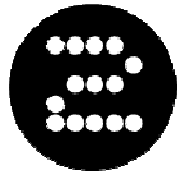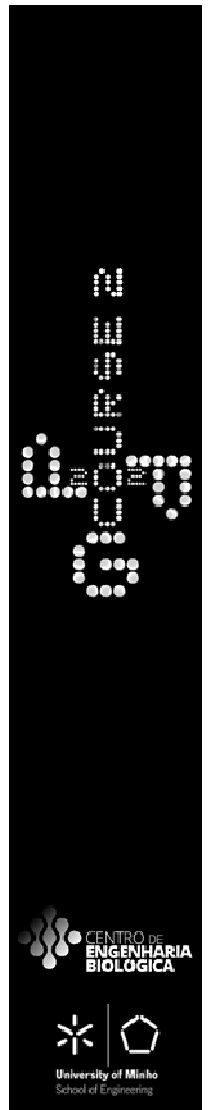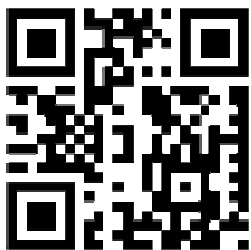# PROTEOME GENE2 PROTEIN³

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

CENTRO DE ENGENHARIA BIOLÓGICA

University of Minho
School of Engineering

BOOK OF THE 2ND PROTEOME 2 GENE 2 PROTEIN (P2G2P) HANDS-ON COURSE

26-29 MARCH 2018, BRAGA, PORTUGAL

EDITORS

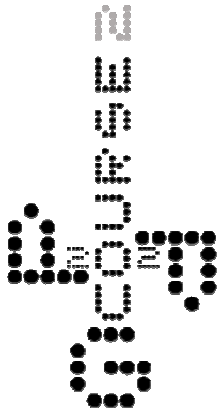Carla Oliveira

Ivone Martins

Sílvio B. Santos

Sónia Silva

Tatiana Q. Aguiar

CONTENTS

PROTEOME ≥ GENE ≥ PROTEIN ≥

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

WELCOME MESSAGE

PROTEOME 2 GENE 2 PROTEIN 2

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

Dear Participants,

We are very pleased to welcome you to the 2nd edition of the "Proteome 2 Gene 2 Protein" (P2G2P) Hands-On Course.

The aim of this 4 days course is to provide you information on proteomic methodologies and the first skills in the techniques of two-dimensional gel electrophoresis, protein identification by mass spectrometry, gene cloning and recombinant protein production and purification.

We wish you all a fruitful course and a pleasant journey in Braga.

The Organizing Committee,

Carla Oliveira (carlaoliveira@deb.uminho.pt)

Ivone Martins (ivone.martins@ceb.uminho.pt)

Sílvio B. Santos (silviosantos@ceb.uminho.pt)

Sónia Silva (soniasilva@deb.uminho.pt)

Tatiana Q. Aguiar (tatiana.aguiar@deb.uminho.pt)
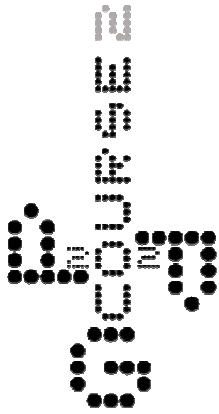
SPONSORS

PROTEOME ≈ GENE ≈ PROTEIN ≈

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

The Organising Committee gratefully acknowledges the support of the following sponsors:

BIOPORTUGAL
Químico, Farmacêutica, Lda.
http://bioportugal.pt

fisher scientific
part of **Thermo Fisher Scientific**
www.fishersci.com

BIO·RAD
http://www.bio-rad.com

frilabo
www.frilabo.pt

Biotecnómica
soluções para biotecnologia e genómica
www.biotecnomica.pt

grisp
Research Solutions
http://grisp.pt

enzymatic
www.enzymatic.pt

nzytech
genes & enzymes
www.nzytech.com

Regina
http://www.chocolatesregina.com

PROGRAM

PROTEOME ⇄ GENE ⇄ PROTEIN ⇄

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

Time	26ᵗʰ March	27ᵗʰ March	28ᵗʰ March	29ᵗʰ March
09h00-12h30		PRACTICAL (P1)	PRACTICAL (P4)	PRACTICAL (P6)
12h30-13h30		LUNCH		
13h30-14h00	Registration			
14h00-14h10				
14h10-14h20	Welcome Session (CEB's Director)			
14h20-14h30	SPBT – Sociedade Portuguesa de Biotecnologia			
14h30-15h00	Lecture 1	PRACTICAL (P2 & P3)	PRACTICAL (P5)	PRACTICAL (P6)
15h00-15h30	Lecture 2			
15h30-16h00	Lecture 3			
16h00-16h45	Lecture 4			
16h45-17h15	Coffee Break			
17h15-18h00	Practical Workshop			

Lecture 1 - Bruno Almeida (ICSV/EM/UMinho)
"2D Electrophoresis: Notions and Application"

Lecture 2 – Clara Sousa (LAQV/REQUIMTE-FF-UP)
"MALDI-TOF MS Principles and Applications"

Lecture 3 - Hugo Osório (IPATIMUP/I3S)
"Applied Protein Mass Spec"

Lecture 4 - Lucília Domingues (CEB-UMinho)
"Recombinant Protein Production: development of new methodologies and biomedical applications"

P 1 - Sónia Silva / Ana Margarida Sousa
2D Electrophoresis

P 2 - Ana Margarida Sousa / Sónia Silva
2D Gel Analysis

P 3 - Carla Silva
Mass Spectrometry Approaches

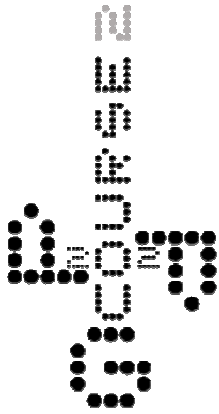**P 4 - Sílvio Santos**
Protein to Gene: Bioinformatics

P 5 - Sílvio Santos / Ivone Martins
Gene Cloning

P 6 - Carla Oliveira / Tatiana Q. Aguiar
Recombinant Protein Production and Purification

LECTURES

PROTEOME ≥ GENE ≥ PROTEIN ≥
Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

Lecture 1 "2-D Electrophoresis: Notions and Application"

Bruno Almeida (ICSV/EM/UMinho) – brunoalmeida@med.uminho.pt

Two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) of proteins has paved the way to the birth of proteomics. Currently, 2-D PAGE is capable of resolving thousands of proteins in a single run, being therefore a key toll of proteomics research. Although it is no longer the only experimental technique used in modern proteomics, it still has distinct features and major advantages. The purpose of this presentation is to give an historic perspective about the field and to give notions about the main steps of the process, from sample preparation to in-gel detection of proteins, commenting the positive features as well as the constraints and caveats of the technique. This will be illustrated by specific examples taken form literature and commented in detail.

Lecture 2 "MALDI-TOF MS Principles and Applications"

Clara Sousa (LAQV/REQUIMTE-FF-UP) – ccsousa@ff.up.pt

During this lecture some mass spectrometry principles will be presented as well as their main limitations which boost the development of MALDI-TOF MS. A detailed presentation about the technique including equipment components, standard experimental procedures and data analysis will be performed. General applications of the technique will be also pointed.

Lecture 3 "Applied Protein Mass Spec"
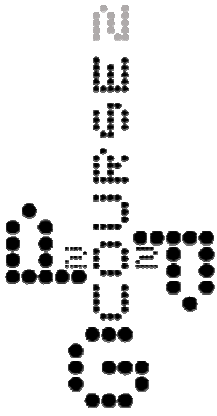
Hugo Osório (IPATIMUP/I3S) – hosorio@ipatimup.pt

Proteomics is one of the best-developed mass spectrometry applications. It encompasses three major topics: 1) high-throughput protein identification; 2) quantitation of protein expression levels; 3) characterization of protein post-translational modifications.

Lecture 4 "Recombinant Protein Production: development of new methodologies and biomedical applications"

Lucília Domingues (CEB-UMinho) – luciliad@deb.uminho.pt

The lecture will give a brief introduction to the mainly used protein expression systems with particular focus on the specificities of each system. Then, an overview of recent developments on novel fusion tags for the bacterial system *Escherichia coli* will be highlighted. Finally, an overview of integrated recombinant protein production and purification strategies in the context of different biomedical applications will be presented.

PRACTICAL SESSIONS
Introduction

Tatiana Q. Aguiar

tatiana.aguiar@deb.uminho.pt

PROTEOME ⮂ GENE ⮂ PROTEIN ⮂

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

The emergence of proteomics, the large-scale analysis of proteins, has been inspired by the awareness that the final product of a gene is inherently more complex and closer to function than the gene itself [1]. Today it is well recognized that any given genome can potentially give rise to an infinite number of proteomes. This happens because the proteome of any biological sample (*e.g.*, cell, tissue, organ, organism, microbial consortium, etc) is dynamic, reflecting the immediate environment in which it is studied. Within a biological system, proteins can be synthesized, degraded, modified by post-translational modifications or undergo translocations in response to internal or external stimuli [1]. Thus, analyzing the proteome of a biological sample is a "snapshot" of the protein environment at that particular given time [1].

Techniques such as two-dimension polyacrylamide gel electrophoresis (2D-PAGE) and mass spectrometry (MS) enable the comparison of proteomes between samples that differ by some variable, providing information that allow us to, among others: (i) characterize proteins involved in protein signaling, disease mechanisms, protein-drug interactions, etc; (ii) identify novel proteins with interesting biological activities (for the pharmaceutical/biotechnological industry); (iii) understand how the expression of certain proteins provides to a biological system its unique structural characteristics. However, often the full analysis of the protein of interest is hampered by the insufficient amount in which it is available in the original sample, even after enrichment. To overcome this limitation, the recombinant production and purification of such protein is often used to obtain suitable high amounts of pure protein for subsequent applications/analyses.

