UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Omics approaches to assess the effect of agro-food nutritive extracts for pcDNA-FLAG-p53 biosynthesis

Andreia Vanessa Guedes Amaral

Dissertação para obtenção do Grau de Mestre em
Biotecnologia
(2º ciclo de estudos)

Orientador: Prof. Doutor Luís António Paulino Passarinha
Co-orientador: Mestre Fátima Raquel Milhano dos Santos

Covilhã, outubro de 2016

Agradecimentos

Em primeiro lugar quero agradecer ao meu orientador Professor Luís Passarinha, pela sua orientação, recomendações e pelo apoio que disponibilizou ao longo da realização deste trabalho. Gostaria de agradecer, em particular à minha co-orientadora Mestre Fátima Santos, por todo o seu apoio na área da proteómica, pelo companheirismo que demonstrou desde que começou a acompanhar este trabalho, e por toda a ajuda que me disponibilizou.

Quero agradecer à Mestre Mafalda Matos pela ajuda, dicas e pelo tempo que me disponibilizou na primeira parte deste trabalho. Agradeço ao professor Guilhermo Gosset (Universidad Nacional Autónoma de México) e ao Professor José Teixeira (Universidade do Minho), por terem disponibilizado os recursos necessários a realização deste trabalho.

Além disso, também gostaria de expressar minha gratidão a todas as pessoas envolvidas no Centro de Investigação em Ciências da Saúde da Universidade da Beira Interior particularmente ao grupo de Biotecnologia e Ciências Biomoleculares por todo apoio e ajuda em especial à mestre Margarida Grilo e Leonor Gaspar. Um obrigado à Dr.ª Ana Martinho pelo convite e por ter possibilitado a participação no concurso CAP – Cultiva o teu futuro: Inovação no setor de leite e lacticínios.

Agradeço aos meus amigos por todos os bons momentos que vivemos nestes cinco anos na Covilhã, especialmente á Regina, ao César, Joana, Ivo, Carla, Quicas obrigada pelo companheirismo e apoio que me deram especialmente no último ano, obrigada pelos nossos cafés, pelas nossas horas de "corte e custura". Às minhas queridas "*miguxas*" Cátia, Vera e Ana Isabel um obrigada especial pelo apoio e suporte que me deram nos momentos menos fáceis, por terem escutado sempre os meus problemas e por terem suportado o meu mau feitio nos últimos anos.

Finalmente, agradeço às pessoas mais especiais da minha vida, a minha família, em especial aos meus pais por todo amor, carinho, paciência e por todas as dificuldades que ultrapassaram ao longo dos anos para me ajudarem sempre, amo-vos muito.

Resumo

Nos últimos anos, o DNA plasmídico tem sido usado como um vetor em terapia génica e em vacinas de DNA, e por essa razão, a capacidade de produzir grandes quantidades de plasmídeo é importante no que diz respeito ao processo de produção de vacinas de DNA, em larga escala. Assim, o objetivo deste trabalho foi o estudo metabólico e proteómico da *Escherichia coli* (estirpe VH35) durante a síntese de pcDNA-FLAG-p53, usando como meio de fermentação compostos alternativos provenientes da indústria agro–alimentar, como o soro de queijo, o "corn steep licor" e o extrato de levedura.

Combinando estes compostos alternativos como meio de crescimento, e usando a estirpe VH35 como hospedeiro, procedeu-se então à otimização da produção de DNA plasmídeo, em termos de rendimento e qualidade. Após essa otimização, analisou-se o consumo de açúcares no meio de fermentação ao longo do tempo, e as alterações proteómicas que ocorrem devido a adaptações metabólicas.

O trabalho desencolvido sugere que a estirpe VH35 utiliza meios agro-alimentares como fonte de energia, consumindo assim a lactose presente no meio, uma vez que estes meios não possuem açúcares PTS (fosfoenolpiruvato). Devido ao facto desta alteração poder levar a modificações no proteoma da célula, foi feita uma análise proteómica por eletroforese bidimensional, que revelou variações no proteoma da célula quando comparado com diferentes meios de crescimento. Por fim, essas variações proteómicas foram analisadas por MALDI-TOF/TOF, e foi possível identificar proteínas diferencialmente expressas, como por exemplo a proteína MreB, a desoxirribose fosfato – aldolase, chaperonina ClpB,as quais se encontram principalmente relacionadas com o metabolismo e a síntese de nucleótidos.

Palavras-chave

Proteómica, Eletroforese Bidimensional, DNA plasmídico, Soro de Queijo, "Corn steep licor", *Escherichia coli* VH35.

Resumo Alargado

Nos últimos anos, tem-se verificado um aumento das aplicações terapêuticas do DNA plasmídico (pDNA) no tratamento de inúmeras doenças como o cancro, infeções virais, hepatite, doenças cardiovasculares, entre outras. Devido à sua crescente ocorrência e complexidade, o cancro tornou-se o maior alvo da terapia génica e das vacinas de DNA, tendo surgido, nos últimos anos, novas abordagens de resposta terapêutica. Muitas destas novas metedologias baseiam-se na proteína p53, devido ao seu papel regulador em inúmeros processos celulares, como por exemplo a apoptose das células, diferenciação celular e mecanismos de reparação.

Deste modo, tornou-se extremamente relevante a obtenção de elevadas quantidades de DNA plasmídico, na sua forma superenrolada e com elevado grau de pureza, para que possa ser usado terapeuticamente segundo as normas da Food and Drug Administration (FDA). Assim, o objetivo deste trabalho centra-se na análise das várias etapas de produção de pDNA. A otimização deste processo engloba várias etapas, incluindo a construção do vetor, escolha do hospedeiro e das condições de crescimento, sendo necessária a seleção da composição ideal do meio de crescimento.

O meio de crescimento possui um enorme impacto na produção de alto rendimento de pDNA, e por essa razão, deve haver um compromisso entre o custo e a eficácia do processo. Deste modo, os meios agro-alimentares surgem como alternativa às fontes de crescimento comuns, constituindo na sua maioria subprodutos da indústria alimentar como é o caso do soro de queijo e do "corn steep licor". Estes já foram descritos como meio de crescimento na produção de inúmeras substâncias como é o caso da produção de etanol.

Assim, o trabalho nesta dissertação visa utilizar estes dois substratos na formulação do meio de crescimento para a produção de pDNA, usando como hospedeiro *Escherichia coli* VH35. Esta estirpe encontra-se descrita na literatura como sendo de baixa produção e acumulação de acetato, permitindo assim um aumento do tempo de crescimento e uma diminuição dos danos na estrutura e função celular. Após uma primeira etapa de otimização de produção, foi utilizado o HPLC (cromatografia liquida de alta eficiência) para a análise dos açúcares e monitorização do seu consumo ao longo do tempo, bem como da produção de acetato. Seguidamente, realizou-se um estudo proteómico através da eletroforese bidimensional, seguida de identificação de proteínas diferencialmente expressas por MALDI-TOF/TOF.

Nos últimos anos, inúmeros estudos relativos ao proteoma da *E.coli* têm demonstrado possíveis interações que ocorrem entre as alterações ao nível do seu proteoma e as modificações metabólicas causadas pela adaptação ao meio. Assim, utilizando a eletroforese bidimensional, foi possível o estudo proteómico de células de *E.coli* VH35 produzidas em diferentes meios de

crescimento com composições distintas A estratégia teve como intuito a análise das alterações ocorridas devido à utilização de substratos agro–alimentares, em vez de substratos comerciais, e ainda a análise de possíveis alterações nos meios não convencionais que possam ocorrer pela adição do extrato de levedura ou de aminoácidos aromáticos.

Depois da otimização das condições de eletroforese bidimensional para o estudo do proteoma da *E.coli* VH35, foi possível proceder-se à análise proteómica dos extratos produzidos nos diferentes meios de crescimento. Esta análise foi fundamental para a identificação das proteínas envolvidas quer no metabolismo quer na síntese de pDNA.

Assim, o objetivo inicial do trabalho centralizou-se na obtenção de elevadas quantidades de pDNA usando como meio de crescimento subprodutos agro-alimentares, sendo possível concluir que a adição suplementar de 5g/L de extrato de levedura ao meio de crescimento, foi benéfica na produção de pDNA. Deste modo, de forma a aumentar a sua produção, optou-se por adicionar aminoácidos aromáticos ao meio. Esta abordagem permitiu aumentar a produtividade de pDNA, assim como a pureza deste, atingindo-se uma produção superior a 40 mg/L de pDNA.

Foi também monitorizado o consumo de açúcares e a produção de acetato ao longo do tempo de fermentação, verificando-se que esta estirpe, na presença de substratos não convencionais onde a quantidade de glucose é mínima, tem a habilidade de se adaptar ao meio de cultura. Esta adaptação consiste na capacidade de metabolizar açúcares não PTS (fosfoenolpiruvato), como é o caso da lactose. Verificou-se também que a produção de acetato é residual, não atingindo assim valores considerados tóxicos para a célula.

Por fim, observou-se que as alterações na composição do meio levam a variações quer a nível do número de proteínas quer na quantidade destas, e após análise por IMAGEMASTER 2D 7.0 software foi possível verificar a presença de proteínas diferencialmente expressas nos meios. Assim sendo, e após análise dos spots que contêm estas proteínas, por MALDI-TOF/TOF, foi possível identificar proteínas diferencialmente expressas como por exemplo a proteína MreB, a desoxirribose fosfato – aldolase, chaperonina ClpB e constatou-se que estas estão principalmente envolvidas em processos metabólicos, possuem atividade catalítica, e regulam as principais vias de síntese de nucleótidos.

Abstract

In the last years, plasmid DNA has been used as a vector for gene therapy and DNA vaccines, and for this reason, the ability to produce large quantities of plasmid DNA is important concerning the DNA vaccines production process, on an industrial scale. Thus, the objective of this research was to study the metabolism and proteome of *Escherichia coli* (strain VH35) during the pcDNA- FLAG-p53 biosynthesis using as alternative media agro–food industry compounds, as cheese whey, corn steep liquor and yeast extract.

Combining these alternative compounds as a growth medium, and using strain VH35 as a host, the optimization of plasmid DNA production was performed in terms of yield and quality. After this optimization, it was analyzed the consumption of sugars in the fermentation medium overtime, and the proteome changes that occur due to metabolic adaptations.

The developed work suggests that strain VH35 uses agro-food media as an energy source, thus consuming the lactose present in the agro-food medium, since it does not contain PTS–sugars. Due to the fact that this alteration can lead to proteome modifications in the cell, proteomic analysis was performed using two-dimensional electrophoresis, which showed that the protein composition of strain VH35 was different among the compared growth mediums. Lastly, these proteome changes were analyzed by MALDI-TOF/TOF, and it was possible to identify differentially expressed proteins, such as anthranilate synthase component 1, chaperone protein ClpB, deoxyribose-phosphate aldolase, that are related principally to metabolic pathways and nucleotides synthesis.

Keywords

Proteomics, Two-dimensional electrophoresis, Plasmid DNA, Corn steep licor, Cheese whey, *Escherichia coli* VH35,

Table of Contents

List of Figures

List of Tables

List of Acronyms

2-DE	Two-Dimensional Electrophoresis
% recovery	Percentage of Recovery
AAA	Aromatic Amino Acids
AB	Ammonium Bicarbonate
ACN	Acetonitrile
ATP	Adenosine Triphosphate
BOD	Biochemical Oxygen Demand
CDM	Cell Dry Mass
CHAPS	3-(3-Cholamidopropyl)dimethylammonium)-1-propanesulfonate
CID	Collision Induced Dissociation
CO_2	Carbon Dioxide
CSL	Corn Step Licor
CW	Cheese Whey
DNA	Deoxyribonucleic Acid
DOC	Dissolved Oxygen Concentration
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene-diamine tetraacetic acid
EMP	Embden Meyerf Parnas pathway
$FADH_2$	Flavin Adenine Dinucleotide
G6P	Glucose 6-phosphate
GalP	Galactose Permease
gDNA	Genomic Deoxyribonucleic Acid
HPLC	High Performance Liquid Chromatography
IAA	Iodoacetamide
LB	Luria-Bertani
MALDI	Matrix-Assisted Laser Desorption/Ionization
MALDI-MS	Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry
MALDI-TOF/TOF	MALDI Time-of-Flight Time-of-Flight
Mg/BAC	Galactose ABC Transporter
MS	Mass Spectrometry
NADH	Nicotinamide Adenine Dinucleotide
oc pDNA	Open Circular Plasmid DNA
OD	Optical Density
pDNA	Plasmid DNA
PEP	Phosphoenolpyruvate

phe	Phenylalanine
pI	Isoelectric point
PMF	Peptide Mass Fingerprinting
PPP	Pentose Phosphate Pathway
PTS	Sugar Phosphotransferase System
Pyk A	Pyruvate kinase II
Pyk F	Pyruvate kinase I
RNA	Ribonucleic acid
sc pDNA	Supercoiled Plasmid DNA
SDS	Sodium Dodecylsulfate Solution
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
TCA	Citric Acid Cycle
TFA	Trifluoroacetic Acid
TOF	Time-Of-Flight
tryp	Tryptophan
tyr	Tyrosine
YE	Yeast Extract

Chapter 1 - Introduction

1.1. Gene therapy and DNA vaccination

During the last decades, plasmid DNA (pDNA) has gained considerable importance as therapeutic agent in gene therapy and DNA vaccination, and it is considerate a new generation of biotechnological products [1,2]. The interest in therapeutic applications has been demonstrated by the progressive increase of patents, as can be seen in Figure 1, and clinical trials [1–3]. Indeed, treatments based on pDNA vaccines have been developed for several diseases including cancer (breast, liver, kidney, skin, prostate cancer and lymphoma), hepatitis B and C, malaria, influenza, fir control of dengue, ebola and human immunodeficiency viruses [2].

