | - | - |

**Table 7.2: *R. rubrum* growth, tested on 869 plates under vacuum/aerobic, light/dark conditions at 30/22°C**

| Growth on Sistrom without amino acids | | |
|---|---|---|
| Temperature | 22°C | 30°C |
| Vacuum; Dark | ++, red | ++, pink to pale |
| Open; Dark | ++, pink to pale | ++, pink to pale |
| Vacuum; Light | ++, centre of colony is red | ++, red |
| Open; Light | ++, red | ++, red |

**Table 7.3: *R. rubrum* growth, tested on Sistrom (- amino acids) plates under vacuum/aerobic, light/dark conditions at 30/22°C**

869 medium didn't allow growth of *R. rubrum* in light conditions. In sistrom medium *R. rubrum* was never shown to be inhibited for growth. Pigmentation was observed in rich medium in dark conditions and in minimal medium in light conditions, at any temperature and in vaccuum as wel as in open conditions. In dark Sistrom conditions pigmentation was only observed in vaccuum, at 22°C.

## 7.4 *Testing dilution drop growth with R. rubrum*

*R. rubrum* was grown during 1 month on MELiSSA medium, under anaerobic light conditions at 30°C. 1.5ml was harvested in a fresh eppendorf and diluted with a salt solution ($MgSO_4$ 10mM). OD was checked on the undiluted, 10 x diluted and 100 x diluted suspension at 680nm.

| | | |
|---|---|---|
| OD: | undiluted | 2.616 |
| | 10 x | 0.730 |
| | 100 x | 0.058 |

The OD of the undiluted sample is larger than 1, and can therefore not be used as a reliable OD measurement. The theoretical OD of the undiluted sample was hence: 10*(0.730+0.580)/2=6.55

### 7.4.1 CORRELATING OD WITH CFU

The culture was diluted up to $10^{-6}$. For each dilution 20µl was dropped on Sistrom enriched with 3% peptone and 1% yeast and incubated at 30°C. At dilution $10^{-5}$, 61 colonies could be counted. At dilution $10^{-6}$ only 8 were counted. Because only 20 µl was dropped on the plate, it was possible to calculate that (61*5) colonies would be counted in a 100µl drop. Taking into account the dilution of $10^{-5}$ and after counting the cfu's in function of 1ml, a concentration of $3.05*10^8$ cfu/ml could be calculated. Using the same procedure with the 8 counted colonies in the $10^{-6}$ dilution, a concentration of $4*10^8$ was obtained. The mean value is hence $3.5*10^8$cfu's/ml.

An OD of 6.55 was hence comparable with $3.5*10^8$ cfu's/ml. Hence for OD 1 would correlate with $5.38*10^7$ cfu's/ml.

### 7.4.2 TESTING THE OD MEASUREMENT USING TWO DIFFERENT WAVELENGTHS AND GROWTH MEDIA

Because it was only possible to use a photospectrometer, measuring OD using a wavelength of 600 nm in the 'lab' in Baikonour, for the MESSAGE experiment, it was needed to compare OD measurements at 600 nm as well as 680 nm. The culture was grown in Sistrom medium unless otherwise stated. Before OD measurements the culture was spinned down, washed in $MgSO_4$ and diluted in $MgSO_4$-solution as indicated.

| OD | 680 | 600 |
|---|---|---|
| Blank | 0 | 0 |
| Rub. Culture | 1.1 | 1.131 |
| 10th dilution | 0.165 | 0.134 |

| Dilution | 1/2 | 1/4 | 1/8 | 1/16 | 1/32 |
|---|---|---|---|---|---|
| 600nm | 0.977 | 0.542 | 0.253 | 0.170 | 0.048 |
| 680nm | 0.943 | 0.506 | 0.232 | 0.104 | 0.027 |

| Media | 869 | Sistrom | Melissa |
|---|---|---|---|
| 600nm | 1,250 | 0,585 | 0,861 |
| 680nm | 1,232 | 0,620 | 0,860 |

**Table 7.4: OD measurements of *R. rubrum* cultures, measured in different media, dilutions and wavelenghts.**

Hence one can conclude that OD measurements using 600 nm can likewise be used for *R. rubrum* cultures.