In this course, as a case study (Fig. 1), the proteome of the cellular extracts of the cellulolytic bacterium *Clostridium thermocellum* grown in two different conditions (with (E1) and without (E2) a cellulosic substrate) will be analyzed by 2D-PAGE aiming at identifying and characterizing proteins whose expression is induced by cellulose. After 2D-PAGE profile analysis of these two samples, one spot from the gel corresponding to a protein whose expression was highly induced by cellulose will be selected, excised, and then prepared for MS sequencing. Using bioinformatic tools, the amino acid sequence retrieved after analysis of the MS data will be converted into a gene sequence with optimized codons for recombinant expression in the bacterium *Escherichia coli*. Primers will be

manually designed for gene cloning into a bacterial commercial expression vector and bioinformatics tools will also be used to simulate the *in silico* molecular cloning. Afterwards, the gene of interest will be experimentally amplified, cloned into the expression plasmid and used to transform a suitable *E. coli* expression strain. The recombinant production of the protein of interest will then be induced and finally purified from this bacterial host. SDS-PAGE analysis will enable to monitor protein production and purification. In the end, it is expected that enough amounts of pure protein for subsequent studies/applications are obtained.
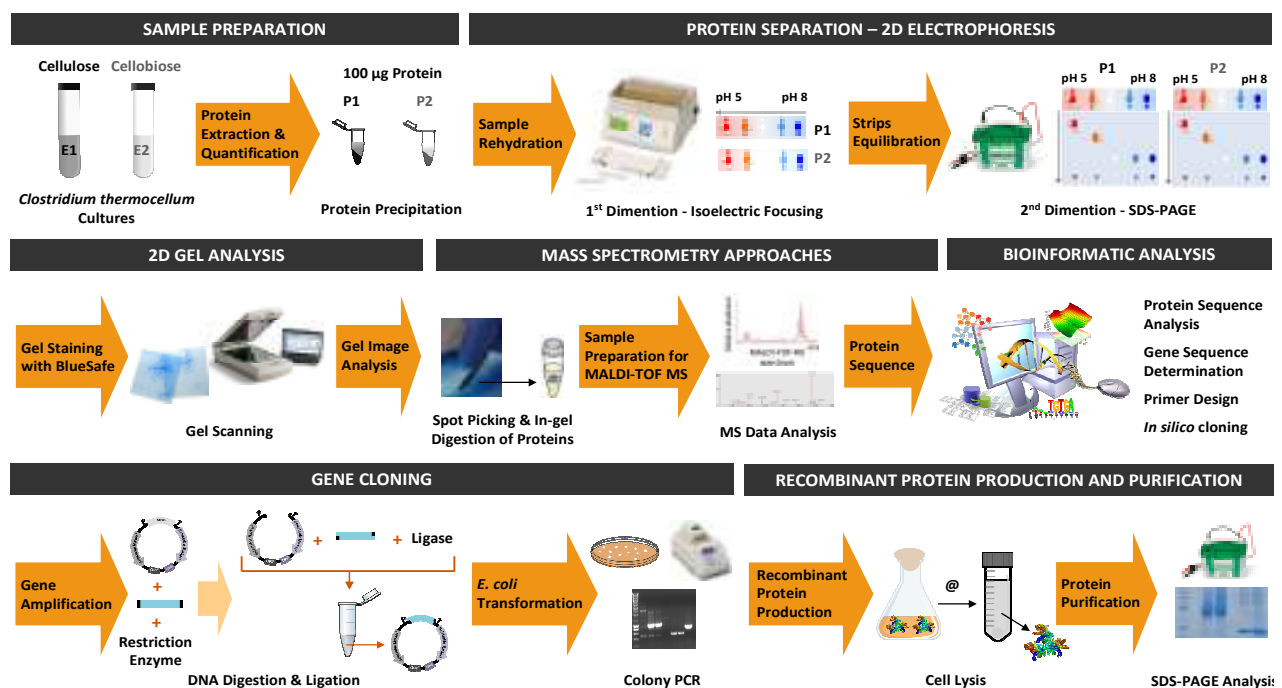
Fig. 1 Workflow of the "Proteome 2 Gene 2 Protein" Hands-on course.
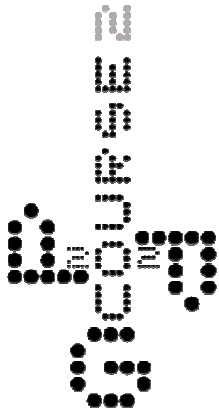
Sónia Silva and Ana Margarida Sousa

soniasilva@deb.uminho.pt

PROTEOME ⇄ GENE ⇄ PROTEIN ⇄

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

The use of proteomic analysis to investigate the particular physiology of microorganisms has been increasing. The protein profiles of microorganisms growing in different conditions reveal alterations in protein expression induced by a specific environmental condition. Protein expression profiles, which significantly vary between two different growth conditions, can be obtained by two-dimension polyacrylamide gel electrophoresis (2D-PAGE).

Clostridium thermocellum is an anaerobic, Gram-positive thermophilic bacterium capable of cellulosome-mediated breakdown of (hemi)cellulose. The *C. thermocellum* cellulosome was discovered in the early 1980s, and its proposed structural model consisting of multienzyme complexes has been supported by several studies [2]. During growth on cellulose, the cellulosome is attached to the cell in early exponential phase, released during late exponential phase, and is found attached to cellulose during stationary phase [3]. The cellulosome expression has been shown to be negatively regulated by growth on cellobiose via a carbon catabolite repression mechanism, and positively regulated by growth on cellulose [3]. To study the global protein expression levels of *C. thermocellum* in the presence of cellulose (E1) and cellobiose (E2), the following protocol will be used to obtain the proteomic profiles of *C. thermocellum* cellular extracts.

2. SAMPLE PREPARATION

2.1. MATERIALS AND REAGENTS

2.1.1. Bacterial cell growth and disruption

1. Anaerobic Balch tubes (26 mL; Bellco Glass Inc., Vineland, NJ).

2. 1191 medium: 1.5 g/L KH_2PO_4; 4.2 g/L $Na_2HPO_4.12H_2O$, 0.5 g/L NH_4Cl, 0.18 g/L $MgCl_2.6H_2O$, yeast extract (BD 212750), 2.0 g; 0.25 mg/L resazurin, 0.50 mL/L vitamin solution, and 1 mL/L mineral solution.

3. Vitamin solution: 20 mg/L biotin, 50 mg/L p-aminobenzoic acid, 20 mg/L folic acid, 50 mg/L nicotinic acid, 50 mg/L thiamine, 50 mg/L riboflavin, 50 mg/L lipoic acid (thioctic acid), and 10 mg/L cyanocobalamin.

4. Mineral solution: 20.2 g/L trisodium nitrilotriacetate, 2.1 g/L $FeCl_3.6H_2O$, 2.0 g/L $CoCl_2.6H_2O$, 1.0 g/L $MnCl_2.4H_2O$, 1.0 g/L $ZnCl_2$, 1.0 g/L $NiCl_2.6H_2O$, 0.5 g/L $CaCl_2.2H_2O$, 0.5 g/L $CuSO_4.2H_2O$, and 0.5 g/L $Na_2MoO_4.2H_2O$.

5. Reducing solution: prepare under nitrogen using sodium sulfide crystals in distilled water to a final concentration of 200 mM.

6. Avicel PH-101 (Sigma-Aldrich).

7. Cellobiose (Sigma-Aldrich).

8. 1x PBS buffer: 8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L Na_2HPO_4, 0.24 g/L KH_2PO_4, adjust pH to 7.4 with HCl.

9. Lysis buffer: 20 mM Tris-HCl (pH 7.0), 20 mM NaCl, 5 mM $CaCl_2$, 1 mM PMSF (phenylmethylsulfonylfluoride).

10. Refrigerated centrifuge.

11. Sonicator.

12. Ice.

2.1.2. Protein quantification

1. Protein samples.

2. BCA Kit.

3. 96 well plates.

4. Microplate Reader.

2.1.3. Protein Precipitation

1. Refrigerated centrifuge.

2. Eppendorfs.

3. Ice.

4. 100% (w/v) ice-cold Trichloroacetic acid (TCA).

5. 100% (w/v) ice cold acetone.

2.2. BACTERIAL CELL GROWTH

C. thermocellum cells were previously grown in the presence of cellulose (E1) and cellobiose (E2) using the following protocol.

1. Grow *C. thermocellum* cells at 60ºC in anaerobic Balch tubes containing 10 mL of 1191 medium (pH 7.0) with 5 g/L carbon source (cellobiose or cellulose).
2. Reduce each tube with 0.1 mL of reducing solution after gassing and degassing them (1:4 min) four times with 100% nitrogen.
3. As inoculum use 10% (v/v) of seed cultures taken during logarithmic growth on cellobiose.
4. Harvest cells at the early exponential phase by centrifugation at 10000 *g*, 5 min, 4°C.
5. Wash cell pellets 3 times with 500 μL of 1x PBS buffer and store at -80°C until further use.

2.3. BACTERIAL CELL DISRUPTION AND PROTEIN EXTRACTION

The method used to disrupt cells is dependent on the type of microorganisms. Usually, physical, such as sonication is the most used method for disruption of bacteria and fungi. Concerning the lysis buffers, it is of upmost importance that they contain protease inhibitors (*e.g.*, PMSF) to prevent protein degradation.

1. Resuspend cell pellets in 1 mL of lysis buffer and sonicate cells for 3 min on ice using the following parameters: 30 s ON and 30 s OFF; amplitude of 35%.
2. Remove cell debris by centrifugation at the maximal speed for 30 min at 4ºC and determine protein concentration in the supernatant (called cell-free extract or lysate).

2.4. PROTEIN QUANTIFICATION

After cell lysis, it is of upmost importance to strictly quantify total sample protein. There are several protocols to determine the total protein concentration, such as the Bradford and Lowry assays, but there are also several commercial kits available that are based on these two reference methods. Here, the protein content will be determined using the Bicinchoninic Acid Protein (BCA) assay kit [4].