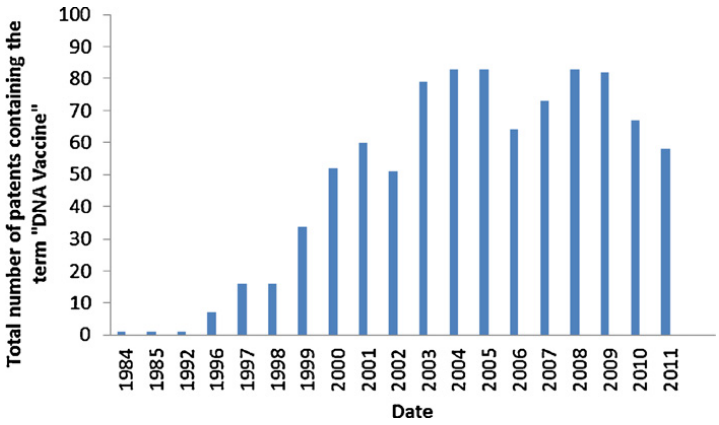


Figure 1- Number of patents in the field of DNA Vaccines (adapted from [2]).

The pDNA molecule consists of two ends of DNA strands covalently linked and highly negatively charged, capable of autonomous replication independently of chromosomal DNA [2, 5]. The pDNA is capable of replicating autonomously with a suitable host and, because it is part of the mobilome (mobile genetic material in a genome), it is often associated with the conjugation mechanism. This mechanism consist in transfer of genetic material between bacterial cells by direct cell-to-cell contact [2].

The helix axis of pDNA can be coiled in space to form a highly ordered structure named supercoiled (sc) pDNA. This structure has been deemed desirable for clinical applications since its conformation presents least risk of recombination and integration into genomic DNA [2,4].

As seen in figure 2, pDNA molecules can exist with varying levels of another two topological conformations: circular (oc) and linear form, however these forms are more subject to rapid intracellular degradation by endonucleases [2,4]. The major advantage of the pDNA molecules is the possibility of using them as cloning vectors into which foreign DNA can be inserted and replicated [5].

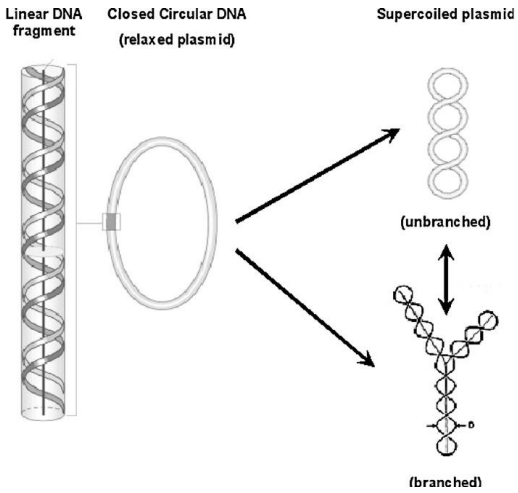


Figure 2 - Schematic representation of DNA structure. Linear, open circular and supercoiled topologies (adapted from [2]) .

Gene therapy involves the introduction of genetic material into the cells or tissues to repair, resulting in effects at the genetic level that may provide novel cellular functions or regulate them [2,6,7]. Also, in vaccination based on gene therapy, the treatment consists in the injection of pDNA containing selected genes encoding antigens to be express transiently in transfected cells, providing a protective immune response against pathogens stronger than conventional vaccines [2,3,6,7].

However, for a successful gene therapy, it is required to design an efficient and safe delivery system capable of transferring the therapeutic gene and of maintaining its stability and functionality. As referred above, pDNA is susceptible to degradation by endonucleases and, thus, it must be ensured that sufficient quantities of pDNA are present to induce the desired immune response [4,7].

Nowadays, gene delivery system can be divided into viral and non-viral vectors [7]. Viral vectors (for example retroviruses, adenoviruses) have higher gene delivery and expression efficiencies, but they present a much worse safety profile: high immunogenicity, cytotoxicity and potential for insertional mutagenesis [1,3,4]. Non-viral vectors already allowed the delivery of several genes encoding molecules such as adhesion molecules, replication inhibitors, tumour suppressor, antigens, and cytokines. These systems are considered to be very safe but less efficient than viral systems, comprising physical and chemical systems, including cationic liposomes, polymers, gene gum, among others [3,8,9]. Therefore, it is possible an effective application of pDNA in cancer treatment, as for example in the recovery of p53 protein function [10,11].

1.1.1. The tumor suppressor p53

The p53 protein is a key tumor suppressor, and its functional inactivation has been associated to many tumor types. Mutations in p53 gene were found in 30 - 50 % of lung, esophageal,

colorectal, head and neck, and ovarian cancers and it was estimate that around 80 % of human tumors have a dysfunctional p53 [12].

The P53 is an unique transcription factor and is considered to be the "guardian of the genome" due to its capability to induce a wide range of biochemical events including apoptosis, cell cycle arrest (G1 and G2 phases), senescence and enhanced DNA repair and modulating metabolic processes [12,13]. So, gene repressor mediated by p53 protein contribute to tumor suppressive activity. Nevertheless, recent studies show that mutant p53 proteins, besides of losing its functions, gain new oncogenic functions that are independent of wild-type p53, including promoting tumor cell proliferation anti-apoptosis, angiogenesis, metastasis and metabolic changes [12,14], as can be seen in figure 3.

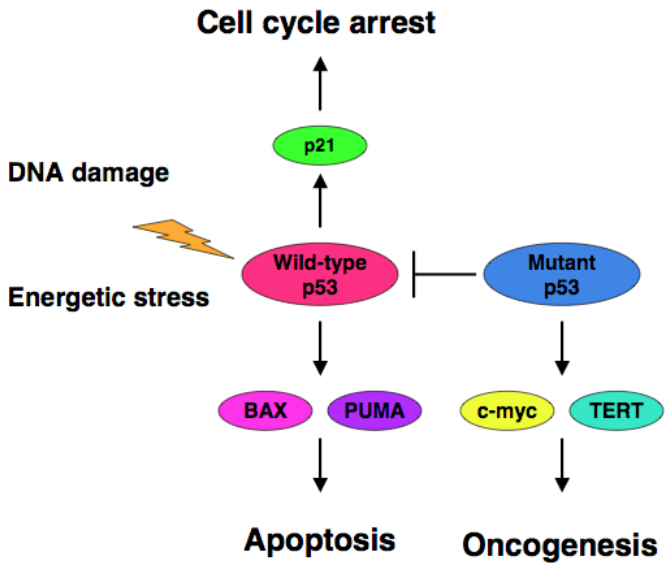


Figure 3- Dominant- negative effect of mutant p53 on wild-type p53. Pro-apoptotic function of p53 is significantly inhibited by certain p53 mutants which induce malignant transformation (adapted from [15]).

Actually, several strategies targeting the reactivation of p53 have been developed, including gene therapy to restore p53 function, inhibition of p53-Mdm2 interaction, reactivation of mutant p53 to wild-type p53, eliminating mutant p53 and p53- based vaccines, as seen in Figure 4 [9,10,12]. Mdm2 is a potent inhibitor of p53 since it binds the transcriptional activation domain to p53, controlled the antitumor activity but in many tumors Mdm2 is overexpressed binds to p53 and inhibits the p53 function [9,16].

Therefore, gene therapy is used to restore p53 functions and it consists in the integration of therapeutic gene product into the human genome and in the delivery of wild-type p53, which triggers a dramatic apoptosis and induces a senescent phenotype [17].

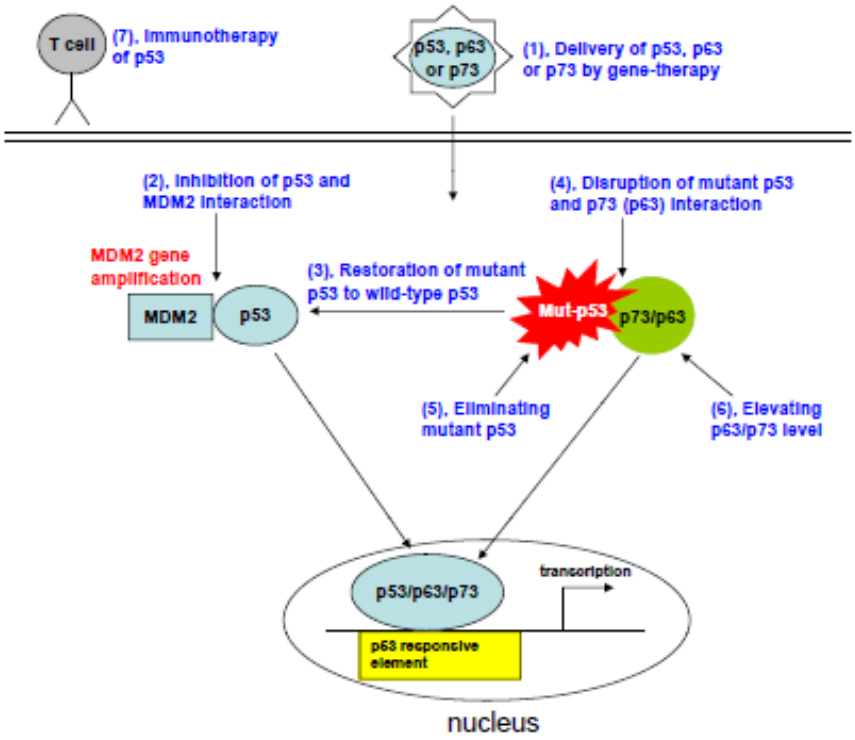


Figure 4- Strategies for targeting restoration of p53 function: (1), delivery of p53, p63 and p73 by gene-therapy; (2), inhibition of p53 and MDM2 interaction; (3), restoration of mutant p53 to wild-type p53; (4), disruption of mutant p53 and p73 (p63) interaction; (5), eliminating mutant p53; (6), Elevating p63/p73 level; and (7), immunotherapy of p53.(adapted from [9]).

1.2. Plasmid DNA Production

To achieve an efficient and economic production of DNA vaccines, it is essential to obtain high yields of functional pDNA using a suitable expression system. The production of DNA vaccines comprises three steps: plasmid design and selection of the host strain, production of the pDNA (upstream process) and recovery and purification of pDNA (downstream process) [4,18,19]. The diagram represented in figure 5 outlines the steps involved in the development and production of DNA vaccines.

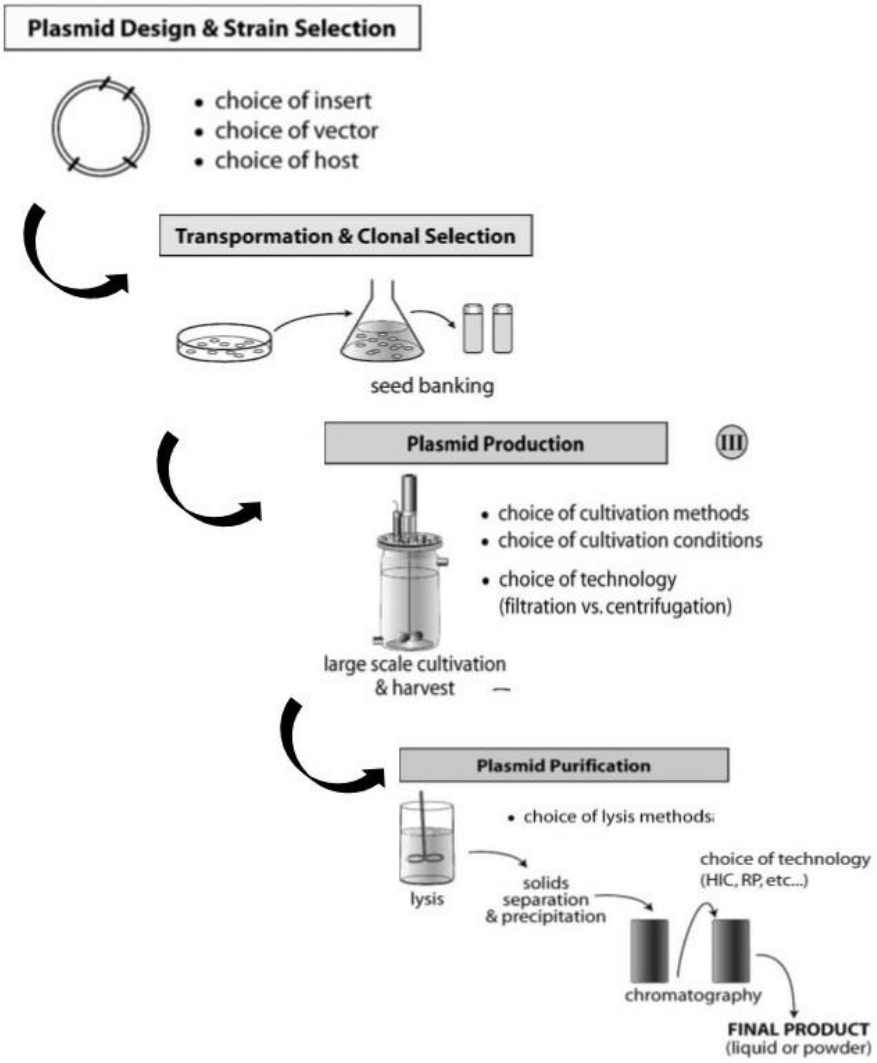


Figure 5- Process steps for the development of plasmid DNA vaccines (adapted from[3]).