## 7.5 *RNA extractions with R. rubrum*

*R. rubrum* was grown during 1 month on MELiSSA medium, under anaerobic light conditions at 30°C. 20 µl, 50 µl and 70 µl of a 10 to 100 times diluted culture were spotted on 869 and incubated during 10 days at 30°C, aerobically in the dark.

(R1+R4), (R2+R5), and (R3+R6) are patches grown from a culture of dilution $10^0$, $10^{-1}$ and $10^{-2}$ respectively. They were disolved in 3ml glycerol 15%, NaCl 0.85%. 1 ml was used for RNA extraction. R7, R8, and R9 are samples from a liquid culture in MELiSSA medium grown for 1 month (OD 2.460, at 680 nm). The samples were pelleted and resuspended in a glycerol 15%, NaCl 0.85% solution resulting in different final OD values, as indicated in the table.

The Promega RNA extraction kit was used to extract RNA and stored at -80°C.

| Sample | OD 600 nm | 260 nm | 280 nm | 320 nm | Ratio | Conc. µg/ml | Volume µl | Total conc µg |
|---|---|---|---|---|---|---|---|---|
| R1 | 1.850 | 1.675 | 0.791 | 0.015 | 2.138 | 66.45 | 100 | 6.645 |
| R2 | 1.802 | 2.208 | 1.045 | 0.012 | 2.126 | 87.80 | 100 | 8.780 |
| R3 | 1.765 | 1.401 | 0.667 | 0.040 | 2.170 | 54.45 | 100 | 5.445 |
| R4 | 1.844 | 1.472 | 0.710 | 0.041 | 2.138 | 57.25 | 100 | 5.725 |
| R5 | 1.696 | 1.230 | 0.582 | 0.013 | 2.140 | 48.70 | 100 | 4.870 |
| R6 | 1.608 | 1.385 | 0.648 | 0.017 | 2.165 | 54.70 | 100 | 5.470 |
| R7 | 1.080 | 0.000 | 0.016 | 0.029 | 2.200 | -1.15 | 100 | / |
| R8 | 0.570 | -0.078 | -0.025 | 0.026 | 2.050 | -4.15 | 100 | / |
| R9 | 0.290 | -0.134 | -0.052 | 0.019 | 2.140 | -6.15 | 100 | / |

**Table 7.5: RNA extraction yields from *R. rubrum* patches grown from differently concentrated inoculation drops (R1-R6) as well as from 1 month old cultures (R7-R9)**

- The extractions of the patches were satisfying in relation to the RNA yield and purity.
- RNA extraction of a 1 month old culture does not result in a detectable yield of RNA, probably due to lack of metabolic activity at the far stationary phase.

## 7.6 *Total genomic DNA extraction from R. rubrum , adapted from Bron & Venema (1972)*

### 7.6.1 SOLUTIONS AND EQUIPMENT

- 3 waterbaths (for 37°C, 60°C and 65°C)
- SSC (Salina Sodium Citraat) (175.32g NaCl; 88.23g Na-Citrate; 1L bidest)
- Lysozyme solution (stock of 200 mg/10ml)
- 10% SDS (Sodium Dodecyl Sulfate)
- proteïnase K (20µg/ml)
- Fenol solution (Tris saturated)
- Chloroform solution
- RNase solution (10µg/ml)
- Na-acetate solution (3M Na-acetate, saturated with acetic acid (11.5ml/100ml); pH4.8)
- Ethanol (100%, 75%)

### 7.6.2 PROTOCOL

1. Grow a 3ml culture in Sistrom medium for 5 days.

2. 1.5 ml of the culture is spinned down for 5 minutes at 11000 rpm.