1. Mix 1000 μL of "A" solution with 20 μL of "B" solution BCA reagents (50:1) for 1 min.
2. Add 25 μL of each sample into a 96 well plate (in triplicate), add 200 μL of mixed reagents, homogenize during 30 s and incubate 30 min at 37ºC.
3. Measure the absorbance at 562 nm using lysis buffer as blank, prepared similarly to the pure or diluted proteins samples.
4. Determine the protein concentration in accordance with the standard curve (Fig. 2).
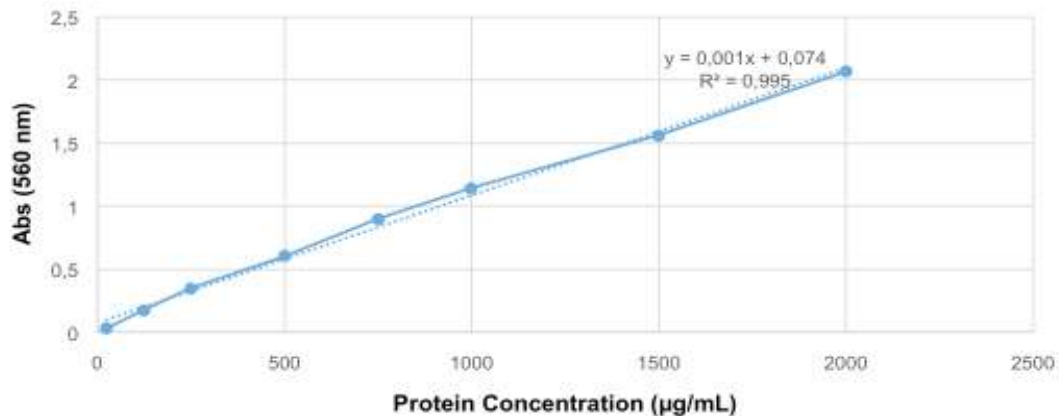
Fig. 2 Standard curve of bovine serum albumin concentration determined with the BCA Kit.

5. Calculate the volume of protein sample to obtain 100 μg of total protein of each sample.

2.5. PROTEIN PRECIPITATION WITH TCA/ACETONE

1. Add a volume of 100% TCA, in order to obtain a final concentration of 20% TCA, to the volume of the protein sample calculated.
2. Vortex the samples.
3. Incubate the mixture for 20 min on ice and centrifuge at 150000 g, 4ºC for 10 min.
4. Remove the supernatant and add 200 μL of ice-cold acetone to wash the pellet.
5. Centrifuge the mixture for 10 min at 15000 g and at 4ºC.
6. Remove the acetone containing the supernatant and dry air the pellet.
7. Resuspend the pellet in an appropriate buffer dependent on the proteomic approach.

3. PROTEIN SEPARATION – 2D ELECTROPHORESIS

2D gel electrophoresis is a form of proteomic methodology commonly used to analyze proteins profiles. In this methodology, the mixtures of proteins are separated by two properties. Firstly, proteins are separated according to their isoelectric point (IP), a process called isoelectric focusing (IEF). Thereby, a gradient of pH is applied to a gel (Ready Strip IPG gel) and an electric potential is adapted across the gel. At all pH values, proteins will be charged. If they are positively charged, they will be pulled towards the more negative end of the gel and if are negatively charged they will be pulled near the positive end of the gel. Thus, the proteins applied in the first dimension will be moved along the gel and will stop at their IP; that is the point at which the overall charge on the proteins is zero (a neutral charge). Following, the proteins complexes will be separated by applying the denaturing SDS-PAGE (second dimension gel) [5, 6].

3.1. MATERIALS AND REAGENTS

3.1.1. Isoelectric Focusing

1. 7 cm Ready Strip IPG gel (Bio-Rad).
2. Protean IEF apparatus (Bio-Rad).
3. Rehydration buffer (Bio-Rad).
4. Mineral oil.
5. SDS-PAGE Equilibration Buffer I (with DTT): 6 M urea, 0.375 M Tris, pH 8.8, 2% SDS, 20% glycerol, 2% (w/v) DTT.
6. SDS-PAGE Equilibration Buffer II (with Iodoacetamide): 6 M urea, 0.375 M Tris, pH 8.8, 2% SDS, 20% glycerol, 2.5 % (w/v) iodoacetamide.
7. Overlay agarose: 0.5% (w/v) low-melt agarose, 1x Tris-glycine-SDS, 1% (w/v) bromophenol blue.
8. Filter paper.
9. 1x Tris-glycine-SDS running buffer.
10. Wicks - specific piece of paper (Bio-Rad).

3.1.2. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

1. SDS-PAGE system (includes a tank, lid with power cables, electrode assembly, cell buffer dam, casting stands and frames, combs and glass plates).
2. Acrylamide/Bis-acrylamide (30%/0.8% (w/v)) (Grisp).
3. Deionized water.
4. 5 M Tris-HCl, pH 6.8.
5. 1.5 M Tris-HCl, pH 8.8.
6. 10% (w/v) SDS.
7. 10% (w/v) ammonium persulfate (APS).
8. Tetramethylethylenediamine (TEMED).
9. Buffer TGS: 25 M Tris, 192 mM glycine, 0.1% (w/v) SDS.

3.1.3. Gel Staining

1. BlueSafe solution (NZYTech).
2. Deionized water.
3. Fixation solution: 50% (v/v) ethanol, 12% (v/v) acetic acid and 0.05% (v/v) formaldehyde.
4. Wash solution: 20% (v/v) ethanol.
5. Sensibility solution: 0.02% (w/v) sodium thiosulfate.
6. Coloration solution: 0.08% (v/v) formaldehyde, 0.2% (w/v) silver nitrate.

7. Reduction solution: 0.05% (v/v) formaldehyde, 0.0004% (w/v) sodium thiosulfate, 6% (w/v) sodium carbonate.

8. Stop solution: 12% (v/v) acetic acid.

3.2. Isoelectric Focusing

3.2.1. Rehydration and Sample Application

1. Add 125 μL of rehydration buffer to each eppendorf with protein samples.

2. Vortex the samples.

3. Transfer each protein sample as a line along the back edge of a channel in a focusing tray.

4. Place the Ready Strip IPG gel side down in the IEF focusing tray with the marked side with + in contact with the anode of the tray. Ensure that the gels make contact with the electrodes.

5. Between the electrodes and the Ready Strip IPG gels place one electrode wick paper *per* connection. The electrode wicks should be wet with water before.

6. Add 500 μL of mineral oil as a line along each strip.

7. Rehydrate under active conditions.

8. Program the Protean IEF apparatus with: 50 V, 20ºC during 12-16 h.

3.2.2. Focusing conditions

It is important to highlight that focusing will vary with sample composition, sample complexity IPG range and the strip IPG gel size. However, the current should not exceed 50 μA/strip and each protocol needs to be optimized in accordance to each apparatus under used. The total of time required for ramping will mainly depend on the protein samples and the strip IPG gel size (e.g. 7 and 18 cm). Three steps for pH 3-10 strips of 7 cm should be applied in accordance with Table 1.

Table 1 Focusing conditions program for 7 cm Ready-IPG strips.

	Voltage	Set Time	Ramp	Temperature
Step 1	250 V	15 min	Rapid	20ºC
Step 2	4,000	1 h	Slow	20ºC
Step 3	4,000	1 h	Rapid	20ºC

3.2.3. Strips equilibration

Prior to running the second dimension it is necessary to equilibrate the IPG strips in SDS-containing buffers. The 2-step equilibration also ensures that cysteines are reduced and alkylated, which minimizes or eliminates vertical streaking that may be visible after staining of the second dimension gels. Equilibration Buffer I contains DTT, which reduces sulfhydryl groups, while Equilibration Buffer II contains iodoacetamide, which alkylates the reduced sulfhydryl groups [5, 6].

Note 1

If the strips were stored at -70ºC, they should be removed from the freezer and placed onto the lab bench to thaw at this time. The IPG strips required 10-15 min to thaw. It is best not to leave the thawed IPG strips for longer than to 15-10 min as the diffusion of the proteins can result in reduced sharpness of proteins spots.

1. Remove the mineral oil from the Ready Strip IPG strips by placing the strips (gel side up) onto a piece of dry filter paper and blotting with a second piece of wet filter paper.
2. Add 1 mL of the equilibration buffer I with DTT to an equilibration/rehydration tray, using a channel per strip.
3. Place the tray on an orbital shaker for 10 min.
4. Discard the used equilibration buffer by carefully decanting the liquid.
5. Add 1 mL of equilibrium buffer II with iodoacetamide to each strip.
6. Return the tray to the orbital shaker for 10 min.
7. During the incubation, melt the overlay agarose solution in a microwave oven.
8. Discard the equilibrium buffer II by decanting.
9. Fill a 100 mL graduated cylinder or a tube that is the same length as or longer than the IPG strip length with 1x Tris-glycine SDS running buffer. Use a Pasteur pipette to remove any bubbles on surface of the buffer.
10. Use pre-caste gels or finish preparing the SDS-PAGE gels in accordance described in **section 3.2.4**.
11. Remove an IPG strip from the disposable tray and lay the strip, with the gel side towards you, onto the back plate of the SDS-PAGE gel above the IPG well. Repeat this process for any remaining IPG strips.
12. Place overlay agarose solution into the IPG well of the gel.

13. Using the tweezers, carefully push the strip into the well, taking care not to trap any air bubbles beneath the strip.

14. Mount the gel into the electrophoresis cell following the instructions provided with the apparatus.

15. Fill the reservoirs with 1x TGS running buffer and begin the electrophoresis.

3.2.4. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The SDS-PAGE is used when the aim is the separation of the proteins by their molecular weight, since SDS will confer the same charge to all proteins. A SDS-PAGE gel includes a running and stacking gel. During this class we will use 12% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad), but the protocol for gels preparation is described below.

1. Prepare the 10% running gel solution as described in Table 2 and mix the solution thoroughly.

Table 2 Volumes for 40 mL of running gel and for 5 mL of stacking gel, corresponding to a gel of 18 x 19 cm dimension.