1.2.1. Plasmid Design

DNA vaccines and genetic therapy are based on bacterial plasmid that have been engineered to contain a specific gene which express the disease related antigens. Bacterial plasmids used in DNA vaccination have several common features: promoter elements that are active in mammalian cells (gene of interest and polyadenylation (poly A) sequence), transcriptional terminator to terminate transcription and select marker to facilitate production (for example antibiotics such as kanamycin), as it can be seen in figure 6 [4].

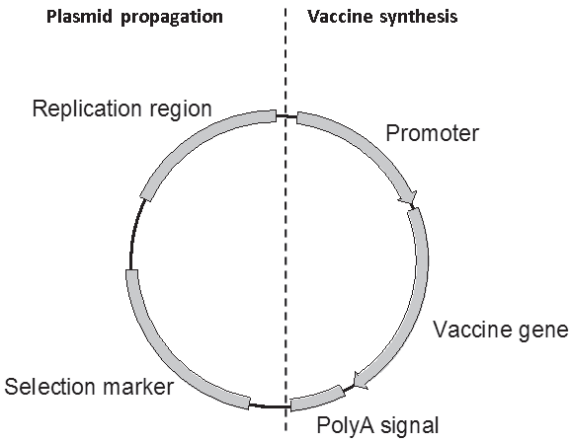


Figure 6 - A schematic structural overview of plasmid DNA (Adapted from [20])

Bacterial replication is defined by the origin of replication (ORI), a specific sequence of vector from which pDNA replication is initiated. The type of the origin of replication determines the copy number of plasmid per cell and influences the pDNA production [18,21]. Currently, the copy-number of pUC ori ColE1-derived is widely applied to reliably obtain a higher copy number of plasmids. The introduction of random or defined mutations into the pUC ori can further increase the plasmid yield. The most commonly used Ori are ColE1 and pUC origin, which induces high plasmid copy number at 37-42 ºC, not at 30-32 ºC, and presents a high-yield pDNA production [22,23].

The most used selection markers are often based on antibiotics such as ampicillin, kanamycin, and tetracycline. However, Food and Drug Administration (FDA) does not recommend the use of ampicillin and other ß-lactam antibiotics due to potential hypersensitivity reactions in patients.

The plasmid construction is difficult due to several factors that need to be taken into account, such as assurance of its structural, segregational and isoform stability. It is also required that the product is homogeneous regarding to its structural form and DNA sequence [19,24,25]. Table 1 summarizes the different factors affecting plasmid stability, once its instability usually induces a decrease in pDNA productivity. It has been demonstrated that the plasmid instability is determined by many factors, which are associated to plasmid copy number, genetic fidelity and segregational stability of plasmid [7] .

Table 1 - Factors affecting plasmid stability (according to [7]).

Type of plasmid instability		
Structural	**Segregational**	**Isoform**
Plasmid size	Host strain metabolic burden	Culture composition
PolyA sequence	Plasmid multimers	Culture medium
Direct, Invert repeats	Lack of selective pressure	Host strain growth phase
Culture conditions	Dissolved oxygen concentration	
Increased expression of transposons	Low pH	
	Culture medium composition	
	Fermentation strategy	

1.2.2. Host selection

Another step involved in pDNA production is the host selection. The principal criteria for a suitable selection of the host strain is to maximize specific and volumetric production, the cell density and the number of copies in order to reduce the production time and the endotoxins levels [26]. It is very important that the host strain is capable of maintaining genetic stability and is amenable to the downstream process [27].

Several alternative host cells have been proposed for pDNA production such as *Saccharomyces cerevisiae, Pichia pastoris, Hanesula polymorf*a [28]**. H**owever, *Escherichia coli* (*E.coli)* remains the most used for pDNA production. It is preferable due to its relative simplicity, inexpensive and fast high density cultivation. It is the most studied organism, capable of fast growth using minimal nutrition [29]. The optimal conditions for pDNA production in *E. coli* will be discussed below. Common problems in pDNA production are plasmid instability, acetate accumulation, substrate inhibition and endotoxin production, which remains associated with the correct folding and lack of post-translational modification [18,28].

1.2.3. Production phase

The production phase consists in two revelant stages: the medium selection and fermentation process development [3,18,30], which are phases that have a significant effect in downstream process [2].

The cultivation medium formulation can dramatically influence the performance of microbial process. Medium composition is involved on the physiology of the host by influencing their intricate regulatory systems and, thereby, the plasmid copy number [21,31]. There are factors to be taken into account in the development of a medium formulation such as the effect

of components on plasmid yield and quality on the further downstream process and in the host metabolism and regulation [21,31].

The culture media can be divided into three types: a minimal media composed only by salts, carbon (C) and nitrogen (N) source and trace elements; a complex media containing in addition to salts and carbon source, one or several complex extracts [32]; and a semi-defined media which is a defined media containing one or more complex extracts [21]. Defined media usually require more components and are more difficult to prepare than complex media [6].

1.2.4. Downstream processing

Downstream processing consists in lysis, isolation and purifications steps, since cellular debris along with all impurities, such as salts, endotoxins and plasmid isoforms must be removed from the final formulations of sc pDNA [2,4].

Generic flow-chart for downstream processing of plasmid DNA consists in cell lysis (alkaline, thermal, mechanical), followed by clarification (filtration, centrifugation), contaminant precipitation (precipitation by chaotropic salts, detergents), plasmid precipitation (precipitation by alcohols, detergents) and finally chromatographic purification (hydrophobic interaction, ionic, affinity chromatography), followed by concentration and final filtration queir[4,33–35].

1.3. Plasmid DNA production in *Escherichia coli*

Escherichia coli, a gram-negative bacterium, is the most used for pDNA production on the industrial scale, because is a simplest and robust expression system, capable of fast growth with minimal nutritional requirements. During fermentation, the percentage of sc pDNA isoform is constant and can lead to high pDNA yields. This organism can be easily manipulated, still it has some disadvantages: problems associated with the correct plasmid folding, endotoxin production and genetic instability [7,28,36].

The *E.coli* strains that have been widely used for pDNA production includes strains as K-12, DH5, DH5α, DH10Β, JM108, JM101, and BL21 [31]. In the last years, several strategies have been developed to enhance pDNA production: optimization of the gene and of the expression plasmid, modification and improvement of host strain and optimization of the fermentation media conditions [28,31].

Emerging strains such as VH33 and VH35, with genetic changes in glucose transport system (PTS), are characterized by producing low amounts of acetate, maintaining pDNA supercoiling with high plasmid copy numbers and high plasmid retention levels (segregational stability) [1,37].

1.3.1. Medium Composition

Tipically, *E. coli* grows in both rich complex organic media and salt-based chemically defined media as long as carbon source is present [6], but have reporter the complex media with a elevate loss of sc isoform and low specific pDNA yields. However, *E.coli* growth in defined media has low yield rates and low reproducibility due to prolonged fermentation times [27,33,38].

The association of medium composition with specific conditions contributes for the controlof plasmid copy number stability and the amount of produced biomass [39]. Elements such as hydrogen, carbon, nitrogen, oxygen, sodium, magnesium, phosphorus, potassium and calcium must be included in the formulation of the medium used for the cultivation of *E.coli* because they have crucial functions for cell growth and plasmid production [6]. When formulating a culture media, there are components that should be taken into account including the carbon source (glucose or glycerol are the most used), the nitrogen source, salts, minerals and trace metals [40,41].

The proportion of carbon source influenced the growth of biomass and plays a key role in the cell yield and in the related acetate production. Besides the two carbon sources normally used for pDNA production, glucose and glycerol, there has been some research regarding alternative carbon sources such as maltose, mannitol, lactose and sucrose. However, both cell mass and plasmid productivity are reduced when these alternative carbon sources are used [42,43].

Some studies have shown that, when glycerol is the preferable carbon source, its metabolism is slower than glucose in order to reduce the maximum specific growth rate during batch fermentation and the acetate production during batch fermentation [44]. Although glucose is the most common carbon source and is essential for the synthesis of nucleotides, this sugar generates more acetate and produces highest volumetric yield [6,42].

Alternativly, nitrogen source is required for biosynthesis of nitrogenous compounds such as amino acids, purines and pyrimidines. Bacterial requirements for nitrogen can be satisfied by several inorganic or organic sources [5] as described in Table 2.

Table 2 - Nitrogen source in minimal, defined and complex media.

Nitrogen source	
Minimal Media	**Semi-defined media or complex media**
Inorganic ammonia and ammonium salts such as ammonium chloride or sulphate [42]	yeast extract, peptone, tryptone and/or casamino acids [24,27,35]

In order to choose an appropriate complex nitrogen source, the amino acid content should be taken into account. For pDNA production in *E.coli*, media formulation is composed by high amounts of yeast extract (YE) [32], a substrate that also contains carbohydrates which is a limiting factor for the carbon consumption rate. Tryptone is not commonly used as nitrogen source in culture media for pDNA production but, as it is deficient in carbohydrates, this substrate may be very useful to establish the consumption rate of the previously selected carbon source [7].

In the culture medium formulation, the C/N ratio must be taken into account, because it has a major impact on plasmid specific yield. Specifically, a C/N ratio of 2.78:1 was considered as optimal for pDNA production [3].

1.3.2. Low-cost fermentative media for plasmid DNA production

The medium growth formulation can dramatically influence the overall cost of the pDNA production process, representing almost 30% of the cost of the fermentation process [45]. So, it is important to reduce the cost of growth medium on both laboratorial and industrial scales. In the last years, many low-cost such as molasses, corn step licor (CSL), cheese whey (CW) and olive mill wastewater have been used in the production of bioethanol, hydrogen and production of rhamnolipids, production of cellulose and production of ß-Galactosidase [46–50].

The cheese whey is the most common by-product of dairy industry and results in the coagulation of milk. This product retains about 55 % of milk nutrients and represents an important environmental problem for dairy industry because of the high volumes and high organic matter content produced [46]. CW contains approximately 55 g/L lactose which is a sugar that can be used in fermentation process as carbon source, and due to whey proteins, it also has a good applicability in the medium showing high nutritional and functional values [48,51,52]. The CW can be used as an alternative carbon source, replacing more common substrates as glucose and glycerol, in cultivation media for *E.coli* with the aim of reducing production costs [53].

The corn step licor is a major by-product of corn starch processing and is considered a low-cost nitrogen source, acting as a replacer for YE and peptone. CSL also is a low-cost source of amino acids, minerals, vitamins and trace elements, and its supernatant composition was determined as: 65% water, 6.1% ashes, 3.4% free reducing sugars, 2.2% total kjeldahl nitrogen, 2.6% fat, density 1.14 and pH 4 [8,46,54]. Over the last few years, CSL have been combined with YE on the production of ethanol, improving its production [46,52,55].

1.3.3. Culture conditions

The pDNA yield has been influenced by several factors such as changes in growth conditions including temperature, nutrient concentration, oxygenation, induction strategies and growth phase [5].

1.3.4. Temperature and pH control

Temperature and pH are external factors that must be controlled during the production process because of their importance for *E.coli* growth rate and cell density. The optimal pH for *E.coli* growth is 5,5 to 8,5 and the temperature is about 37^0 C [23,32]. The growth of *E.coli* cells is inhibited at extreme pHs and, so, at these ranges the cell growth rate will significantly decrease and cause cell death.

1.3.5. Oxygen demand

The dissolved oxygen concentration (DOC) is a major factor and can be used to control and optimize pDNA production. Depending on the desired products, aerobic or anaerobic conditions can be used. The aerobic culture favors faster growth but anaerobic conditions are needed for the formation of certain products including ethanol or lactic acid [56,57].