3. Addd another 1.5 ml of the culture and spin down for 5 minutes at 11000 rpm.

4. Add 300µl of the SSC-lysozyme solution (with 20mg lysoszyme/10ml), add 3.3 µl RNase solution and incubate for 10-20 min at 37°C.

5. Add 17 µl 10%SDS to the mixture and incubate for 10 minutes at 65°C.

6. Add 3.3 µl proteinase K and incubate at 60°C for 60 min.

7. Add 300 µl phenol solution to the mixture and mix gently. Centrifuge for 20 min at 11000 rpm at 4°C.

8. Remove yellowish upper phase, add 200µl bidest sterile water and spin again 20 min 11000 rpm at 4°C.

9. Take the upper phase, add 300 µl Chloroform solution. Centrifuge for 20 min at 11000 rpm at 4°C.

10. Take the upper phase and add 1/10 of the Na-acetate solution and 100% Ethanol, so that the final concentration of ethanol is 75%.

11. Incubate at -20°C for 2hours.

12. Centrifuge at 11000 rpm for 20 min to pellet the DNA.

13. Wash the DNA 2 times with 75% ethanol.

14. If necessary additional purification can be accomplished with a final purification step from a commercial kit.

**Fig7.1: DNA extraction of *R.rubrum* cultures.**

DNA extraction using the adapted protocol, resulted in high yields of pure DNA.

## *7.7 Growth of R. rubrum on C1 effluent*

80 ml of reactor 1 (low pH) or reactor 2 (normal pH) was centrifuged. Supernatant was recuperated and the pH was adjusted to 7. 40 ml of both supernatant were filtered.

*R. rubrum* was grown in 3ml Sistrom medium , shaken, at 30°C, aerobically in the dark. Each time 100µl of the *R. rubrum* culture was used to inoculate:

- sistrom medium (positieve control)
- sterile supernatant from reactor 1
- sterile supernatant from reactor 2
- non sterile supernatant from reactor 1
- non sterile supernatant from reactor 2

All cultures were incubated under anaerobic conditions, in the light at 30°C. OD was measured regularly (680nm). *R. rubrum* grew best in the sterile and non sterile effluent of reactor 1 (acid pH), suggesting that *R. rubrum* was not stressed by growing on the C1 effluent.

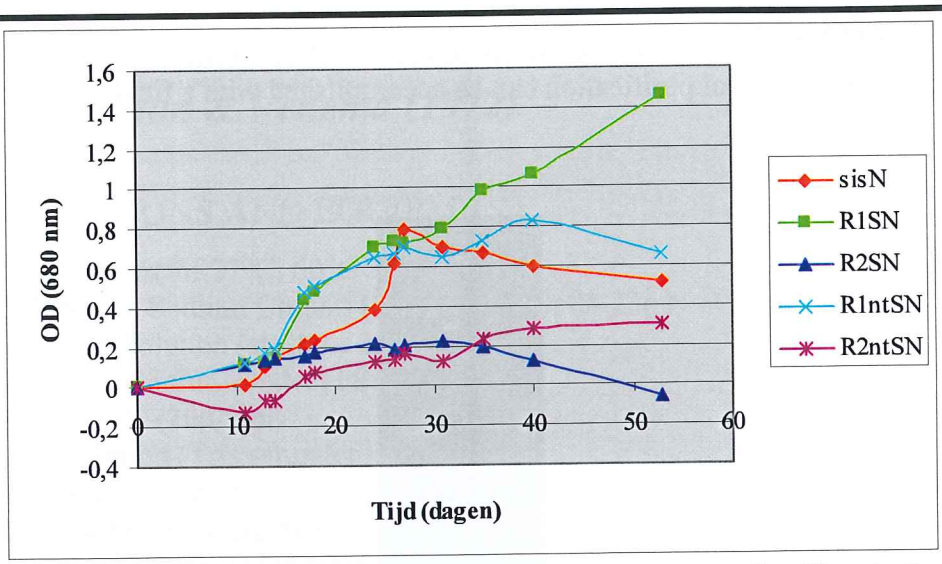



**Fig. 7.2: Growthtest of *R. rubrum* in sterile and non-sterile effluent of compartment 1 (with either inhibited methanogenesis R1, or allowed methanogenesis R2). Sistrom medium was used as control.**