	Running gel			Stacking gel
	Acrylamide			
Reagent	10%	12%	15%	4%
H₂O	15.2 mL	12.8 Ml	8.8 mL	2.98 mL
Acrylamide/Bis-acrylamide (30%/0.8% w/v)	13.6 mL	16 Ml	20 mL	0.67 mL
1.5 M Tris, pH 8.8	10.4 mL	10.4 Ml	10.4 mL	-
0.5 M Tris-HCl, pH 6.8	-	-	-	1.25 mL
10% (w/v) SDS	0.4 mL	0.4 mL	0.4 mL	0.05 mL
10% (w/v) APS	400 μL	400 μL	400 μL	50 μL
TEMED	40 μL	40 μL	40 μL	5 μL

2. Pipette the separating gel solution between the glasses, leaving about 1 cm below where the comb ends. Cover the gel with water or ethanol and wait about 20 to 30 min for the polymerization to occur.

3. Prepare the stacking gel solution, like described in Table 2.

4. Discard the water or ethanol that is in the top of the separating gel and pipette the stacking gel solution until all space is full.

5. Insert the comb, making sure that there are no bubbles under the teeth.
6. Let the gel polymerize.

3.2.5. Gel staining

Proteins separated by gel electrophoresis can be visualized using different staining procedures. The choice of staining technique depends on the desired sensitivity required to visualize the spots and, in many cases, the availability of imaging equipment in the lab. Here, BlueSafe staining will be used however the protocol of nitrate silver staining is described below.

#### 3.2.5.1. Silver nitrate staining

Silver stains offer the highest sensitivity, but with a low linear dynamic range. Often, these protocols are time-consuming and more complex than Coomassie blue staining and BlueSafe staining protocols.

1. Transfer the gel from the casting frames to a container with fixation solution and incubate at least two hours (it is possible to leave in fixation solution overnight).
2. Discard the fixation solution and wash the gel with washing solution for 20 min. Change the solution 3 times to be sure that all SDS detergent is removed.
3. Discard the washing solution and add enough sensibility solution.
4. Incubate for 2 min.
5. Discard the sensibility solution and wash the gel twice, 1 min each, with deionised water.
6. Add the silver nitrate cold solution and incubate under agitation for 20 min. Make sure that the gel is incubating without any light.
7. Substitute this solution for a large quantity of deionized water for 20 to 60 s to remove free ions. Repeat this step once again.
8. Add reduction solution for a short period of time.
9. Discard the solution.
10. Add new volume of the reduction solution for 2 to 5 min. The reaction should be stopped when intensity is desired.
11. Add stop solution. Agitate for 10 min.
12. Remove the stop solution and conserve the gel in water.

3.2.5.2. Staining with BlueSafe

BlueSafe is a protein stain that consists of a safer alternative to the traditional Coomassie Blue staining for detecting proteins in SDS-PAGE. BlueSafe is a highly sensitive single step protein stain and it is safer than Coomassie Blue staining because does not contain any methanol or acetic acid in its composition. Furthermore, it does not require the use of a destaining solution.

1. Transfer the gel from the casting frames to a container with deionised water to wash the gel.
2. Add BlueSafe Blue solution until the gel is submerse and incubate, under agitation, for 30 min at room temperature.
3. Remove the solution and conserve the gel in water.
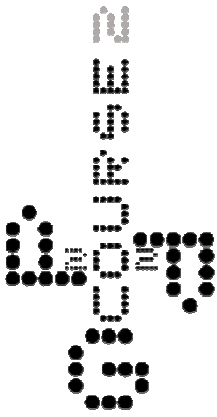
Ana Margarida Sousa and Sónia Silva

anamargaridasousa@deb.uminho.pt

PROTEOME ⮂ GENE ⮂ PROTEIN ⮂

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

Data derived from 2D gel electrophoresis is abundant and quite complex. In general, thousands of spots are detected and to find out differences in spot presence/absence, intensity (up or down regulated) or location (due to post-translational modifications) among gels is very tricky. Image capturing tools (CCD camera, laser scanner, and optical scanner) enable the transformation of 2D gel information into quantitative, computer-readable data that can be further analysed by image analysis software, which helps extracting biologically relevant information from the experiment. Through 2D gel analysis software it is possible to detect and quantify spots, compare gels, perform matching between gels and provide a complete statistical analysis. This protocol takes the main steps of 2D gel analysis using the SameSpots software.

2. INSTALATION OF THE SAMESPOTS TRIAL VERSION

You can freely explore the SameSpots software available at: http://totallab.com/samespots/tutorials/. Download the zip file "2D single stain tutorial". In this file, you will find the programme (.exe file), the tutorial images and the user guide.

3. IMAGE ACQUISITION

Before 2D gel image analysis, gels must be digitized into a computer-readable data. There are a variety of devices, including scanners, camera systems, densitometers, phosphor imagers, and fluorescence scanners. The most adequate device depends on the gel staining method performed and the output data format that must be compatible to the 2-D software.

4. 2D GELS ANALYSIS

The first time you use the 2D gel analysis software it is recommended you start with an analysis using the default settings. After the first interaction with the software and understanding what tools are available, and what kind of results are obtained, then the analysis can be improved by setting up advance tools.

4.1. Gel image analysis

The first step in the gel image analysis workflow is image optimization. To accurately compare gels, the images should be normalized in cases of differences in staining intensity or sample load variability. To avoid this step and to ensure results repeatability, it is recommended to use the same protocol along every 2D electrophoresis experiments and, more importantly, to use the same amount of protein.

To perform gel image analysis and spot matching, first it is necessary to choose the reference gel. If a gel image exists that contains all the protein spots of interest, then this gel should be chosen as the reference gel. Otherwise, let the software automatically select the reference gel from the images loaded.

Note 1

In some software, it is necessary to choose a gel named as master gel. The master gel is a virtual composite image that includes all the spots of interest in the experiment (similarly to the reference gel in SameSpots). If a gel image exists that that contains all the protein spots of interest, then this gel should be chosen as the master gel. Otherwise, let the software automatically select the master gel from the images loaded. In the last case, during further analysis steps, all the proteins of interest will be added to the master gel and thus, do not forget that is a virtual gel that the software constructs to compare gel images. The software combines all the protein of interest on a single reference image, the master gel.

4.2. Alignment

Before spot matching, Samespot software asks to select the interest zone of analysis. This feature is kind of interesting if you are looking for a particular protein or group of proteins speeding up the entire analysis.

The alignment phase can be started placing manual vectors between the same spots (usually well-known spots) to support the further automatic alignment. The manual vectors will thus

function as starting points of the automatic alignment. The SameSpots software recommends 5 alignment vectors distributed over the gel image to provide better results. After placing the manual vectors in each gel image run the "Automatic Alignment".

Tip 1 You can edit the Alignment Vectors later. The generated vectors by the automatic alignment can be added and deleted, and the vectors re-calculated.

Note 2
Some software follows a different approach than alignment vectors. To detect the protein spots in gel images and prepare to spot analysis, other software uses spot setting tools, such as "size", "intensity", and "minimal peak".

After completing the alignment, the software will display all the spots detected. If needed, spots can be removed based on their characteristics, such as position, area, normalised volume and combinations of these spot properties.

4.3. Spot detection
Once alignment has been performed, the next step is to verify the spots detected by the software and, if necessary, to improve detection by removing non-spots. You can remove spots based on position, area, normalised volume and combinations of these spot properties. According the parameters selected the software provides you the number of spots that verify these requisites and the option to delete them.

4.4. Normalisation
The remaining spots after the alignment and eventually further spots deletion will be used in the normalisation calculation. During the normalization step, the software compensates the images from variations that are not related with differences in protein sample, but rather reflect inconsistencies in the gel electrophoresis technique. The normalisation data cannot be altered but provides the data points used in the calculation of the normalisation factor for each image. For more information about calculation of the normalisation factor please consult the SameSpots tutorial.

4.5. Reviewing results
At this phase, all the spots are statistically listed based on their p value for the Anova analysis. You can review all spots and exclude a spot from being used in the further analysis.

Note 3

The selection of the spots can be based on: 1) the noted difference on spot intensities or protein quantities, such as increase or drop in a protein quantity as the result of treatment (generally, it is set as 2 fold-change); 2) statistical significance of the spot by using statistical validation; 3) combining set of spots for both quantitative and statistical selection.

Tip 2 You can tag the spots with different colours according the selection criteria. For example, green for those spots that had, at least, 2 fold-change and orange for those spots that were statistically significant. A same spot can be tagged with more than one colour if it respects both tags.

4.6. Statistics

The SameSpots presents several different statistic approaches, including the Principal Component analysis (PCA), correlation analysis, power analysis. If you are not interested in Picking or Calibrating the spots on your image for pI and Molecular Weight (Mw) you can skip this step and generate a report with the results obtained at this stage.

4.7. pI and Mw calibration

All the spots of your experiment can be calibrated according to pI and Mw by using the empirical information provided by Mw ladders used during the electrophoresis run or by using the pI and Mw from identified proteins.

As the image data is aligned, the pI and Mw markers are only required for one image, and then be applied to any of the images in the data set. After calibration, the pI and Mw of the spots are updated and appear in the table.

4.8. Spot picking

Spot picking can be performed by a robot or manually. For more information about spot picking using a robot, please consult the SameSpots tutorial.

4.9. Reporting

In the reporting section, it is possible to choose which information one wants to include in the final report. After selecting the type of information, create the report and, once finalized, the report will open in your default internet browser.

Tip 3 You can use the tags to control the spots that are included in the report.