The DOC attenuates *E. coli* metabolic burden in the aerobic state. The tricarboxylic acid (TCA) cycle operates to oxidize pyruvate with the reductants formed coupling with the electron transport chain in order to generate the proton gradient, which is used for adenosine triphosphate (ATP) production [57].

It was demonstrated that by decreasing the growth rate with limited DOC, the number of plasmid copies per cell and the plasmid specific yield was increased as well as the purity of cell lysates. However, in the presence of a higher DOC, the cell growth and fermentation time decreases, whereas the by-products and mostly acetate increases, leading to cell death [57,58].

1.3.6. Effect of Acetate production

Acetate is predominantly formed in *E.coli* under aerobic conditions, causing several physiological effects and inhibiting cell growth and pDNA production [18,59]. This is demonstrated because the protonated form of acetate is able to cross the cell membrane and act as an uncouple of the proton motive force [60].

Escherichia coli uses the acetate production pathway to produce ATP under anaerobic and even aerobic conditions [61]. The acetate accumulation under aerobic conditions appears to be a result from an imbalance between glycolysis and tricarboxylic acids cycle [18]. The cells produce excess of acetyl-CoA which is converted into acetyl phosphate, an intermediate of acetate pathway, resulting in the posterior production, excretion of acetate and its accumulation in the extracellular environment [61].

1.3.7. Fermentation techniques

Two types of fermentation modes are used to produce pDNA: batch and fed-batch [27]. In the batch process, all the nutrients required for the fermentation process are initially

decontaminated and added into the bioreactor, and the material is only removed at the end of process. The batch fermentation is a simpler process but, when compared to a fed-batch strategy, it is inefficient in terms of plasmid yield and in the conversion of raw materials to product [62].

Fed-batch process starts as batch cultivation, i.e., in a first phase the cells grow until nutrient exhaustion, typically the carbon source. After this nutrient exhaustion, the fed-batch phase begins with the controlled addition of a limiting nutrient, allowing a greater control of culture growth rates [63,64]. This process also provides higher biomass yields and, concurrently higher product yields than in batch cultures. As the substrate is supplied at a controlled rate, its consumption is alomost complete and, so, it never reaches inhibitory concentrations [21,27,65]. This prevents the metabolic overflow resultant from the excess of substrate, reducing the formation of inhibitory concentrations of acetate [66].

1.3.8. Central Carbon Metabolism in *Escherichia coli*

Central carbon metabolism in *E.coli* is constituted by glycolysis, gluconeogenesis, pentose phosphate pathway, and tricarboxylic acid cycle (TCA pathway), and is responsible for transforming carbon into energy, in the form of ATP and into redox cofactors (NADH e NADPH) which are essential for the physiology of *E.coli* [37,67]. These pathways are responsible for biomass and plasmid DNA production and they are intrinsically correlated in the pentose phosphate pathway (PP pathway), composed by the oxidative and the non-oxidative phases [35].

Escherichia coli, as the host organism, has the capability to use a wide range of sugars and sugar alcohols as a carbon source. However, the preferred carbon source is glucose wich lead to a fast growth rate [67–69]. In *E.coli,* the phosphotransferase system (PTS system), which is the main glucose transport system, is responsible for the entrance of carbohydrates and for catalyzing the reaction responsible for glucose phosphorylation. The glucose phosphorylation is coupled to its translocation across the cytoplasmic membrane, mediated by the ratio of pyruvate to phosphoenolpyruvate (PEP). In addition to glucose transport, PTS also regulates sugar uptake in *E. coli* [67,70,71].

The PTS system consists of three distinct soluble proteins that participate in phosphoryl system: pyruvate dehydrogenase (E1), acetyltransferase (E2) and dihydrolipoamide deydrogenase (E3) [67].

Under aerobic conditions, pyruvate enters in the TCA cycle and is oxidized into carbon dioxide (CO_2), where oxygen is a final electron acceptor. This releases large amounts of energy in the form of ATP (adenosine triphosphate), NADH (nicotinamide adenine dinucleotide) and $FADH_2$ (flavin adenine dinucleotide). Fermentation is a process in which cells release energy in the absence of any terminal electron acceptor. This involves the utilization of carbon and the

resulting energy is conserved, at the substrate level, by phosphorylation and by formation of the overflow metabolites (acetate, ethanol, succinate), which are secreted from the cell in order to maintain the redox balance [72].

For example, *E.coli* strains lacking PTS are also capable of transporting glucose across the cell membrane using non-PTS transporters for glucose, such as galactose transporters GalP (galactose permease) and MglBAC (galactose ABC transporter) [72].

In the case of the strain used in this study, strain VH35 (PTS-pykF), the PTS inactivation had a strong effect on the flux distribution and, so, the metabolic flux is directed to biomass formation. As the strain VH35 also lacks the pyruvate kinase isozyme pykF (pyruvate kinase I), the reduction phosphoenolpyruvate (PEP) to pyruvate only depends on pyKA(pyruvate kinase II) activity, which decreases the acetate accumulation during the fermentation process and increases the availability of phosphoenolpyruvate, as it can be seen in figure 7 [37,59].

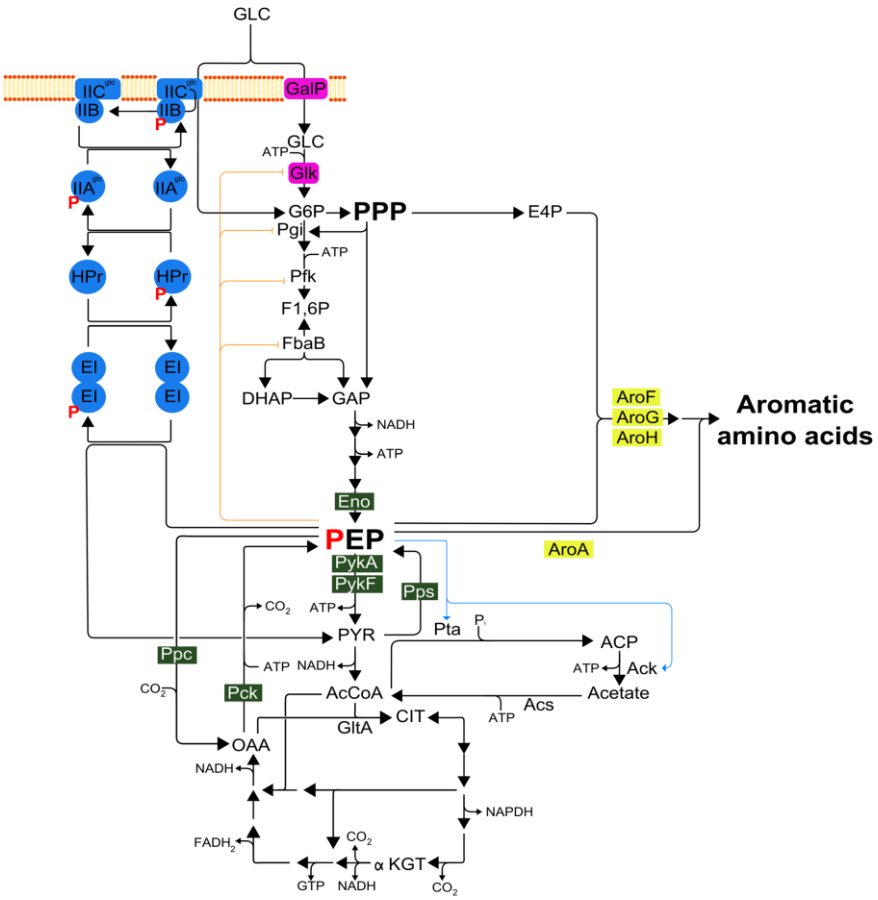


Figure 7 - Central carbon metabolism pathways. The Embden-Meyerhof-Parnas (EMP), pentose phosphate pathway (PPP) and tricarboxylic acid cycle (TCA) in E. coli during its growth in a glucose rich medium. In the figure, PTS (in blue), the non-PTS glucose uptake system (in purple), PEP node (green) and AAA (yellow) enzymes are depicted. The allosteric effects that PEP has in EMP and acetate biosynthesis are represented by an orange line for inhibition and by a blue arrow for activation (adapted from [37]

1.4. OMICS approaches for the study of biological systems

Omics approaches represent a unique opportunity for the quantification and characterization of biomolecules pools, which are present in a specific organism, tissue or fluid. They also have the ability to improve the overall understanding of the functioning of biological systems [73,74]. As a result of technical and methodological improvements, omics areas have emerged to understand and characterize the changes that occur in biological systems, including cell cultures, under certain conditions [73]. Omics include genomics, transcriptomics, proteomics and metabolomics that allow the study of the structure, function and interaction of genes, messenger RNA, metabolites and proteins in an organism [73,75].

1.4.1. Metabolomics

Metabolomics can be defined as a comprehensive and quantitative analysis of metabolites (intermediates or products produced by cellular process) with a view to understand metabolism. Metabolism is constituted by a set of chemical reactions and transformations, and it comprises two parts: the catabolism which is the breakdown of molecules to obtain energy, and the anabolism which is the synthesis of all compounds needed by the cells [76,77].

Under an environmental perturbation, cells are subjected to several physiological stresses and they must recognize the changes and adjust the metabolism systematically [78,79]. Metabolomics studies have been used in order to access information about these responses to environmental stress, by comparing mutants, drug discovery, toxicology and nutrition. This is focused on metabolite target analysis and metabolite profiling of endogenous and exogenous small molecules metabolites (<1500 Da), including peptides, amino acids, nucleic acids, carbohydrates, organic acids, vitamins, polyphenols, alkaloids and inorganic species [76,77].

The greatest advantage of the metabolomics studies is the biological proximity to the system phenotype and hence the rapid observation of its perturbations regarding the metabolome. In recent years, innovational techniques were used in metabolomics studies, including separation, combination and detection techniques. Gas chromatography, capillary electrophoresis, high performance liquid chromatography (HPLC), ultra-performance liquid chromatography and nuclear magnetic resonance spectroscopy are examples of some techniques used in metabolomics studies [76,77].

1.4.2. Proteomics

The term proteomics refers to the study of proteome, which consists in the entire set of proteins expressed by a genome, cell or tissue [74,80–82]. Proteome is extremely dynamic and complex, and its study is gaining interest because it has been proved that the gene expression

has no direct relation to protein expression [74,83,84]. Proteomics also allows the detection of subtle changes that can occur in a studied proteome. This is due to the fact that different conditions can directly affect both the protein expression and DNA synthesis [80,82]. Thus, flux modeling and the determination of the proteome may allow the investigation and interpretation of significant adaptations that occur in metabolic systems at distinct states [82].

In recent years, the study of proteome has made a great deal of progress. The proteomics workflow consists in various stages: acquisition and treatment of material, specific protein profile analysis, and bioinformatics analysis of the obtained data [74,83]. The proteome of several organisms has been studied combining a vast number of methods including two-dimensional polyacrylamide gel electrophoresis (2DE), high performance liquid chromatography (HPLC), mass spectrometry (MS), X-ray crystallography and protein microarrays, as seen in figure 8 [74,85]. These combined methodologies generate raw data that are crossed with databases information using specific algorithms, that are consequently translated into protein information [83]. Table 3 shows some of the techniques which are currently used in proteomics, as well as their advantages and disadvantages. Despite the available techniques, the combination of 2DE and MS has been the preferential workflow for separation and identification of proteins in many biological fields [86].