## *7.8 Growth of R. rubrum in heterotrophic growth medium*

### 7.8.1 CONTENT

| | |
|---|---|
| Biotin soln. (Stock: 1 mg/ml) | 1.0ml |
| HEFN soln. (see below) | 10.0ml |
| MgSO47H2O soln. (Stock: 25 g/l) | 10.0ml |
| CaCl2 soln. (Stock: 10 g/l) | 10.0ml |
| *NH4Cl | 1.0g |
| DL-Malic acid | 4.0g |
| MOPS | 11.0g |
| *Yeast Extract (Difco) | 3.0g |
| *Casamino Acids (Sigma C-0626) | 3.0g |

OMIT starred (*) items for photosynthetic N2-dependent growth.

### 7.8.2 TO PREPARE MEDIUM:

Adjust volume to 800 ml with tap distilled water. Adjust pH to 6.8 with NaOH (ca. 3 g pellets), adjust vol. to 1 liter, dispense 100 ml per 250 ml bottle. Autoclave 20 min. Before use add 0.5 ml 200X PO4 (see below) per liter of medium.

For anaerobic growth in serum vials, sparge medium with N2 and dispense 10 ml/121 ml serum vial under N2. Seal vials and flush and evacuate 2X with N2; relieve over-pressure before autoclaving.

Note that nitrogenase activity accumulates when a poor nitrogen source (e.g. L-glutamic acid at 4 g per liter) is included in the medium and the good N-sources are absent; this obviates the need for N2-dependent growth.

### 7.8.3 OTHER NECESSARY SOLUTIONS

HEFN soln. (per liter distilled water): H3BO3, 0.28g; EDTA, 2.0g; Fe Citrate, 0.4g; Na2MoO4, 0.1g (dihydrate OK). Prepare as follows: to 500 ml distilled water add the Fe Citrate, heat to dissolve (place in a steamer at 100°C). Add remaining components and 300 ml additional distilled water, adjust pH to 7.0. Heat gently to dissolve all components and adjust the volume to 1 liter. Store at 4°C.

200X Ormerod's PO4:18 g K2HPO4 (MW 174.18) + 12 g KH2PO4 (MW 136.09), glass distilled water to 100 ml, pH 7.0, autoclave.

### 7.8.4 HETEROTROPHIC GROWTH

Use rich medium and incubate at 30°C aerobically in the dark or anaerobically in the light.

*Rhodospirillum rubrum* will also grow aerobically on standard rich medium (LB), though more slowly than on the rich medium described above.

Cultures may be preserved as DMSO stocks (0.9 ml culture + 0.1 ml DMSO, store at -80°C) or, for shorter term preservation, as stabs.

CO-dependent growth: Replace the malic acid with 0.82 g NaAcetate (anhyd.), add 200µl of a 50mM stock of NiCl2.H2O. Adjust the medium pH to 7.0 by adding ca. 0.9 g NaOH pellets. Prepare cultures anaerobically and add 0.1 ml of anaerobically-prepared 1% Na2S and 0.25 ml of anaerobically-prepared 0.5 M NaHCO3 prior to inoculation. Note that cultures will use CO oxidation as an energy source, while the acetate serves as the primary non-fermentable carbon source.

This recipe for heterotrophic media was kindly provided by Richard Kerby from the workgroup of Prof. Gary Roberts (University of Wisconsin, Madison).