Mass Spectrometry Approaches

Carla Silva

cmpsilva@ceb.uminho.pt

PROTEOME ⇄ GENE ⇄ PROTEIN ⇄

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

Gel electrophoresis is applied for separation of complex mixtures of proteins prior to mass spectrometric analysis. In-gel proteolytic digestion of separated proteins is performed to cleave the protein of interest present within the polyacrylamide matrix. After comprehensive analysis of gel images candidate proteins are selected depending on the purpose of the study. The spots of interest are excised for further processing. Coomassie Brilliant Blue is widely used for staining and must be removed prior to mass spectrometric analysis. After stain removal, the next steps include reduction and alkylation of the protein residues to denature the protein into its primary structure. Prior to MS identification, proteins are digested to generate peptides. Several proteolytic enzymes are available being the mostly used trypsin. After overnight incubation, peptides generated through proteolytic digestion are extracted. Afterwards the samples are subjected to mass-spectrometric analysis and the first step in MALDI-TOF analysis is the selection of appropriate matrix for the sample. The matrix selection mostly depends on the sample type, molecular weight of the target to be analysed. Different types of matrix are available in the market, with different properties and various applications. Selection of suitable matrix for a specific sample is very important, which can be narrowed down depending upon the properties and functions of the matrix.

2. MATERIALS AND REAGENTS

2.1. In-gel digestion of proteins

1. Ammonium bicarbonate (NH_4HCO_3).
2. Dithiothreitol (DTT).
3. Iodoacetoamide (IAA).
4. Trypsine (Proteomics/sequencing grade).
5. Acetonitrile (ACN).
6. Trifluro Acetic Acid (TFA).

7. Stain removal solution: 1:1 (v/v) ACN: 100 mM NH$_4$HCO$_3$. Dissolve 79.06 mg Ammonium bicarbonate (NH$_4$HCO$_3$) in 10 mL MilliQ water and add 1 mL of ACN to 1 mL of this solution.

8. Dehydration solution: 2:1 mixture of ACN and 50 mM ammonium bicarbonate buffer.

9. Reduction solution: 10 mM DTT in 100 mM ammonium bicarbonate buffer. Dissolve 7.7 mg DTT in 5 mL of 100 mM ammonium bicarbonate buffer.

10. Alkylation solution: 55 mM IAA in 100 mM ammonium bicarbonate buffer. Dissolve 10.16 mg IAA in 1 mL of 100 mM ammonium bicarbonate buffer.

11. Trypsin solution: Dissolve 20 µg lyophilized trypsin powder in 100 µL of 1 mM HCl solution. Mix properly and add 900 µL of 40 mM ammonium bicarbonate solution made in 9% ACN. Store at -20ºC freezer. The concentration of this stock solution is 20 µg/mL. Store this stock solution in different aliquots to avoid multiple freeze thaw cycles.

12. Extraction solution: 0.1% TFA in 50% ACN solution.

2.2. Sample preparation for the MALDI-TOF MS analysis

1. Acetonitrile (ACN).

2. Trifluoroacetic acid (TFA).

3. Spotting matrix (α-cyano-4-hydroxycinnamic acid – CHCA).

4. Protein sample.

5. Sample solvent: TA30 (30 : 70 [v/v] acetonitrile : 0.1% TFA in water)

6. Matrix solubilization: prepare a saturated solution of CHCA in TA30 solvent (30 : 70 [v/v] acetonitrile : 0.1% TFA in water)

2.4. Cleaning of MALDI targets

1. 2-propanol.

2. Deionized water.

3. Solvent TA30.

4. Ultrasonic bath.

5. Lint-free tissues.

3. METHODS

3.1. In-gel digestion of proteins

1. Rinse the entire gel with water for few hours with intermittent changing of water.

2. Excise protein spots/bands with a clean sterile scalpel and place the gel slice into a 1.5 mL microcentrifuge tube.

3. Cut the slices into small cubes (1 x 1 mm) while avoiding too small pieces as those can clog pipette tips.

4. Add 50 µL of stain removal solution (for large gel slices take enough liquid to cover it completely) and rotate on a shaker for 30 min at room temperature for removal of the stain from the gel pieces. Change the solution after every 10 min.

5. Add 50 µL dehydration solution and incubate at room temperature, until the gel pieces become white and stick together. Change the solution after every 10 min. Spin the gel pieces down (at ~2000 g for 30 s) and remove all liquid.

6. Add 30–50 µL of the reduction solution to completely cover the gel pieces. Incubate for 30 min at 56ºC. Chill down the tubes to room temperature; add 50 µL of dehydration solution, incubate for 10 min and remove the liquid.

7. Add 30-50 µL (more for a larger gel slices) of the alkylation solution and incubate for 20 min in dark at room temperature, add 50 µL of dehydration solution, incubate for 10 min and remove the liquid.

8. Add 25 µL trypsin solution (~500 ng) to the dry gel pieces and keep on ice for absorption of the enzyme by the gel pieces.

9. Add 25 µL ammonium bicarbonate buffer (in which trypsin is prepared) and incubate at 37ºC for overnight (12-16 h). Next morning, stop the reaction by keeping the reaction mixtures in ice.

10. Add 50 µL of extraction solution and vortex for 10 min at room temperature. Collect supernatant in a fresh microcentrifuge tube. Repeat the extraction step thrice and pool the collected supernatant.

11. Make small aliquots of the extracted solution and store at -20ºC freezer (if needed to be stored).

3.2. Sample preparation for the MALDI-TOF MS analysis

1. Mix 1 part of saturated CHCA solution with 1 part of sample solution.
2. Deposit 0.5 µL of the matrix/analyte mixture onto MALDI target and allow to dry.

3.3. Cleaning of MALDI targets

1. Wet a tissue with 2-propanol and wipe the sample/matrix spots from the surface of the MALDI target place.
2. Wet a tissue with water and wipe the upper surface of the MALDI target plate.

3. Place the MALDI target into a clean high-sided container and pour in enough 2-propanol to submerge the MALDI target plate. Place the container in the ultrasonic bath and sonicate for 10 min.

4. Place the MALDI target into a clean high-sided container and pour in enough solvent TA30 to submerge the MALDI target plate. Place the container in the ultrasonic bath and sonicate for 10 min.

5. Dry the MALDI target plate using a stream of high-purity nitrogen or compressed air.

Note 1
Do not wipe the upper surface of the cleaned target.

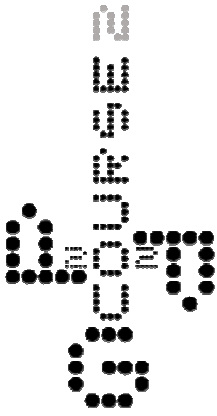### 3.4. MS data analysis
The experimental mass list obtained is further compared with a database of peptide mass values (SwissProt, NCBInr or other) using different search engines such as MASCOT or MS-Bridge available at:

http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF and

http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msbridgestandard,

respectively.

Sílvio Santos
silviosantos@ceb.uminho.pt

PROTEOME ⇄ GENE ⇄ PROTEIN ⇄

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

Bioinformatics is an interdisciplinary field that has become an important part of many areas of biology. It develops methods and software tools for understanding biological data, giving clues on the organisational principles within nucleic acid and protein sequences, and has been used for *in silico* analyses of biological queries using mathematical and statistical techniques. Bioinformatics uses homology to determine which parts of a protein are important in structure formation and interaction with other proteins, by searching unknown proteins against a database of known protein sequences to find if proteins with similar amino acid sequences have been described in other organisms. This information is used to predict the function of a protein once a homologous protein is known, allowing the classification of proteins on basis of their amino acid sequence as well as the identification of functional protein domains.

Moreover, bioinformatics aids in the design of cloning strategies for a given gene, in this particular case, in the design of the cloning strategy for the gene of interest to be heterologously expressed. Heterologous expression enables the isolation and obtainment of the protein of interest for further applications and functional analyses, being of central importance to overcome issues related to insufficient amount of protein in the original sample or interference of other proteins or components. The main structural component of the *C. thermocellum* cellulosome is CipA, a protein scaffold comprising nine type I cohesin modules, a type II dockerin module, and a family III carbohydrate binding module (CBM3) known for its strong affinity for cellulose [2]. In this practical session we will: 1) analyse the amino acid sequence obtained by mass spectrometry (corresponding to the CBM3 from CipA); 2) identify the gene sequence encoding this protein sequence; and 3) design and execute a strategy to clone this gene into an *E. coli* expression vector, so that we can subsequently produce and purify recombinant CBM3 for further use. CBM3 has wide-ranging biotechnological applications, such as improvement of paper properties, and functionalization of biomaterials, among others [7, 8]. Recombinant production is a simple way for obtaining the necessary amounts of pure protein for these and novel applications.

2. BIOINFORMATIC ANALYSIS

2.1. Identification of the protein and corresponding gene

The output of the MS analysis may be the complete protein or just a fragment. To identify the protein and corresponding gene a Basic Local Alignment Search Tool (BLAST) may be carried through the online server (https://blast.ncbi.nlm.nih.gov/Blast.cgi). BLAST is an algorithm for comparing primary biological sequence information (nucleotide or amino acid sequences), that enables to compare a query sequence with a library or database of sequences to find homologs.

We will clone the gene with its original nucleotide sequence although, in some cases, since the protein is usually expressed heterologously (in a different organism), a codon optimization may be needed. In this case the following bioinformatic tools can be used:

- http://eu.idtdna.com/CodonOpt;

- http://www.geneius.de/GENEius/Security_login.action.

2.2. Bioinformatic analysis of the protein

Bioinformatic analysis will enable the identification of homologous proteins as well as functional domains enabling to predict the protein function. Also, they will enable some basic protein characterization.

There is a huge number of bioinformatic tools that can be used. We will use the following online tools:

Output	Name	Web address
Homologous / family domains	Blastp	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Function prediction	IsterProScan	http://www.ebi.ac.uk/interpro/sequence-search
Protein families	Pfam	http://pfam.xfam.org
Protein homology	HMMER	http://www.ebi.ac.uk/Tools/hmmer
Transmembrane domains	TMHMM	http://www.cbs.dtu.dk/services/TMHMM
MW/PI	Compute pI/Mw tool	http://web.expasy.org/compute_pi
Signal peptide	SignalP	http://www.cbs.dtu.dk/services/SignalP
Solubility	SPpred	http://crdd.osdd.net:8081/sppred/submit.jsp

2.3. Selection and analysis of the expression plasmid for gene cloning

Depending on the aim, different expression plasmids can be used. In this work, the goal is to express the gene in *E. coli*, in an IPTG inducible vector, which enables high protein

expression (high copy-number plasmid), with a histidine tag which enables the protein purification. Among possible expression plasmids that present the desired characteristics, we have selected pET21a(+). The map and sequence of this vector can be obtained at: https://www.researchgate.net/file.PostFileLoader.html?id=5452712ed4c118b2108b45d2&assetKey=AS%3A272 165222453292%401441900654531 and https://www.addgene.org/vector-database/2549.