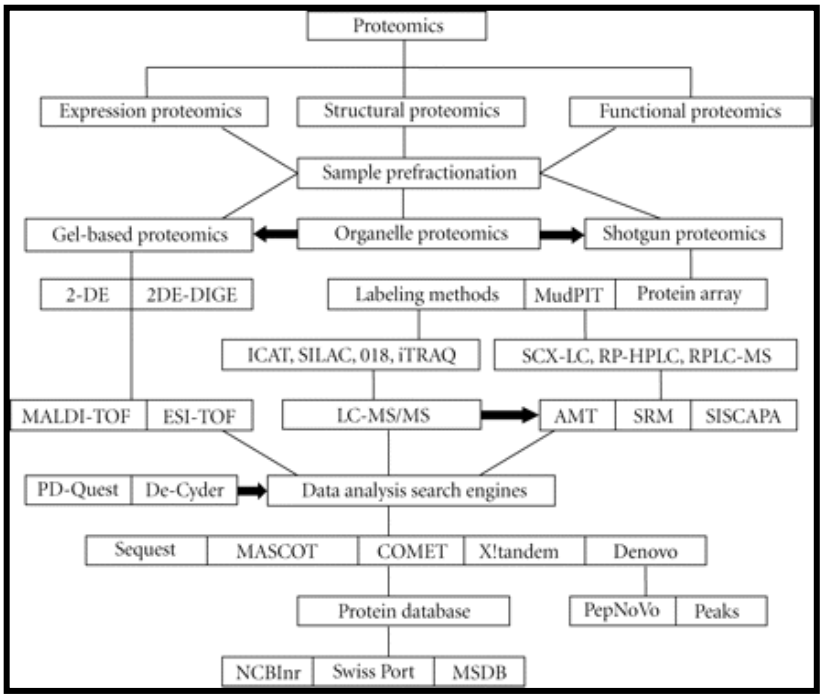


Figure 8- An overview of the available proteomic strategies (adapted from [85]*).*

Table 3- Overview of the currently applied proteomic technologies and its advantages and disadvantages, (adapted from [87]*).*

	ADVANTAGES	**DISADVANTAGES**
PROTEOME PROFILE – 2D-ELECTROPHORESIS	Good resolutions of proteins High sensitivity Detection of postranslational modifications and isoforms Low-cost technique	Poor proteins solubility Limited dynamic range of detection Analysis and quantification are difficult Reproducibility Number of replicates
PROTEIN IDENTIFICATION – MASS SPECTROMETRY	Determination of molecular weight and amino acids sequence information Detection of postranslational modifications High throughout capability	False positives and negatives Secondary assays needed for confirming interactions
IMMUNOAFFINITY	Identifications of protein complex Functional assignment to proteins withouth a priori knowledge of function	High affinity needed between all proteins in the complex Protein loss during purification
HOMOLOGY	Detailed information on specifity and function	Homologous protein structures needed
NONHOMOLOGY	Interation with unknown proteins functional assignment to unknown proteins	Complete genome sequences required limited coverage of genome. Inaccuracy of predictions

1.4.2.1. Sample preparation

In proteomics analysis a suitable sample preparation is crucial since the results of the experiment largely depend on the condition of the starting material. Sample preparation must be laborious and meticulous since the approaches used in proteomic analysis, although have high resolution and high-throughput, are sensible to interferents [85,88]. Samples preparation in proteomics analysis are performed combining several techniques, summarized in table 4, allowing the extraction, solubilization and enrichment of the target proteins, and removal of interfering substances [88,89] .

Table 4 - Methods for samples preparation in proteomic research (adapted from [88,89]*).*

Samples preparation			
Methods of cell disruption:	**Methods for removal of contaminants:**	**Components applied in protein solubiliation:**	**Protein enrichment methods:**
• **Mechanical homogenization** • **Ultrasonic homogenization** • **Pressure homogenization** • **Freeze-thaw homogenization** • **Osmotic and detergent lysis**	• Ammonium sulphate precipitation • Acetone precipitation • TCA acetone precipitation • Ethanol acetone precipitation • Isoelectric precipitation • Chloroform methanol precipitation	• Chaotropes (e.g. urea, thiourea) • Detergents (CHAPS …) • Reductants (e.g. DTT, DTE …) • Protection from proteolysis	• Precipitation • Centrifugation • Membrane proteins enrichment • Prefractionation • Chromatographic techniques • Solid-phase

1.4.2.2. Two-dimensional electrophoresis (2DE)

Two-dimensional electrophoresis is a widely used method for the analysis of complex protein mixtures. It was introduced by O'Farrell in 1975 and consists of two steps of protein separation, a first dimension according to isoelectric point (pI) and a second dimension according to their molecular weight. As referred previously in table 3, 2DE is a relatively low-cost technique and presents high resolution and applicability [90,91]. In a 2DE gel, each spot corresponds to a single protein specie, and information such as pI, molecular weight, and the relative amount of each protein can be obtained. This technique can be applied to analyze the whole proteome, post- and co-translational modifications, which cannot be predicted from the genomic sequence. 2DE is commonly used for detection and identification of potential disease biomarkers but also can be used for bacterial identification, purity check, microscale protein purification and product characterization [90,92]. Despite of its versatility, 2DE technique main disadvantage is the poor performance in the analysis of membrane proteins, a largest category of proteins that remains under-identified [93,94].

As it can be seen in figure 9, the analysis by 2DE is composed by distinct steps, including sample preparation, first dimensional separation by isoelectric focusing (IEF), second dimensional separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (second dimension) and detection of the protein spots by staining.

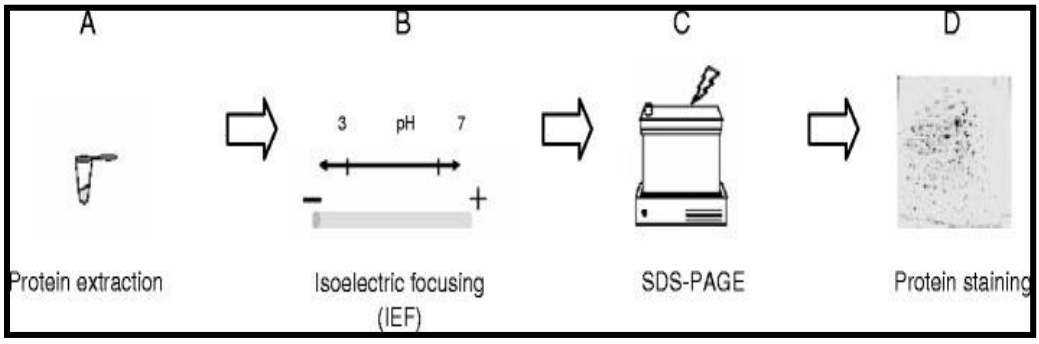


Figure 9 -Typical two-dimensional electrophoresis (2DE) workflow for generating protein maps. 2DE workflow includes (A) sample preparation, (B) isoelectric focusing (1st dimension), (C) SDS-PAGE (2nd dimension) and (D) protein spots staining (adapted from [95]*).*

In 2DE technique, sample preparation is absolutely essential.Ideally, all proteins must be denatured, disaggregated and solubilized before the separation. Also, potential interfering substances must be removed before analysis since 2DE is highly sensible to the presence of interferents. For example, salt concentrations higher than 40–50 mM may interfere with the efficacy of bidimensional separation [96]. In the last years, developments in sample preparation helped to achieve reproducible results. These technical developments include improved extraction buffers and protein precipitation methods and kits for cleaning the interfering substances and for concentrating samples [97–99].

In the second dimension, SDS-PAGE, proteins are separated according to their molecular weight. This technique is based on the traditional process, in which the proteins, negatively charged due the presence of SDS, migrate through the acrylamide gel when a current is applied [98,100]. Second dimension is performed under reducing conditions, so, the complete denaturation and dissociation of proteins is achieved by incubation with dithiothreitol (DTT), resulting on the disruption of its three-dimensional structure by reducing disulfide bonds, unfolding and subsequent complexation with SDS. Iodoacetamide (IAA) alkylates disulfides bonds, converting cysteine residues so that they cannot recombine to form disulfides [100].

For protein visualization, there is a variety of available staining methods including general methods such as Coomassie brilliant blue, silver nitrate and fluorescent stains or specific methods as immunodetection and glycoprotein detection [97,98,101]. The requirements of an ideal detection method include high sensitivity and reproducibility, wide linear dynamic range, low toxicity and should be fully compatible with the posterior MS-based proteomic analysis. Coomassie staining is a simple and low-cost method, suitable for quantitative analysis and is compatible with downstream characterization methods (for example MALDI-TOF/TOF) [102].

Silver staining techniques are a non-radioactive methods based upon saturating gels with silver ions, washing the less tightly bound metal ions out of the gel matrix and reducing the protein-bound silver ions to form metallic silver. Silver ions bind to the amino acid side chains, primarily to the sulfhydryl and carboxyl groups of proteins [102,103]. Fluorescent detection have a higher sensitivity, is a simple and robust staining protocol, and is characterized by its quantitative reproducibility [102,104].

A 2-D image analysis provides statistic evaluation of the protein spots. High quality and reproducible 2-D gels are required to examine patterns and spot intensities to access the difference of the protein expression between samples. The ImageMaster software is the most common in the 2-DE gels analysis, and is designed for automated data processing, however is necessary a correction by user [105,106].

1.4.2.3. Matrix- assisted laser desorption ionization time-of-flight (MALDI-TOF)

For the identification of proteins of interest, spots are extracted from 2DE gels and analyzed by MS. MS has been gaining interest in proteomics field as an analytical technique used for protein identification in unknown samples, but it can also be used for the quantification of several compounds as neuropeptides, antibiotics or various metabolites [107,108]. MS has become the preferred technique for the characterization of the full proteome and for the study of differentially expressed proteins and of post-translational modifications within an organism.

In MS, the analyzed sample is firstly ionized by ionization source and, then, separated in a gas phase through the application of an electric and magnetic fields. The ions are separated according to their mass-to-charge (m/z) ratio in a mass analyzer and the number of ions at each m/z value is recorded by the detector. The vacuum system avoids the collision between the ions and air particles, enabling a free path for the ions from the ion source to the detector [109]. MS analysis are displayed as a mass spectra, where each peak represents the relative abundance of a detected ion as a function of the m/z ratio [109].

The most common ionization method in proteomic studies is the matrix-assisted laser desorption ionization (MALDI) that was introduced by Karas and Tanaka in 1987. Proteins or peptides are mixed with an organic matrix, deposited into a MALDI plate (metal plate) and dried before analysis. Then, the peptide or protein sample co-crystallizes with the matrix, and it is ionized with a laser pulses, under vacuum conditions. The matrix absorbs the most of the energy, protonating the sample and converting it into positively charged ions (+1 charge). Under an electrical field, the formed ions are accelerated into the analyzer and separated according to their m/z ratio. The soft ionization method, MALDI, is commonly combined with the time-of-flight (TOF) analyzer [110,111].

The matrix consists of an organic solid or liquid species and its adequate choice is crucial for the good performance of the technique. [112]. Ideally, the matrix should absorb strongly the laser wavelength, form micro-crystals with the analyte and have a low sublimation temperature, which facilitates the formation of an instantaneous high-pressure plume of matrix-analyte material during the laser pulse duration. The most common acid matrix applied in proteomics approaches are α-Cyano-4-hydroxycinnamic acid (CHCA), sinapic acid (SA) and 2,5-Dihydroxybenzoic acid (DHB) [111–113].

Table 5 - The most common matrix used for the sample preparation for MALDI-TOF/TOF (according to [113]*).*

Matrix	Analyte	Characteristics
CHCA	Low mass proteins and peptides	Forms small homogeneous crystals High internal energy in the analyte High resolution for low mass proteins Solubility: Organic solvents
SA	High mass proteins	Forms small homogeneous crystals Low internal energy in the analyte High resolution for high mass proteins Solubility: Organic solvents
DHB	Peptides, glycoproteins and glucans	Forms big crystal needles Lower resolution Solubility: Organic solvents, water

1.4.2.4. The MALDI Time-of-Flight Time-of-Flight Mass Spectrometer (MALDI-TOF/TOF)

Some MALDI-TOF instruments, due to its configuration, allow tandem MS (MS/MS) analysis, which enables the analysis of the peptide fragments within a selected m/z ratio. These instruments contain a timed ion selector, which is a double-sided deflection gate, that allows only the peptides within a selected m/z pass into the collision cell and can be fragmented through collision induced dissociation (CID). During CID, the peptide is exposed to a chemically inert gas, which separates the peptide into two fragments, where one is ionized and the other becomes uncharged. The gate-selected ions are decelerated by retarding lens prior to their entry into the collision cell. Fragmented ions are reaccelerated and separated by their m/z in the second TOF analyzer [110,111,114]. A schematic view of MALDI ion source connected to TOF/TOF analyzers is shown in the figure 10.

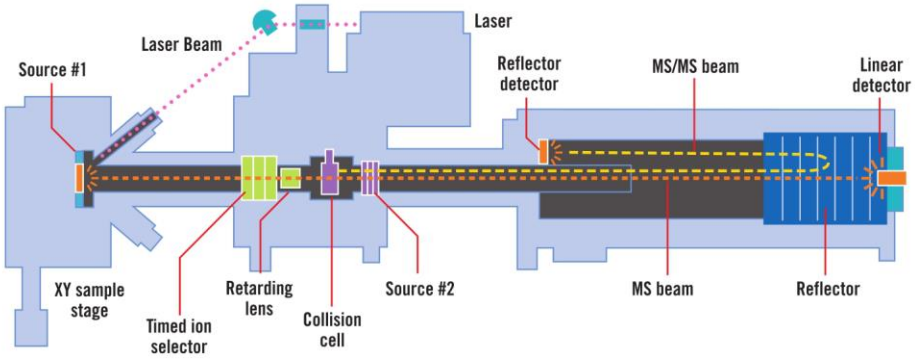


Figure 10 - Schematic view of a MALDI-TOF/TOF mass spectrometer. (Adapted from http://www3.appliedbiosystems.com/cms/groups/psm_marketing/documents/generaldocuments/cms_040103.pdf).