## 7.9 *Growth of R. rubrum in autotrophic medium (2X) [1X medium at 10μM Ni, 2 μM Zn, 10 μM Mn]*

### 7.9.1 CONTENT

| | |
|---|---|
| Biotin soln. (1 mg/ml) | 10.0ml |
| HEFN soln. | 10.0ml |
| MgSO4.7H2O soln. (25 g/l) | 10.0ml |
| CaCl2 soln. (10 g/l) | 10.0ml |
| NH4Cl | 1.0g |
| MOPS | 2.2g |
| NiCl2.6H2O (50 mM) | 200.0μl |
| ZnSO4.7H2O (100 mM) | 20.0μl |
| MnSO4.H2O (59.2 mM) | 169.0μl |

### 7.9.2 PREPARATION

Solution needs to be adjusted to 400 ml with tap distilled water. Adjust pH to 7.0 with NaOH (ca. 0.1 g pellets), adjust vol. to 500 ml, dispense 100 ml per 250 ml bottle. Autoclave 20 min.

For Serum Vials dilute 100 ml medium with 100 ml tap distilled water (to 1X), flush with N2, dispense 15 ml/121 ml serum vial under N2. Seal vials and flush and evacuate 3X with N2; relieve overpressure before autoclaving.

200X Ormerod's PO4:18 g K2HPO4 (MW 174.18) + 12 g KH2PO4 (MW 136.09), glass distilled water to 100 ml, pH 7.0, autoclave.

This medium works OK for photoautrophic growth, though a bit of yeast extract (0.5 g/ 1 liter 1X medium) will speed up the growth (and with the right controls can be shown not to serve as a carbon or energy source). However, this medium should not be considered optimal!

This recipe for autotrophic medium was kindly provided by Richard Kerby from the workgroup of Prof. Gary Roberts.

## 7.10 *Conclusions*

*R. rubrum* grew best on C1 effluent. *R. rubrum* is a versatile organism, able to grow on heterotrophic as well as autotrophic media. 869 was not a very good medium for *R. rubrum* to grow in. In controled conditions Sistrom medium without aminoacids is the medium of choice for quick and easy experimental procedures. In liquid cultures addition of aminoacids will speed up the growth, which is sometimes necessary.

RNA extraction using Promega leads to more than satisfying results. DNA extraction was optimised, based from a total genomic extraction procedure described by Bron and Venema (1972). Using wavelength 680nm, 4.83 * $10^7$ cfu's/ml correlates with OD value 1.

## *7.11 References*

Anderson L, Fuller RC. (1967a) Photosynthesis in *Rhodospirillum rubrum*. II. Photoheterotrophic carbon dioxide fixation. Plant Physiol. 1967 Apr;42(4):491-6.

Anderson L, Fuller RC. (1967b) Photosynthesis in *Rhodospirillum rubrum*. I. Autotrophic carbon dioxide fixation. Plant Physiol. (42) 487-90.

Bron, S. and Venema, G. (1972). Ultraviolet inactivation and excision-repair in *Bacillus subtillis*. Construction of a transformable 8-fold auxotrophic strain and to Ultra-Violet sensitive derivatives. Mut. Res. (15) 1-10.

Luderitz R, and Klemme JH. (1977). Isolation and characterization of a membrane-bound pyruvate dehydrogenase complex from the phototrophic bacterium *Rhodospirillum rubrum*. Z Naturforsch (32) 351-61. German.