Considering the plasmid multicloning site (MCS), two restriction enzymes must be chosen to insert the gene in the correct orientation. Be aware that the restriction sites cannot be present in the gene sequence to avoid cutting the gene.

To identify the restriction map of the gene you can use the online tool NEBcutter (http://tools.neb.com/NEBcutter2).

2.4. Primer design

To amplify the gene of interest in order to insert it into the plasmid, the forward and reverse primers need to be designed. Some principles should be taken into consideration in primer design (adapted from [9]):

i.   primers should be 17-28 bases in length;

ii. base composition should be 50-60% (G+C);

iii. primers should end (3') in a G or C, or CG or GC: this prevents "breathing" of ends and increases efficiency of priming;

iv. melting temperatures (Tms) between 55-80ºC are preferred;

v.   3'-ends of primers should not be complementary (ie. base pair), as otherwise primer dimers will be synthesised preferentially to any other product;

vi. primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided;

vii. runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing) and should be avoided.

More information can be found at:

http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html.

The melting temperatures of the primers can be calculated through many online tools. An example is the Multiple Primer Analyzer at:

https://www.thermofisher.com/pt/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html.

The sequences of the chosen restriction enzymes should be added at the 5'-end of the corresponding primer and an extra sequence should be added to assure or increase the activity of the restriction enzymes. These extra sequences can be found at:

https://www.neb.com/~/media/NebUs/Files/Chart%20image/cleavage_olignucleotides_old.pdf.

You can verify the specificity of your primers through a nucleotide blast against the target sequence or template at http://blast.ncbi.nlm.nih.gov/Blast.cgi.

3. *IN SILICO* GENE CLONING

The gene cloning (PCR amplification, digestion and insertion into the plasmid) can be done *in silico* using, for instance, the software PlasmaDNA. The online version can be found at: https://www.plasmadna.net. The *in silico* gene cloning enables the researcher to confirm all the cloning steps and correct some common mistakes before performing the lab procedure.

PRACTICAL SESSIONS P5
Gene Cloning

Sílvio Santos and Ivone Martins

silviosantos@ceb.uminho.pt, ivone.martins@ceb.uminho.pt

PROTEOME ⇄ GENE ⇄ PROTEIN ⇄

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

Gene cloning is basically the simple act of making copies, or clones, of a single gene using the most diverse and available genetic engineering techniques. The obtained clones of a previously identified gene can be introduced into different organisms to i) give rise to new organisms or ii) alter the DNA sequence of an organism to change a given protein product. The obtained clones and/or genetically manipulated organisms can be used in many areas of biomedical and industrial research. The success of the cloning process depends on other independent but essential procedures, namely: the amplification of the DNA by polymerase chain reaction (PCR); the digestion of the insert DNA and vector DNA (plasmids are often used as vectors), using restriction enzymes; the visualization of the DNA by gel electrophoresis; the ligation of two or more DNA strands (insert + vector) to create recombinant DNA; the process of transferring the genetic material to a vector, into new host cells – transformation - and the subsequent screening method to select the transgenic organisms [9].

2. GENE CLONING

In this practical class we will clone the previously identified gene in P4, into an expression plasmid and introduce it into an *E. coli* strain for further protein production. First, we are going to amplify the DNA by PCR using the primers designed in P4 and the PCR product will be visualized in an agarose gel. Then, both the amplified DNA and the plasmid DNA will be digested with the same restriction enzymes, purified and ligated to obtain a circular plasmid containing the foreign DNA. The ligation product (recombinant plasmid) will be introduced in *E. coli* by electroporation and positive transformants will be screened by colony PCR.

2.1. MATERIALS AND REAGENTS

2.1.1. Gene amplification

1. 10 mM dNTPs (NZYTech).
2. Nuclease-free water.

3. Phusion Green High-Fidelity DNA Polymerase (Thermo Scientific).

4. 5x Phusion Green High-Fidelity Buffer (Thermo Scientific).

5. Primers forward and reverse.

6. Template DNA (10 ng).

7. GreenSafe Premium loading dye (NZYTech) or Safe-Green™ (BioPortugal).

8. 6x NZYDNA loading dye (NZYTech).

9. Agarose (BioPortugal).

10. DNA marker Nzyladder I (NZYTech).

11. 1x TAE (Tris-acetate-EDTA): 242 g Tris-base, 57.1 mL 100% acetic acid, 10 mL 0.5 M EDTA, water up to 1L.

12. Shott flasks.

13. 0.2 mL tubes.

14. Micropipettes and corresponding sterile tips.

15. Centrifuge.

16. Thermal cycler.

17. Microwave.

18. Electrophoresis equipment (casting tray, gel comb, gel tank and power supply).

19. Transiluminator.

2.1.2. Digestion with the restriction enzymes

1. Nuclease-free water.

2. 10x Xpress Buffer (Grisp).

3. Xpress *Nde*I (Grisp).

4. Xpress *Xho*I (Grisp).

5. FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific).

6. 1.5 mL tubes.

7. Micropipettes and corresponding sterile tips.

8. Centrifuge.

9. Heat block.

2.1.2. DNA purification and Ligation

1. Nuclease-free water.

2. 10x T4 DNA Ligase Buffer (Thermo Scientific).

3. T4 DNA Ligase (Thermo Scientific).

4. 1.5 mL tubes.

5. GRS PCR and Gel Band purification kit (Grisp).

6. Micropipettes and corresponding sterile tips.

7. Centrifuge.

8. Heat block.

2.1.3. Transformation (Electroporation)

1. *E. coli* electrocompetent cells (JM109).

2. SOC (Super Optimal Broth with Catabolite Repression) medium (31,54 g/L, NZYTech).

3. Luria-Bertani (LB) medium (25 g/L, NZYTech) sterilized by autoclaving plus 20 mg/mL ampicillin (NZYTech).

4. LB culture plates (LB medium with 20 g/L of agar, NZYTech) plus 20 mg/mL ampicillin (NZYTech).

5. 2 mm Fisherbrand™ Electroporation cuvettes Plus™ (ThermoScientific).

6. 1.5 mL tubes.

7. Micropipettes and corresponding sterile tips.

8. Centrifuge.

9. Eletroporator (Eppendorf).

10. Ice.

2.1.4. Screening of transformants

1. Nuclease-free water.

2. NZYTaq II 2x Green Master Mix (NZYTech).

3. T7 promotor primer forward.

4. T7 terminator primer reverse.

5. LB plus ampicillin culture plates.

6. GreenSafe Premium loading dye (NZYTech) or Safe-Green™ (BioPortugal).

7. Agarose (BioPortugal).

8. 1x TAE.

9. 6x NZYDNA loading dye (NZYTech).

10. DNA marker Nzyladder I (NZYTech).

11. 1.5 mL tubes.

12. Micropipettes and corresponding sterile tips.

13. Centrifuge.

14. Electrophoresis equipment (casting tray, gel comb, gel tank and power supply).

15. Transiluminator.

3.2. METHODS

3.2.1. Gene amplification

Amplification of the target gene should be done with a proof reading polymerase to avoid the insertion of erroneous nucleotides. In this case we will use the Phusion Green High-Fidelity DNA Polymerase according to the manufacturer's instructions.

The reaction should be carried as follows:

Component	50 μL reaction	Final concentration
H_2O	Up to 50 μL	
5X Phusion Green HF Buffer	10 μL	1x
10 mM dNTPs Mix	1 μL	200 μM each
Primer forward	2.5	0.5 μM
Primer reverse	2.5	0.5 μM
Template DNA	2 μL	20 ng
Phusion DNA Polymerase	0.5 μL	0.02 U/μL

With the following cycling instructions:

Cycle step	3-step protocol		Cycles
	Temperature	Time	
Initial denaturation	98ºC	30 s	1
Denaturation	98ºC	10 s	
Annealing	53-63ºC (Gradient)	30 s	34
Extension	72ºC	30 s	
Final Extension	72ºC	10 min	1
	4ºC	Hold	

Amplification should be confirmed in an electrophoresis agarose gel (0.8%) with an appropriate molecular weight marker.

Prepare the agarose gel as follow:
1. Weight 0.40 g of agarose.
2. Add 50 ml TAE 1x.
3. Mix and melt in the microwave.

4. Cool down at approximately 50ºC and add 2.5 µL of GreenSafe Premium or Safe-Green™ to stain the nucleic acids.

5. Pour the agarose into a gel tray with the well comb in place and let it solidify.

To load the samples and run the gel proceed as follows:

1. Add 1 µL of 6x NZYDNA loading dye to 5 µL of sample.

2. Once solidified place the gel into the gel tank, remove the comb and fill the tank with 1x TAE until the gel is covered.

3. Carefully load 5 µL of the NZYDNA Ladder I DNA marker into the first lane of the gel and all the volume of the sample (prepared in 1) in the additional wells.

4. Run the gel at 80-150 V until the dye line is approximately 75-80% of the bottom of the gel (a typical run time is about 1-1.5 h, depending on the gel concentration and voltage).

5. Visualize the gel under UV light in a transiluminator.

3.2.2. Digestion with the restriction enzymes

The gene and the plasmid must be digested with the selected restriction enzymes *Nde*I and *Xho*I (double digestion). We will use the Xpress version of these enzymes, according to the manufacturer's instructions:

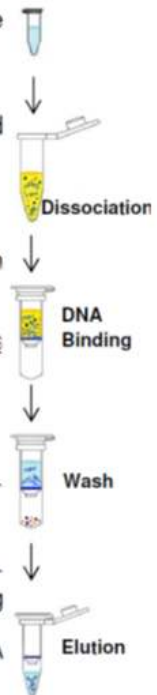|	Volume	
Component	Plasmid DNA	Unpurified PCR product
H$_2$O nuclease-free	Up to 20 µL	Up to 30 µL
10x Xpress Buffer	2 µL	2 µL
DNA	5 µL	10 µL
Xpress *Nde*I	1 µL	1 µL
Xpress *Xho*I	1 µL	1 µL

1. Add 1 µL of FastAP Alkaline Phosphatase to the plasmid (to catalyse the removal of 5'- and 3'-phosphate groups avoiding self-ligation of the vector).