1.4.2.5. Protein Identification

Several different strategies are currently used to identify proteins, based on the comparison of MS or MS/MS spectra with theoretical data that can be carried out through a database dependent or independent searching. Peptide mass fingerprinting (PMF) is the most used strategy to identify proteins separated by 2DE, in-gel digested and analyzed by MALDI-TOF mass spectrometry [115]. This strategy involves searching the experimental peaks list, provided by MS spectra, in a theoretical mass list constructed using a protein sequence database that includes the select protein candidates [116]. Protein database is generated by performing a theoretical digest on protein sequences using specific criteria that closely resembles the experimental conditions [117]. Then, using the proper algorithm, peptides with the highest matching score are selected and identified according to the match protein sequence. However, this strategy only works properly with isolated proteins and, so, requires the previous protein separation. In mixtures exceeding 2 or 3 proteins, a MS/MS based protein identification is required.

The strategies used to identify proteins include the application of the ProteinPilot software, a revolutionary protein database search tool which combines Paragon™ and Mascot™ databases, independently from each other.

Mascot software implements a probability-based scoring algorithm, where a peaks list containing peptide masses is submitted to a database searching to potentially identify proteins from a pool of protein candidates, depending on the obtained score. This is a widely-used database search engine in which the identity score depends on the number and quality of fragment ion assignments [118].

Paragon software allows simultaneous searching of a large number of biological and other modifications, genetic variants and unexpected cleavages.Also performs a statistical analysis regarding the ProtScores from each protein and the total score is based on the sum of all peptides that are related to a certain protein [119]-

1.4.2.6. Proteome of *Escherichia coli*

The organisms respond and adapt them to certain environmental conditions by changing their DNA synthesis, protein expression and, thus, their metabolism. For example, *E.coli* grows in several subtracts and it has the ability to adapt to each medium through changes in mRNA levels and, consequently, by inducing substantial modifications in protein expression and in metabolic pathways. So, as *E. coli* is widely used for the production of pDNA, it is important to understand the host metabolic pathways and the interactions established between the interconnected components. The unbalances between host metabolism and recombinant biosynthesis may hamper the biotechnological bioprocesses efficiency [79,120]. This host is suited for proteome analysis because its complete genome sequence is known, which may facilitate the identification of proteins expressed at a given time in a specifc environment [121,122].

The proteomic analysis became an important tool in order to determine changes in the overall cellular metabolic pathways. The advantage of analyzing these changes using proteomics approaches is the possibility to elucidate global regulatory systems by discovering which proteins are expressed in response to a variety of stress conditions, and getting information related to post-translational modifications and interactions between macromolecules [78,120,123]. Once the proteins of interest are identified, a more precise description of the specific metabolic pathways and mechanisms is trigged in response to changes in the cellular environment. Though 2DE is the most used, a variety of techniques have been used for the study of the proteome and metabolism of *E.coli* strains, as described in table 6. This technique allows the study of changes in a proteome because of its ability to resolve complex protein mixtures into individual polypeptide [79,120,124].

Table 6 - Proteomics studies regarding the analysis of Escherichia coli cultures.

Organism	Objective	Methodology	Observations	Reference
***E. coli* K12 BW25113**	Study of the response to chlortetracycline	Liquid chromatography Matrix assisted laser desorption ionization Mass spectrometry	723 proteins were identified (184 downregulated and 147 upregulated), from which 51.3% were involved in metabolic pathways such as pyruvate metabolism, pyrimidine metabolism, the TCA cycle and butanoate metabolism-	[120]
***E. coli* VL8**	Identification of sinigrin-induced bacterial proteins potentially involved in the metabolism of glucosinolate	2DE LC–MS/MS	16 proteins are involved in carbohydrate metabolism, oxidoreduction system and sugar transport.	[125]
***E. coli* K-12 strains**	Identification phosphate limitation	2DE and MALDI TOF/TOF	Proteome analysis confirms that phosphate limitation occurs.	[121]
E.coli	Identification of proteins induced by acid or base, during long-term aerobic or anaerobic growth conditions in complex medium	2DE and MALDI TOF /TOF	Identification of complex relationships between pH and oxygen and a novel permeant acid-inducible gene, YfiD.	[126]
***E. coli* K12**	Regulation of the main metabolic pathways of cells grown aerobically in different carbon sources (glucose, acetate, gluconate or glycerol).	2DE and MALDI TOF /TOF	Acetate was used as a carbon source: pfkA, pykF, ppc and zwf were downregulated, while fbp, pckA, ppsA and mez were significantly up-regulated. Glycerol as a carbon source, fbp and TCA cycle enzymes were up-regulated, while ackA was significantly downregulated.	[78]

Chapter 2 - Objectives

Gene therapy and DNA vaccination has become a promising alternative for treating many serious diseases, such as cancer, neurodegenerative diseases and genetic disorders. In the last years, several vectors have been developed and gene therapy clinical trials have been completed or are ongoing. In particular, a system based on the gene encoding a p53 protein, a tumour suppressor, has been improved for cancer treatment.

The main goal of this work is to study the metabolism and proteome of *E.coli* (strain VH35) during the pcDNA-FLAG-p53 biosynthesis using agro–food industry compounds as an alternative media. This work involves metabolic analysis of growth media and proteomic analysis of *E.coli* lysates in order to understand how specific molecules, such as sugars, amino acids, and nucleosides metabolically can interact in order to improve pDNA production. In particular, the study of the proteome of *E.coli* lysates aims to understand the adaptations of *E.coli* metabolism to several growth media.

For this purpose, the production of therapeutic pDNA was optimized, in terms of yield and quality of pDNA, using *E.coli* VH35 as a host and combining CW and CSL in the growth medium, as sources of carbon and nitrogen, respectively. It was also important to explore the capacity of *E.coli* VH35 to produce pDNA from alternative carbon sources and to verify the variation in consumption of sugars that occurs throughout the fermentation process.

Proteomics analysis was also performed in order to provide a comprehensive description of proteome changes of *E.coli* VH35 that may occur in their metabolic adaptation to different conventional and non-conventional growth media.

Chapter 3 - Materials and Methods

3.1. Materials

Acrylamide 4K-solution 40% was obtained from PanReac AppliChem (Darmstadt, Germany). Coomassie brilliant blue G-250 was obtained from Fluka Chemika (Buchs, Switzerland). CHAPS was obtained from Amresco (Ohio, USA). Sodium phosphate, sodium chloride and glycine were obtained from Fisher Scientific (Loughborough, UK). Dithiothreitol (DTT) was obtained from HiMedia (Mumbai, India). Methanol (HPLC-grade) was acquired from VWR Internacional (Carnaxide, Portugal). The ethylenediaminetetraacetic acid (EDTA), glucose, sulfuric acid, acetic acid (HPLC grade), L-tyrosine, L-phenylalanine and L-tryptophan (≥99 % purity) were acquired from Sigma-Aldrich (Sintra, Portugal). Sodium sulfite was acquired to José M. Vaz Pereira, S.A. (Sintra, Portugal). Tris (hydroxymethyl) aminomethane (Tris), Ammonium bicarbonate, trypsin from porcine pancreas Trifluoroacetic acid (TFA) and iodoacetamide (IAA) were obtained from Sigma-Aldrich (Missouri, USA). Ultrapure reagent-grade water used for preparing solutions was obtained from the Mili-Q system (Milipore/Waters).

3.2. Plasmid and strain

The 6.7 Kbp pcDNA3-FLAg-p53 plasmid was purchased from Addgene (Cambridge, MA, USA). The vector pcDNA3-FLAG-p53 encodes for the human p53 protein conjugated with a FLAG tag, and contains the ampicillin resistance gene and SV40 virus mammalian expression promoter.

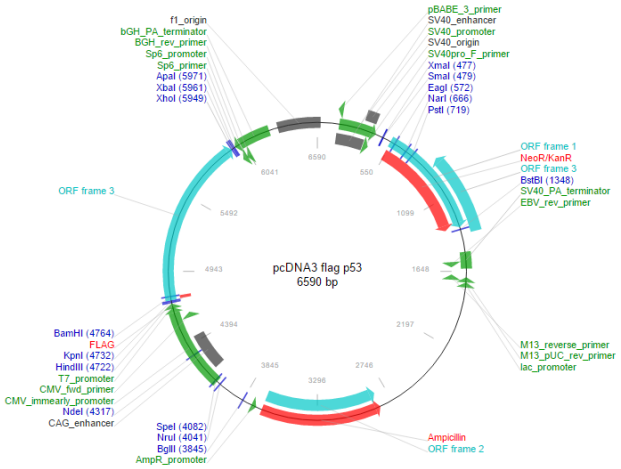


Figure 11- Plasmid pcDNA3-FLAG-p53 backbone (adapted from https://www.addgene.org/10838/).

The strain used was *E. coli* VH35, which was gently provided by Professor Guillermo Gosset from the *Instituto de Biotecnología* from *Universidad Nacional Autónoma de México*. Strain VH35 derivate from *E. coli* W3110 was modified for L-phenylalanine production, as consequence

of PTS and PykF inactivation. The *E. coli* VH35 was transformed with pcDNA3-FLAg-p53 plasmid according to the protocol previously described [59].

3.3. Bacterial growth conditions

E. coli VH35 strain containing the pcDNA3-FLAg-p53 plasmid, was first cultivated in a Luria-Bertani (LB) agar plate, supplemented with 100 µg/ml of ampicillin. To define the optimal conditions for pDNA synthesis, several media composed by CW and CSL were tested for the fermentation of *E.coli* VH35. The tested growth media consisted in the following conditions: 1 and 2 % (w/v) of CW, 2 % (v/v) of CSL combined with 0, 2.5, 5, 7.5 and 10 g/L YE.

For media preparation, CW power and YE were diluted with sterile water and autoclaved, followed by the removal of the precipitate by centrifugation (1 hour at 3900 g). CSL was diluted to 50% with sterile water, pH was adjusted to 7 and the CSL solution was autoclaved and centrifuged at 3900 g during 40 min to remove the insoluble material.

In order to improve pDNA synthesis through the control of the aromatic amino acids pathway, from the previously optimized media, were supplemented with 50 µg/l tyrosine, 78.5 µg/ml phenylalanine and 100 µg/ml tryptophan, according to the previously described [59]. The aromatic amino acids used in experiments were dissolved in Mili-Q water at a final concentration of 10 mg/mL, sterilized using a 0.22 µm syringe filter and stored at 4ºC in absence of light.

Specific fermentation assays were carried out at 37 ^{0}C and 250 rpm, in 500 mL Erlenmeyer containing 125 ml of complex medium, supplemented with 100 µg/ml ampicillin and salts (90 mM KH_2PO_4, 40 mM $(NH_4)2SO_4$, 20 mM NaCl, 1.6 mM $MgSO_4 \cdot 2H_2O$, 0.05 mM $CaCl_2$, 0.072 mM $FeSO_4 \cdot 7H2O$). Cell growth was evaluated by measuring the optical density of the culture medium at wavelength of 600 nm (OD_{600}). All cultures started with an OD_{600} of approximately 0.2 by inoculation from the pre–culture medium. The cells were grown in 500 mL shake flasks containing 125 mL of medium (equal to fermentation medium), at 250 rpm at 37ºC. Grown cells were suspended at late log phase and the cells were recovered by centrifugation at 3800 g for 45 min at 40ºC (Beckman coulter Allegra X22 centrifuge). Lastly, the bacterial pellets were stored at -20 ºC.

3.4. Quantitative and qualitative analysis of the pDNA

3.4.1. Alkaline Cell Lysis

The bacterial pellets were thawed and dissolved vigorously in 8 mL of resuspension buffer containing 50 mM glucose, 25 mM Tris- HCl and 20 mM EDTA, pH 8. Alkaline lysis was performed by adding 8 mL of a 200 nM NaOH, 1% (w/V) sodium dodecylsulfate (SDS) solution and, after 5 min of incubation at room temperature, cellular debris, gDNA and proteins were precipitated by adding and mixing 7 mL of a pre-chilled solution of 3M potassium acetate, pH 5.0. After a 20 min incubation in ice, the precipitate was removed by centrifugation at 20 000 g, during 30 min at 4 ^{0}C, using Beckman Coulter Allegra 25R centrifuge. A second centrifugation step was carried out, under the same conditions, in order to separate the remaining suspended material. For the following analysis, the extracts resulted from the second centrifugation were used without further purification steps.

3.4.2. Plasmid DNA Quantification

All experiments were performed using on ÄKTA Pure System (GE Healthcare Biosciences, Uppsala, Sweden) with the Unicorn control System Version 6.3, equipped with 20 µL sample loop. The HPLC method based on hydrophobic interaction chromatography [34], was performed in order to measure the concentration and purity of synthetized pDNA using a 4.6/100 mm HIC Source 15 PHE PE column (Amersham Biosciences). Briefly the system was prepared with 1.5M $(NH_4)_2SO_4$ in 10 mM Tris-HCl buffer pH 8.0 in pump B and 10 mM Tris-HCl buffer pH 8 in pump A. The hydrophobic column was equilibrated with 100 % of buffer B at a flow rate of 1 mL/min. After injecting the lysis extract, the non-retained species were removed from the column for 0.8 min by maintaining the initial conditions. The concentration of ammonium sulphate was then instantaneously decreased to zero in order to elute bound species [34]. The absorbance at 260 nm and the conductivity were monitored during the assay.