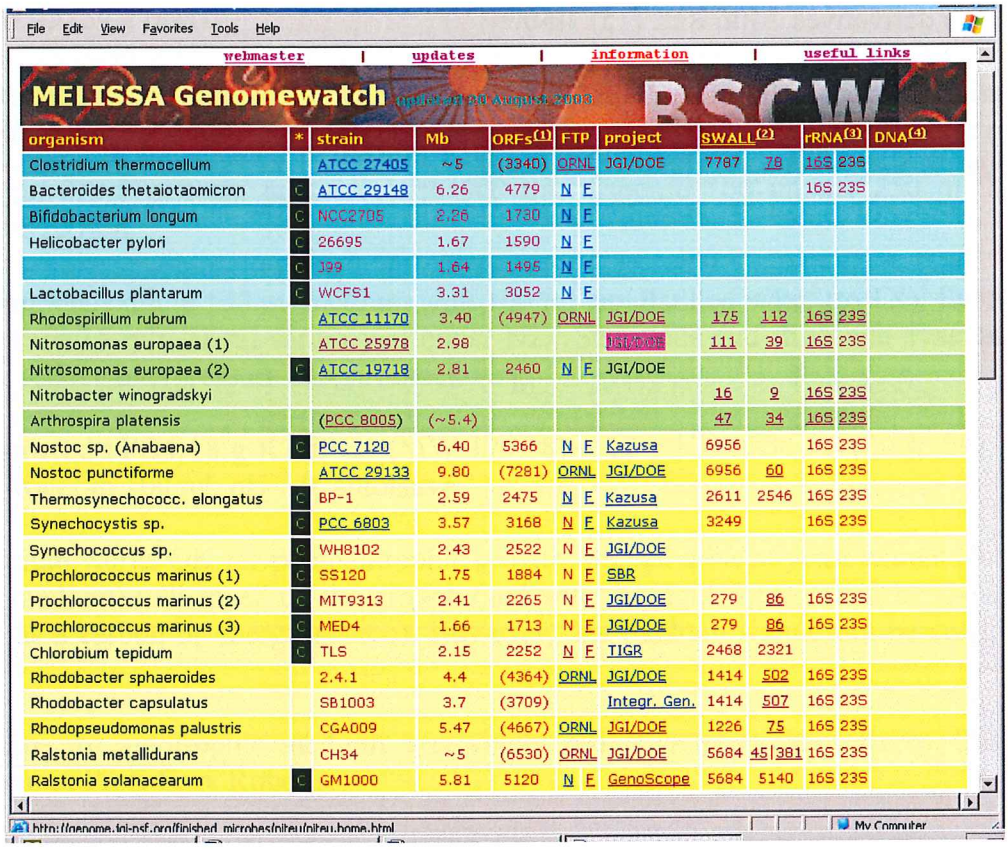
# 8 GENOMEWATCH WEBSITE

## 8.1 Introduction

The MELiSSA Genome Watch Website (MGWW) was established to keep track of all the new genomic data relevant to the MELiSSA project and available via the world wide web.

## 8.2 The MGWW website

The general layout of the MELiSSA Genome Watch Website (MGWW) has been explained in detail in TN70.1 (section 5).