2. Mix gently and spin down.

3. Incubate at 37°C in a heat block or water thermostat for 5-15 min.

4. Inactivate the enzymes for 5 min at 80ºC.

3.2.3. DNA purification and ligation

1. Purify the plasmid and gene DNA with the GRS PCR and Gel Band purification kit following the manufacturer´s instructions:

PROTOCOL FOR PCR CLEAN UP

1) Transfer up to **100µl** of the **PCR reaction solution** to a 1,5-ml microcentrifuge tube. Add 5 volumes of **Gel Solubilization Solution** and mix by vortexing. (e.g. 25µl PCR reaction volume plus 125µl gel solubilization solution).

2) **[optional]** If the mixture has turned purple, add 10 µl of 3M Sodium Acetate pH 5.0 to adjust pH and mix thoroughly.

3) Place the **DNA fragment mini spin column** in a 2-ml collection tube, and transfer the sample mixture to the column.

4) Centrifuge at 14.000g-16.000g for 30 seconds.

5) Discard the collection tube containing the flow-through, and place the spin column back in the collection tube.

6) Add 600µl of **Wash Buffer 2*** and let stand for 1 minute. (check if ethanol is added first time the kit is used; see notes above).

7) Centrifuge at 14.000g-16.000g for 30 seconds and discard the flow-through

8) Place the spin column back in the collection tube and centrifuge at 14.000g-16.000g for another 3 minutes to dry the matrix of the column.

9) Transfer the spin column to a new 1,5ml microcentrifuge tube and pipet 20µl-50µl **Elution Buffer** directly to the centre of the spin column without touching the membrane. Incubate at room temperature for 2 minutes.
Notes: Yield could be increased using pre-warmed Elution Buffer (60ºC). Instead of Elution Buffer, DNA can also be eluted with TE or water; pH should be 8.0-8.5.

10) Centrifuge for 2 minutes at 14.000g-16.000g to elute purified DNA. Discard the spin column and use DNA immediately or store at -20ºC.

2. To insert the gene into the plasmid vector a ligase is needed, commonly a T4 DNA ligase is used. We will use the T4 DNA Ligase from Thermo Scientific, according to the manufacturer's instructions:

Component	Volume
Plasmid	2 µL
Insert DNA	1 µL (above 3:1 molar ratio over plasmid)
10x T4 DNA Ligase Buffer	2 µL
T4 DNA Ligase	0.5 µL
Water	Up to 20 µL

3. Incubate 10 min at 22°C.

3.2.4. Transformation (Electroporation)

E. coli electrocompetent cells (JM109) will be used to eletroporate as follows:

1. Thaw on ice an aliquot of 20 µL of eletrocompetent cells.

2. Add 3 µL of the ligation product and mix.

3. Incubate for 1 min on ice and then transfer to a pre-cooled 2 mm cuvette.

4. Place the cuvette in the chamber slide of the eletroporator and pulse with 2.5 kV.

5. Remove the cuvette and add to the cells 1 mL of SOC medium.

6. Transfer the mixture to an eppendorf tube and incubate at 37ºC for 1 h to allow the cells to recover.

7. Inoculate a LB plus ampicillin culture plate with 100-200 µL of the eletroporated cells and incubate at 37ºC overnight.

3.2.5. Screening of transformants

Transformation will likely result in a multitude of colonies with an associated percentage of correct transformants. This percentage will reflect the success of the whole process. The presence of incorrect transformants or clones requires the use of a screening method. Colony PCR is an easy way to rapidly screen ligation reactions for positive clones, saving both time and money.

1. Prepare five eppendorf tubes with 25 µL of LB plus ampicillin.

2. Selected and pick with, a pipette tip, five colonies and resuspended in the prepared tubes.

3. Incubate at 37ºC, 1 h.

4. Prepare the PCR reaction as follows using the NZYTaq II 2x Green Master Mix:

Component	Volume
NZYTaq II 2x Green Master Mix	25 µL
T7 promoter primer Forward*	1 µL
T7 terminator primer Reverse*	1 µL
Template DNA (resuspended colony)	2 µL
Water	Up to 50 µL

* These commercially available primers flank the inserted gene, annealing in the original plasmid. This way we have a built-in control for the PCR reaction, because even negatives (vector religation) will yield a ~250 bp amplification product.

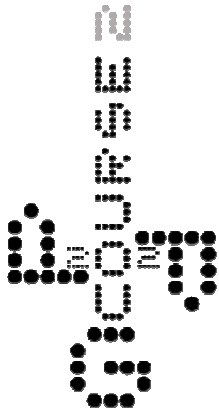| Cycle step	3-step protocol		Cycles
	Temperature	Time	
Initial denaturation	95ºC	10 min	1
Denaturation	95ºC	45 s	
Annealing	49ºC	45 s	34
Extension	72ºC	30 s	
Final Extension	72ºC	10 min	1
	4ºC	Hold	

5. Run the PCR product on an agarose gel (0.8%) to analyse product size (*see* **section 3.2.1**).

6. Use the remaining portion of two positive colonies to inoculate a LB plus ampicillin culture plate for downstream applications.

Note 1

Plasmid DNA is extracted from one of these colonies after overnight growth and purified to be sequenced across the entire insert, in order to verify the exact sequence of the insert (use a sequencing service, *e.g.*, GATC). Plasmid purification can be carried through a miniprep kit according to the manufacturer's instructions (*e.g.*, NucleoSpin® Plasmid from MACHEREY-NAGEL).

Note 2

After confirming the correct insertion of the gene of interest, the purified plasmid (see *Note 1*), will be used to transform the *E. coli* strain BL21 (DE3) for heterologous expression.

PRACTICAL SESSION **P6**

Recombinant Protein Production and Purification

Carla Oliveira and Tatiana Q. Aguiar

carlaoliveira@deb.uminho.pt

PROTEOME ⊇ GENE ⊇ PROTEIN ⊇

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. INTRODUCTION

The bacterial strain *Escherichia coli* BL21 (DE3)/pET21a+CBM3 constructed in the practical section P3 will be used to produce and purify recombinant CBM3 in practical sections P4 and P5, respectively. The *cbm*3 gene was cloned in the pET21a vector in frame with a tail of six histidines (histag) at its C-terminal to allow its purification by IMAC (immobilized metal ion affinity chromatography). The histag is one of the most widely used tags in recombinant protein purification (for recent reviews see, for example, [10-12]). The recombinant CBM3 has 208 amino acids and a predicted molecular weight (MW) of 22.7 kDa. 1D denaturing polyacrylamide gel electrophoresis (SDS-PAGE) [9] will be used for monitoring protein production and purification and to assess potential purification levels.

2. MATERIALS AND REAGENTS

2.1. Production of recombinant CBM3 from *E. coli* BL21 (DE3)

1. Luria-Bertani medium (LB): sterilized by autoclaving.
2. 100 mg/mL Ampicillin: filter-sterilized.
3. 1 M IPTG (isopropyl-β-D-thiogalactopyranoside): filter-sterilized.
4. 50 and 100 mL Erlenmeyer flasks: sterilized by autoclaving with cotton plug.
5. 1.5 and 50 mL centrifuge tubes.
6. Micropipettes and corresponding sterile tips.
7. Incubator with orbital shaking.
8. Spectrophotometer.
9. Refrigerated centrifuge.
10. Ice.

2.2. Purification of recombinant CBM3 by IMAC

All buffers should be filtered with 0.45 µm filters and maintained at 4ºC until use.

1. HisPur™ Ni-NTA Resin (Thermo Scientific).

2. NZY Bacterial Cell Lysis Buffer (NZYTech) containing 0.1 mg/mL lysozyme and 4 μg/mL DNase I.

3. 0.2 M PMSF (phenylmethylsulfonylfluoride).

4. 8x Phosphate buffer, pH 7.4: 160 mM sodium phosphate, 4 mM NaCl.

5. Equilibration and wash buffer: 1x phosphate buffer, 40 mM imidazole, pH 7.4.

6. Elution buffer: 1x phosphate buffer, 300 mM imidazole, pH 7.4.

7. 1.5, 2 and 15 mL centrifuge tubes.

8. Microcentrifuge.

9. Tube rotator.

10. Dry block heater.

11. UV spectrophotometer (e.g., Nanodrop).

2.3. SDS-PAGE analysis

1. 15% SDS-PAGE gel; per 10 mL: 3.6 mL 40% acrylamide, 2 mL 2% *N,N*-methylene-bis-acrylamide, 2.5 mL 0.5 M Tris-HCl pH 6.8 (stacking gel buffer) or 1.5 M Tris-HCl pH 8.8 (resolving gel buffer), 100 μL 10% (w/v) SDS (sodium dodecyl sulfate), 50 μL 10% (w/v) APS (ammonium persulfate), 5 μL TEMED (tetramethylenediamine).

2. 1x and 5x SDS-PAGE sample buffer (NZYTech).

3. PageRuler™ Unstained Broad Range Protein Ladder (Thermo Scientific).

4. 10x Running Buffer; for 250 mL: 7.575 g Tris base, 36 g glycine, 2.5 g SDS.