The standards used for calibration curve were prepared with sc pDNA obtained from *E.coli* fermentation extracts purified with a Quiagen Commercial Kit (Hilden, Germany). Standards were obtained by diluting the sc plasmid solution with the mobile phase to concentrations ranging between 1 and 100 µg/mL. The calibration curve obtained from the correlation between the peak area and the concentration sc pDNA is depicted in appendix I. The purity degree was defined as the percentage of pDNA peak area related with the total area (area of all peaks on the chromatogram).

3.4.3. Plasmid DNA quality analysis

For pDNA quality analysis, the isoforms proportion was assessed by agarose gel electrophoresis, a technique that applies an electric field to an agarose gel matrix to separate a mix of DNA molecules. The samples were analysed by horizontal electrophoresis using a 1% agarose gel and 0.5 µg/mL green safe. Electrophoresis was carried out at 110 volts, during 30 min, using tris-acetate-EDTA (TAE) as running buffer. The agarose gels were visualized under a UV light in a Viber Lourmat System (ILC LDA Lisbon, Portugal).

3.5. Determination of sugars and acetate by High Performance Liquid Chromatography

All experiments were performed using a HPLC from Agilent (Waldbronn, Germany) with a refractive index detection (RID – Agilent 1260 Infinity). Sugars and acetate were simultaneously analysed onto an analytical column HI-PLEX H (7.7 × 300 mm and 8 µm of pore size). The mobile phase consisted of Mili-Q Water with 5 mM H_2SO_4. The compounds were eluted under the following conditions: flow rate at 0.6 mL/min in an isocratic mode, temperature of 50 ^{0}C and pressure of 50 mmHg [59].

Samples consisted on collecting 1000 µL of cell culture broth, and were centrifuged at 10000 rpm for 7 min. The supernatant was filtered (0. 22 µm) and previously degasified. A calibration curve was determined for glucose, galactose, lactose and acetate by diluting mother solution at a concentration ranging from 1- 100 g/L and 1-40 g/L for galactose.

3.6. Two Dimensional Gel Electrophoresis

3.6.1. Sample preparation

After the optimization process, the fermentation conditions, shown below in table 7, which demonstrated increased production of pDNA, were selected for further analysis of the proteome *E. coli* by 2DE.

Table 7- Medium Composition

MEDIUM COMPOSITION
CSL 2% (v/v)+ 1.2% *CH* (w/v)
CSL 2% (v/v)+ 1.2% *CH* (w/v)+ + 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan
CSL 2% (v/v)+ 1.2% *CH* (w/v)+5 (g/l) ye
CSL 2% (v/v)+ 1.2% *CH* (w/v)+5 (g/l) *YE*+ 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan
20 g/l *YE*, 20 g/l glucose, 24 g/l triptone and 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan

Upon completion of the fermentation process, 125 mL of cell broth was harvested by centrifugation at 3800 g, 4 ^{0}C for 40 min, and washed three times with PBS Buffer 1X for 20 min, at same conditions. Cells were suspended in lysis buffer (7 M Urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM DTT, 10 mM TRIS- HCL) and lysed by mechanical glass bead disruption (seven cycles of vortexing during 1 min, followed by 1 min of incubation on ice). The mass of glass beads and the volume of lysis buffer were added to the mass of cell pellet in a proportion of 2:2:1. To the resultant lysates was added 25 µg/ml of the DNAse, followed by incubation on ice for two min.

The proteins from lysates were extracted using chloroform/methanol precipitation, in which methanol, chloroform and Mili-Q water were added to 200 µl of sample, in a volume proportion of 4:1:3:1. Then, the mixture was vortexed and centrifuged at 15 000 g, 4 ^{0}C for 10 min. The aqueous layer above the proteins precipitate, formed between the hydrophilic and hydrophobic phases, was removed and more 4 volumes (800 µl) of methanol were added. The protein pellet was obtained after a centrifugation at 10 000g for 5 min. The supernatant was removed, and the pellet was solubilized in the 2DE sample buffer (7 M Urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM DTT, 2 % (w/v) IPG buffer 3-10, 0.002 % bromophenol blue solution).

3.6.2. Protein quantification

The 2D Quant Kit (GE Healthcare, UK) was used to determine the accurate quantity of protein in the samples previously prepared, based on manufacturer's instructions. Bovine serum albumin (BSA, 2 mg/mL) was used as standard for the construction of calibration curve to determine the protein mass, in an range from 2.5 to 5 µg. Briefly, 125 µL of precipitant reagent was added to each sample/standard and, after 3 min incubation at room temperature, 125 µL of co-precipitant reagent was added. Protein sample was centrifuged at 10000 rpm for 10 min, 4 ^{0}C and the supernatant was removed. Pellets were solubilized with 100 µL of a copper solution (25% cooper solution and 75 % Mili-Q water), 200 µL of working color reagent was added to each tube and the solution was transferred to a 96-wells plate. After a 20 min incubation at room temperature, the plate absorbance was read at 490 nm.

3.6.3. Two-Dimensional gel electrophoresis

For the first dimension, immobiline DryStrips pH 4-7, 18 cm (GE Healthcare Life Sciences, Sweden) were rehydrated on an immobiline Drystrip dehydratation tray (GE Healthcare Life Sciences), for 12 h at room temperature, with 500 µL of rehydratation solution (7M Urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM DTT, 0,5 % (w/v) IPG buffer, 0.002 % bromophenol blue solution), as previously performed [92].

A total of 400 µg of protein sample was applied onto each strip using a cup loader. The rehydrated strips were then subjected to isoelectric focusing in Ettan IPGphor III (GE Healthcare Life Sciences) using the following conditions: 500 V for 1 hour, an increasing voltage gradient until 1000 V for 1 hour, a gradient from 1000 to 10000 V for 4 hours and, finally, the voltage was set to 10 000 V for 2h30. Biological and technical duplicates were performed for all 2DE experiments.

After isoelectric point, IPG strips were incubated with reducing equilibration buffer (6 M urea, 29.3% (v/v) glycerol, 2% (w/v) SDS, 0.02% (w/v) bromophenol blue, 75 mM Tris-HCl buffer, 1 % (w/v) Dithiothreitol (DTT)) for 15 min, followed by a 15 min incubation with alkylation equilibration buffer (6 M urea, 29.3% (v/v) glycerol, 2% (w/v) SDS, 0.02% (w/v) bromophenol blue, 75 mM Tris-HCl buffer, 2.5 % (w/v) iodoacetamide).

The second dimension was performed on 12.5% acrylamide gels, using Ettan DALTSIX Large Vertical System (GE Healthcare Life Sciences, Sweden). The 12.5% acrylamide gels were casted into 1 mm Ettan DALT gels casting cassettes. Strips were positioned on the top of acrylamide gels and sealed with agarose solution (0.5% (w/v) agarose, 0.002% (w/v) bromophenol blue, 25 mM Tris, 125 mM glycine, 1% (w/v) SDS). The second-dimensional gel electrophoresis was carried out using a 1x electrophoresis buffer into the under chamber and 2x buffer into the upper chamber. The electrophoresis running buffer was prepared by diluting the stock 10x

composed by 250 mM Tris, 1.92 M glycine and, 1% (w/v) SDS, pH 8.3. Gels were initially run at 2 mA/gel for 45 min to allow the proteins entrance in the gel and, then the amperage was increased to 17 mA/gel in order to allow the proteins separation into the resolving gel.

After electrophoresis, gels were fixed (40% ethanol, 10% acetic acid) during a day, stained with colloidal coomassie brilliant blue solution during 5 days, and distained (1% acetic acid, 20% ethanol) overnight, according to the previously described procedure [101].

3.6.4. Image acquisition and analysis

The gels were also subjected to a densitometry scan on ImageScanner III (GE Healthcare Life Sciences, Sweden). The images were analyzed using the ImageMaster 2D Platinum v7.0 (GE Healthcare) software. The protein spots were automatically detected in the 2DE gels (smooth 2, minimum area 40.0 and saliency 1) and minimal manual correction was performed to remove artifacts and edges of the images. Manually assigned vectors (landscapes) were implemented on gel images to correct technical variances. T-test (ANOVA), with p-values ≤ 0.05, to identify the proteins with significant expression differences, namely differentially expressed proteins.

3.6.5. Mass Spectrometry

3.6.5.1. Trypsin digestion

For identifying the differentially expressed proteins, protein spots were excised from 2DE gels, and were digested with trypsin. Gels spots were destained with 50 % acetonitrile (ACN) and 25 mM ammonium bicarbonate (AB) at 37ºC, overnight. After removing the wash solution, pure ACN was added to rehydrate the gels and proteins were reduced with 10 mM DTT, at 56º C for an hour, and alkylated with 55 mM IAA, at room temperature for 30 min. Then, gel spots were rehydrated using 30 µL of 10 ng/µL trypsin solution prepared in digestion solution (25 mM AB and 9% ACN) and incubated in ice for one hour. After the absorption of trypsin solution, gel spots were covered with digestion solution and proteins were digested overnight at 37ºC. Lastly, the tryptic peptides were firstly extracted, using 0.1% TFA in water (at 37ºC for 15 min) and, then, with 0.1% TFA in 50% ACN at the same conditions. The extracted tryptic peptides were polled in the same tube and dried by vaccum centrifugation at room temperature.

3.6.5.2. MALDI-TOF/TOF analysis

For clean-up peptides, dried samples were acidified with 1% TFA to obtain a pH lower than 3. Zip-tip pipette tips C18 0.1-10 µL pipette tips (Millipore®, Molsheim, France) were activated 5 times using pure ACN and equilibrated 10 times with 0.1% TFA in LC-MS grade water. The tryptic peptides were loaded on the column and afterwards bound peptides were eluted with 5 µL of 80% ACN with 0.1% TFA solution. Thus, with tryptic peptides desalted and concentrated, it was possible to perform MALDI TOF/TOF analysis.

The MS and MS/MS spectra were acquired on 4800 plus MALDI-TOF/TOF analyzer (Applied Biosystems), equipped with a 355 nm laser. Initially, CHCA matrix was prepared by adding 60 µL of 0.1% TFA in 50% ACN in order to obtain a matrix solution with 5 mg/mL concentration. The standards used for the equipment calibration were a mixture of CALMix 1 and CALMix 2, obtained from AB SCIEX Peptide Mass Standard Kit (Massachusetts, USA), according to the manufacter's instrcutions. MALDI-TOF/TOF was initially calibrated, using components from the calibration mixture such as des-arg-bradykinin (904.4681m/z), angiotensin I (1296.6853m/z), glu-fibrinopeptide B (1570.6774m/z) and ACTH (2093.0867m/z).

Samples were prepared by combining 5 µL of clean-up peptide samples with 5 µL of CHCA matrix solution. All spots were acquired in a positive MS reflector mode in the range 800 to 4000 m/z by averaging 1500 laser spots. The eight more intense MS ions per spot that satisfied the precursor criteria (200 ppm fraction-to-fraction precursor exclusion, S/N ratio >25) were selected for subsequent MS/MS analysis. MS/MS analysis was performed using 1 keV collision energy of 1 kV with a total of 1500 laser shots per spectrum. Peak lists were export to an MGF file using the function Peaks to Mascot 4000 Series Explorer™ Software (Applied Biosystems).

3.6.5.3. Protein Identification

Protein identification was performed using Paragon algorithm from ProteinPilot™ Software 4.5, from AB SCIEX (Massachusetts, USA), under a 95% confidence, where each peak list was searched against the *Escherichia coli* database (1,447,752 entries) downloaded from Swiss-Prot *Escherichia coli* in FASTA format, at 15 September 2016. The "unused" and "total protein" scores are ProteinPilot specific terms that stands for peptides specific for protein identification and all peptides measured for protein identification, respectively. In this study an unused/total protein score of 1.3/1.3 was used as a cutoff for protein identification. The search parameters took into consideration cysteine modification by methyl-methanethiosulfonate, digestion of peptides with trypsin and default biological modification settings [105].

After protein identification, STRING10 was used to share protein-protein associations. To get a more profound knowledge about biological process, cellular compartment and molecular function of each identified protein, STRAP software was used. STRAP (Software Tool for Rapid Annotation of Proteins) is a program that automatically annotates a protein list with information which helps in the meaningful interpretation of data from mass spectrometry and other techniques.

Chapter 4 - Results and Discussion

4.1. Optimization of media based on agro-industrial by-products on the production of plasmid DNA by *E. coli* VH35

In the last years, the demand for obtaining high yields of pure and effective pDNA has been increased. Several procedures for pDNA production and purification have been proposed, although much effort has been direct towards to downstream process. However, there is still much room for improvement at the upstream level, namely on the production phase [19,35].