| organism | * | strain | Mb | ORFs(1) | FTP | project | SWALL(2) | | rRNA(3) | DNA(4) |
|---|---|---|---|---|---|---|---|---|---|---|
| Clostridium thermocellum | | ATCC 27405 | ~5 | (3340) | ORNL | JGI/DOE | 7787 | 78 | 16S 23S | |
| Bacteroides thetaiotaomicron | C | ATCC 29148 | 6.26 | 4779 | N F | | | | 16S 23S | |
| Bifidobacterium longum | C | NCC2705 | 2.26 | 1730 | N F | | | | | |
| Helicobacter pylori | C | 26695 | 1.67 | 1590 | N F | | | | | |
| | C | J99 | 1.64 | 1495 | N F | | | | | |
| Lactobacillus plantarum | C | WCFS1 | 3.31 | 3052 | N F | | | | | |
| Rhodospirillum rubrum | | ATCC 11170 | 3.40 | (4947) | ORNL | JGI/DOE | 175 | 112 | 16S 23S | |
| Nitrosomonas europaea (1) | | ATCC 25978 | 2.98 | | | JGI/DOE | 111 | 39 | 16S 23S | |
| Nitrosomonas europaea (2) | C | ATCC 19718 | 2.81 | 2460 | N F | JGI/DOE | | | | |
| Nitrobacter winogradskyi | | | | | | | 16 | 9 | 16S 23S | |
| Arthrospira platensis | | (PCC 8005) | (~5.4) | | | | 47 | 34 | 16S 23S | |
| Nostoc sp. (Anabaena) | C | PCC 7120 | 6.40 | 5366 | N F | Kazusa | 6956 | | 16S 23S | |
| Nostoc punctiforme | | ATCC 29133 | 9.80 | (7281) | ORNL | JGI/DOE | 6956 | 60 | 16S 23S | |
| Thermosynechococc. elongatus | C | BP-1 | 2.59 | 2475 | N F | Kazusa | 2611 | 2546 | 16S 23S | |
| Synechocystis sp. | C | PCC 6803 | 3.57 | 3168 | N F | Kazusa | 3249 | | 16S 23S | |
| Synechococcus sp. | C | WH8102 | 2.43 | 2522 | N F | JGI/DOE | | | | |
| Prochlorococcus marinus (1) | C | SS120 | 1.75 | 1884 | N F | SBR | | | | |
| Prochlorococcus marinus (2) | C | MIT9313 | 2.41 | 2265 | N F | JGI/DOE | 279 | 86 | 16S 23S | |
| Prochlorococcus marinus (3) | C | MED4 | 1.66 | 1713 | N F | JGI/DOE | 279 | 86 | 16S 23S | |
| Chlorobium tepidum | C | TLS | 2.15 | 2252 | N F | TIGR | 2468 | 2321 | | |
| Rhodobacter sphaeroides | | 2.4.1 | 4.4 | (4364) | ORNL | JGI/DOE | 1414 | 502 | 16S 23S | |
| Rhodobacter capsulatus | | SB1003 | 3.7 | (3709) | | Integr. Gen. | 1414 | 507 | 16S 23S | |
| Rhodopseudomonas palustris | | CGA009 | 5.47 | (4667) | ORNL | JGI/DOE | 1226 | 75 | 16S 23S | |
| Ralstonia metallidurans | | CH34 | ~5 | (6530) | ORNL | JGI/DOE | 5684 | 45\|381 | 16S 23S | |
| Ralstonia solanacearum | C | GM1000 | 5.81 | 5120 | N F | GenoScope | 5684 | 5140 | 16S 23S | |

**Fig. 8.1: layout of the MELiSSA Genome Watch Website (as of 20 February 2004). Blue: relevant to compartment C1. Green: (principle organisms) relevant to compartments C2-C4. Yellow: organisms that bear some relevance to compartments C2-C4. Details in TN70.1, section 5.)**

The entry page was updated and put in a new format. Information on relevant genome projects (Fig. 1) were constantly added to the MGWW entry page with the number of completed genomes rising (marked with a bold "C"). For the principle MELiSSA strains,

very little additional information was obtained in spite of circa 500 ongoing genome projects and an exponential increase of genome data from other sources.

According to the COGENT (COmplete GENome Tracking) database (Janssen *et al.*, 2003)[2], there are now 139 genomes completely sequenced and published, representing a total of nearly 560,000 proteins as compared to a total of 1,430,000 sequences in the non-redundant protein database at the NCBI. The number of protein sequences originating from complete genome sequencing projects is expected to reach over 2,000,000 entries in the next 15 months.

## *8.3 Conclusions*

Genomic data were updated roughly every two months, until August 2003 (the numbers of available protein sequences per organism were obtained as explained in TN70.1, section 5). In general, there is very little increase in available data, and fluctuations are mainly due to regular updates in the various protein databanks (i.e. reannotations and emended taxonomic positions). The major changes are due to new species whose genome were completely sequenced. These new data have been added to the website in hyperlink format. The last update is from 20-02-2004. Further updates will be done every 4-6 months.

[2] Janssen et al. (2003). Bioinformatics (In Press) (http://maine.ebi.ac.uk:8000/services/cogent/)