5. Mini-Protean III apparatus (Bio-Rad).

6. BlueSafe (NZYTech).

7. Containers.

3. PRODUCTION OF RECOMBINANT CBM3 FROM *E. COLI* BL21 (DE3)

3.1. Preparation of the pre-culture

1. Add 5 mL LB medium plus 5 μL ampicillin stock solution to a 50 mL culture flask.

2. Innoculate the medium with the biomass from a fresh plate, using a tooth pick.

3. Shake the culture at 150 rpm, 37ºC, overnight.

3.2. Culture inoculation and protein induction

1. Prepare 20 mL LB medium with 20 μL ampicillin stock solution in a 100 mL Erlenmeyer flask.

2. Add 0.20 mL of the pre-culture to the flask (*see* **Note 1**).

3. Shake the culture at 200 rpm and 37ºC until reaching an OD600 = 0.5-0.6.

4. Take a 100 µL aliquot before induction (**BI**), centrifuge at the maximal speed for 10 min in a table-top centrifuge, discard the supernatant and resuspend the cell pellet in 50 µL of 1x SDS-PAGE sample buffer. Boil it for 10 min and freeze until use.

5. Add 20 µL IPTG stock solution (final concentration 1 mM) to the culture and shake at 150 rpm and 37ºC for 3-4 h.

6. Take a 100 µL aliquot after induction (**AI**) and proceed as in step 4.

7. Centrifuge the culture in a 50 mL falcon tube at 10000 rpm for 15 min at 4ºC, discard the supernatant and place the cell pellet on ice (*see* **Note 2**).

4. PURIFICATION OF RECOMBINANT CBM3 BY IMAC

4.1. Cell lysis

1. Add 4 mL of lysis buffer containing 20 µL PMSF stock solution (*see* **Note 3**) to the total cell pellet and resuspend it gently by pipetting.

2. Transfer the mixture into a 15 mL falcon tube and incubate for 20 min at room temperature, with gentle shaking.

3. Centrifuge at 10000 rpm for 15 min at 4ºC.

4. Carefully remove the supernatant (called cell-free extract or lysate) (*see* **Note 4**).

5. Keep the lysate on ice until its use (***soluble fraction* sample - SF**).

4.2. IMAC spin-protocol at eppendorf-scale

1. Gently shake the bottle in which the HisPur™ Ni-NTA Resin is supplied (stored at 4ºC) until the medium is homogeneous.

2. Remove 0.5 mL of slurry from the bottle and transfer it into a 2 mL eppendorf tube (*see* **Note 5**).

3. Sediment the resin by centrifugation for 2 min at the maximal speed in a table-top centrifuge and carefully remove and discard the supernatant using a micropipette.

4. Add 1 mL of equilibration buffer to the resin, mix by inverting the tube until the resin is fully suspended (~3-4 times), and proceed as in step 3.

5. Incubate 1.5 mL lysate with the resin for 30 min in a tube rotator.

6. Re-sediment the resin as in step 3 and save the supernatant for analysis (***after resin* sample - AR**).

7. Add 1 mL of wash buffer to the resin, mix and proceed as in step 6 (*1ˢᵗ wash* **sample** - **W1**).

8. Repeat resin washing (step 7) five times (*2ⁿᵈ to 5ᵗʰ wash* **samples** – **W2 to W5**).

9. Add 0.5 mL of elution buffer to the resin, mix and proceed as in step 6 (*1ˢᵗ elution sample* - **E1**).

10. Repeat protein elution (step 9) two more times (*see* **Note 6**), saving each supernatant fraction in a separate tube (*2ⁿᵈ and 3ʳᵈ elution* **samples** - **E2 and E3**).

11. Save the resin for further regeneration (*see* **Note 7**).

12. Mix 16 μL aliquots from samples **SF**, **AR**, **W1**, **W5**, **E1**, **E2** and **E3** with 4 μL of 5x SDS-PAGE sample buffer in 1.5 mL eppendorf tubes. Boil them for 5 min, keep on ice, and proceed for SDS-PAGE analysis (**section 5**).

13. Estimate protein concentration in each eluted sample by measuring the absorbance at 280 nm (*see* **Note 8**).

14. Calculate purification yield in mg of protein per mL of culture.

5. SDS-PAGE ANALYSIS

5.1. Gel running

1. Load 15 μL of each denatured sample from the key steps of recombinant production and purification in the SDS-PAGE gel.

2. Load 5 μL of molecular weight marker per gel.

3. Run the gel in 500 mL of fresh 1x Running Buffer at 15 mA (for only 1 gel) or 30 mA (for two gels) until the running front (indicated by the blue line) is positioned at ~0.5 cm from the end of the gel.

5.2. Gel staining

1. After running the gel, remove it from the frames and place it in a container with distilled water.

2. Pour out the distilled water and add 25 mL of BlueSafe *per* gel.

3. Incubate with gentle agitation and wait for the protein bands to appear (10-15 min to see 60 ng, and longer for lower concentration proteins).

4. Remove the gel from the staining solution and place it in distilled water.

5. Photograph and analyze the gel (*see* **Note 9**).

6. NOTES

1. The dilution of the pre-inoculum in the culture is usually of 100 times.

2. Frozen pellets are easier to lysate. Pellets can be stored at -20ºC for several weeks for long-term processing.

3. PMSF (which inhibits serine proteases) is the protease inhibitor most commonly used. Other protease inhibitors can also be used; however, they should be compatible with IMAC purification. For example, EDTA (used as metalloprotease inhibitor) is incompatible with IMAC, as the nickel ions in the affinity matrix are chelated by EDTA. Therefore, EDTA is used for stripping the nickel ions from the resin when necessary.

4. The lysate must be filtered through a 0.45 μm filter if it is to be directly applied into a pre-packed purification column, to prevent column clogging.

5. The binding capacity of this product is 60 mg per milliliter of settled resin (for a 28 kDa 6xHis-tagged protein). The amount of resin should be adjusted to each recombinant protein expression level and scaled as needed. The protocol herein described was optimized for the purification of recombinant CBM3, and therefore, it does not have the exact guidelines of the resin manufacture.

6. For downstream applications, imidazole must be removed from the eluted protein samples, using, for instance, a dialysis or a desalting method.

7. The resin may be reused (about 5 times) for the purification of the same protein without affecting protein yield or purity, with no need for stripping and recharging of the nickel ions. Follow the manufacturer's instructions for resin regeneration.

8. The spectrophotometric determination of protein concentration is based on the absorbance and molar extinction coefficient of the protein at A280 (*see* Beer-Lambert Law). This coefficient can be calculated from the primary protein sequence, using, for instance, the ProtParam on-line tool (http://web.expasy.org/protparam/) [10]. For the recombinant CBM3 produced as described here use the following relation, as given by this tool: A280 ~ 1.6 is equivalent to 1 g/L of pure CBM3.

9. Silver nitrate is more sensitive than Coomassie blue staining; allowing the detection of protein amounts down to 2 ng, and thus, is more adequate to evaluate protein purity [9].
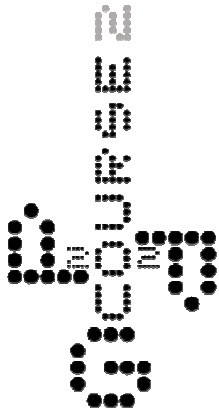
PROTEOME ⇄ GENE ⇄ PROTEIN ⇄
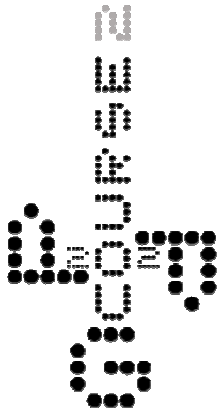
Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

1. Graves PR, Haystead TA (2002) Molecular biologist's guide to proteomics. Microbiology and Molecular Biology Reviews 66(1):39-63.

2. Xu Q et al. (2016) Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. Science Advances 2:e1501254.

3. Rydzak T et al. (2012) Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression. BMC Microbiology 12:214.

4. User Guide: Pierce BCA Protein Assay Kit.

5. Rabilland et al. (2011) Two-dimensional gel electrophoresis in proteomics: a tutorial. Journal of Proteomics 74:1829-41.

6. ReadyStrip IPG strip instructions Manual Cat163-2099.

7. Oliveira C, Carvalho V, Domingues L, Gama FM (2015) Recombinant CBM-fusion technology – applications overview. Biotechnology Advances 33:358-69.

8. Oliveira C, Sepúlveda G, Aguiar TQ, Gama FM, Domingues L (2015) Modification of paper properties using carbohydrate-binding module 3 from the *Clostridium thermocellum* CipA scaffolding protein produced in *Pichia pastoris*: elucidation of the glycosylation effect. Cellulose 22:2755-65.

9. Oliveira C, Aguiar TQ, Domingues L (2017) Principles of Genetic Engineering. *In* A Pandey & JA Teixeira (Eds.), Current Developments in Biotechnology and Bioengineering: Foundations of Biotechnology and Bioengineering. Oxford: Elsevier Ltd. pp. 81-127. ISBN 978-0-444-63668-3.

10. Oliveira C, Domingues L (2018) Guidelines to reach high-quality purified recombinant proteins. Applied Microbiology & Biotechnology 102:81-92.

11. Costa SJ, Almeida A, Castro A, Domingues L (2014) Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli*: the novel Fh8 system. Frontiers in Microbiology 5: 63.

12. Yadav DK, Yadav N, Yadav S, Haque S, Tuteja N (2016) An insight into fusion technology aiding efficient recombinant protein production for functional proteomics. Archives of Biochemistry and Biophysics 612: 57-7.

LIST OF PARTICIPANTS

PROTEOME ∶ GENE ∶ PROTEIN ∶

Hands-on course | 26-29 Mar 2018 | Braga-PORTUGAL

Surname	Name	Institution	E-Mail
Araújo	Daniela	CEB-UMinho	daniela.araujo@ceb.uminho.pt
Augusto	Diana	UTAD	dianah@utad.pt
Azevedo	Nuno	UMinho	nunoazevedo015@gmail.com
Brandão	Ana	CEB-UMinho	anacatarinabr@hotmail.com
Costa	Lara	CEB-UMinho	laracosta@ceb.uminho.pt
Dias	Patrícia	CEB-UMinho	pattydias30@gmail.com
Machado	Manuela	CEB-UMinho	mmmachado@net.sapo.pt
Oliveira	Catarina	FCT/UNL	css.oliveira@fct.unl.pt
Silva	Elda	UMinho	essabino@hotmail.com
Souza	José António	IST/UL	josesouza@tecnico.ulisboa.pt