The production phase consist of two important components: the medium selection and fermentation process [3,30,59,127] . The cultivation medium can represent almost 30% of the total cost for a pDNA production and, then, complex medium are not economically attractive because requires more expensive components [6,45]. So, this study attempted to find economical and profitable alternatives for the preparation of complex culture media, by the application of agro-industrial by-products.

4.1.1. Effect of the supplementation of media based on agro-industrial with yeast extract on the production of pcDNA3-FLAg-p53 plasmid

CSL and CW are considered agro-industrial wastes/by-products [55]. CW contains high levels of lactose, which can be used as a carbon source in the complex medium for the production of compounds by fermentation. However, CW must be supplemented with other extracts, such as CSL and YE, to meet the metabolic needs of the cells during fermentation process [8,51]. CSL is an alternative nitrogen source and its combination with YE may increasing cells growth and expression [8]. Previous results from our research group show that CSL and CW as alternative substrates in culture media for production of pDNA in *E.coli* VH35, a combination of 1.2% CW and 2% CSL demonstrate to be a optimal medium for pDNA biosynthesis.

Thus, in a first approach, the effect of the supplementation of medium based on by-products with YE for the production of pDNA was studied. So, for improve the production of pDNA, 2% (v/v) CSL and 1.2% (w/v) was combined with several YE concentrations, as shown in the table 8.

Table 8- Composition of the different fermentation media tested in this study with 1.2 % (w/v) CH and 2% (v/v) CSL

Medium	YE (w/v)
A	0.0
B	2.5
C	5.0
D	7.5
E	10.0

Regarding the pDNA production yields, from the five analyzed media, the medium C showed the higher production of pDNA, unlike the medium, A which presents low levels of production of pDNA. The results for the pDNA volumetric and specific yield revealed that the introduction of YE in medium leads to an increase in final pDNA specific yield (Table 9).

According the table 9, pDNA specific yield increase from 1,048± 0,022 µg pDNA per cell dry mass (CDM) in medium A to 4.983 ±0.37 µg pDNA per CDM in the medium B and C for 10.234 ± 0.656 µg pDNA per CDM in medium C. The composition C presented the most significant increase in the volumetric and specific yield but led to a decreasing growth rate of *E. coli* cells (2.972±0.193 g/L of CDM). Indeed, by comparing the CDM obtained with media A and B, it is apparent that the increase of YE levels did not result in a significance increase CDM value, as can be seen by CDM values presented in table 9 (4.18± 0.047 for medium A and 5.29± 0.129 for medium B). This indicates that the presence of YE in fermentation medium, at moderate concentrations, may allow the increase of *E. coli* metabolic flux towards the pDNA production, rather than for cell growth [59].

Table 9 - Effect of the supplementation of the media with yeast extract on bacterial growth, plasmid DNA volumetric, specific yields, and purity.

Medium	*$OD_{600\ nm}$*	*Cell dry mass (CDM) (g/L)*	*pDNA volumetric yield (µg/mL)*	*pDNA specific yield (µg pDNA /mg CDM)*	*Purity (%)*
A	9.825 ± 0.643	4.18± 0,047	4.379±0.096	1.048± 0.022	0.192± 0.037
B	11.073±0,843	4.983 ±0.379	8.654±0.348	4.983 ±0.379	0.616 ±0.037
C	6.613±0.430	2.972±0.193	35.126±1.067	11.803±0.700	2.902±0.043
D	9.267 ±0.379	4.170 ±0.171	21.798 ±2.242	5.228 ±0.766	2.079 ±0.159
E	11.7750.286	5.299 ±0.129	10.262 ±0.814	5.299 ±0.129	0.894 ±0.095

However, for higher concentrations of YE, above 7.5 and 10 g/L, the strong impact of YE on the production and consequent plasmid yield is softened. Despite of showing higher production than

the obtained with medium A, it was possible to obtain volumetric yields of 21.798 ± 2.242 and 10.262 ± 0.814 with media D and E, respectively, lower values taking into account the results for medium C.

The provided results shows that the presence of moderate concentration of YE in the medium had a strong impact on pDNA production, maybe due to the disturbance of carbon/nitrogen ratio in the medium [40,55]. Finally the medium C appears to be a promising alternative for the production of the pDNA using CSL and CW as a components of the growth medium, showing volumetric and specific yield of 35.126±1.067 and 11.803±0.700, respectively. So, the medium C was selected for a further optimization of pDNA production.

It was achived 35.126±1.067 µg pDNA/ mg cell dry mass and also was confirmed that the use of agro –food nutritive extracts is viable for pcDNA3–FLAG–p53 biosynthesis. In fact the results are similar to those obtained with commercial semi-defined medium with glucose and aromatic amino acids supplementation [59].

The quality of the obtained plasmid is an important factor to be considered and it was evaluated using gel electrophoresis. The analysis by gel electrophoresis demonstrate that the tested agro-food medium could produce high degree of sc pDNA and the presence of the other isoforms can also be observed (Figure 12).

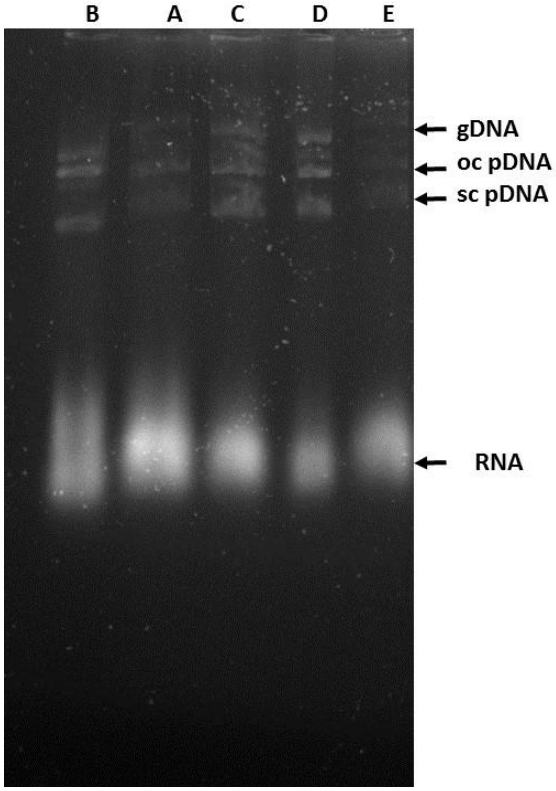


Figure 12 - Agarose gel electrophoresis of target plasmid DNA during fermentation. Lane A, B, C,D,E are fermentations of indicated in table 4.

By the analysis of agarose gels in the figure 12, the plasmid produced at the medium C was predominantly in the sc form, in contrast the medium A and the medium B showed a low yield of sc pDNA. The presence of other pDNA topologies occur as a result of degradation, during the process, caused by intrinsic and extrinsic factors, per example, the presence of nucleases or the unsuitability of fermentation process [128,129].

4.1.2. Effect of the supplementation of media based on agro-industrial by-products with aromatic amino acids on the production of plasmid DNA

In the next phase, different approaches based on the addition of aromatic amino acids to growth medium were applied to optimize the biosynthesis of pDNA. The manipulation of the initial concentration of aromatic aminoacids can increase the pDNA production. The previous proposal model demonstrate that the influence of aromatic amino acids on their pathway provides necessary precursors for the nucleotides network [59].

Similarly, it was intended to increase production of pDNA by chemical manipulation of the metabolic pathways of *E. coli* VH35. To achieve this, the previously optimized media, based on agro-industrial by-products, were supplemented with aromatic amino acids such as tyrosine (tyr), phenylalanine (phe) and tryptophan (tryp). Tyr, phe and tryp were applied at concentrations of 10 µg/mL, 78.5 µg/mL and 100 µg/mL, respectively, according to previous results from our research group [59].

Two of the previously optimized media conditions were tested: the medium A and C. Aromatic amino acids were added to medium A in order to increase the pDNA production and avoid the addition of yeast extract in the medium, in order to lower the cost of the process [46]. Subsquently, the aromatic amino acids was added to medium C in order to improve the pDNA yield and purity. The results obtained for the pDNA volumetric and specific area and purity are depicted in the table 10.

Table 10- Effect of the supplementation of the media with aromatic amino acids on bacterial growth, plasmid DNA volumetric, specific yields, and purity.

Medium	*$OD_{600\ nm}$*	*Cell dry mass (CDM) (g/L)*	*pDNA volumetric yield (µg/mL)*	*pDNA specific yield (µg pDNA /mg CDM)*	*Purity (%)*
A + aromatic amino acids	4.955±0.075	2.230±0.143	6.019±0.227	2.699±0.143	0.337±0.023
C + aromatic amino acids	6.00±0.616	2.00±0.616	42.061±1.012	15.578±1.597	3.373±0.043

The maximum pDNA yield was obtaining with the medium C supplemented with aromatic amino acids, with values of pDNA volumetric and specific yield of 42.061±1,012 µg/mL and 15.578 ±1.597µg/CDM, respectively. Comparing to the previous results, the supplementation with aromatic amino acid led to an increased in both pDNA volumetric and specific yield, from 35,126±1,067 µg/mL and 11.803±0.700 µg/CDM, respectively, to 42.061±1.012 µg/mL and 15.578 ±1.597 µg/CDM.

Regarding the medium A supplemented with aromatic amino acids, values of pDNA volumetric and specific yield of 6.019±0.227 µg/mL and 2.699±0.143 µg/CDM were obtained. Although the addiction of aromatic amino acids to the medium A increased production of pDNA in about 2 µg/mL, the values of pDNA production yields and purify are still low when compared to the other media tested, especially the medium C with aromatic amino acids.

As referred above, the quality of the obtained plasmid is an important factor to be considered in pDNA production. Likewise, as the previous results, plasmid produced at the medium C supplemented with amino acid was predominantly in the sc form, but the medium A with amino acids produced less sc form.

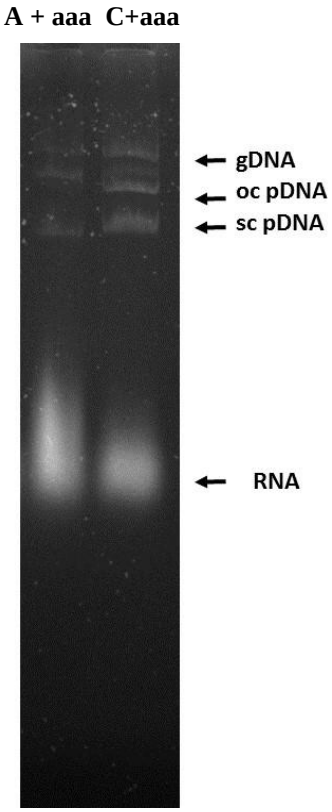


Figure 13 - Agarose gel electrophoresis of target plasmid DNA during fermentation with aromatic amino acids (aaa). Lane A and C are fermentations of indicated in table 4.

4.2. The influence of different carbon sources in the metabolism of *E. coli* VH35

E.coli is on the well-studied organism and one of the reasons for its choice as host is its capability to utilize a wide range of sugars and sugars alcohol as a carbon source [44]. In some cases, it seems to be a hierarchical utilization of sugars by *E. coli*, in which glucose is the most preferred carbon source. Only after glucose complete consumption, *E.coli* utilizes the secondary carbon sources. However, other reports indicate that, when cultivated in media with several carbon sources, *E. coli* frequently consumes these subtracts simultaneously [44,130].

In the case of VH35 strain, used in this work, hierarchical utilization of sugars may be changed due to the genetic mutation of the PTS, responsible for glucose transport across the cell membrane, in strain VH35. However, others galactose transporters, galactose permease are expressed and transport other sugars across the cell membrane [131,132]. Also, it was reported that the strain VH35 generally produces low acetate levels due its lack of PyK A activity, the enzyme responsible by the reduction PEP to PYR [37]. So, to fully understand the *E. coli* VH35 metabolism, its behavior in the presence of non-conventional media, such as agro-industrial extracts, should be evaluated.

In this work, a study was performed with the purpose of exploring the capacity of *E. coli* VH35 for producing pDNA from alternative carbon sources and verify the variation in the consume of sugars that occurs throughout the fermentation. For this purpose, the consumption of the sugars and the production the acetate throughout the fermentation were monitored by the analysis of glucose, lactose, galactose and acetate. For this study, the next medium was included for comparison with the media based on agro-food extracts [59].

The results for conventional fermentation medium are represented in the figure 14, in the presence of glucose, the cells experience a period of adjustment (4 hours), followed by an exponential growth period, in which glucose concentration decrease over time. As can be seen in figure xiii, glucose is a preferred carbon source, whereas the low galactose concentrations are maintained over time. Regarding acetate production, HPLC of measurements indicate that the strain VH35 produced a maximum of 2 g/L acetate in this assay.

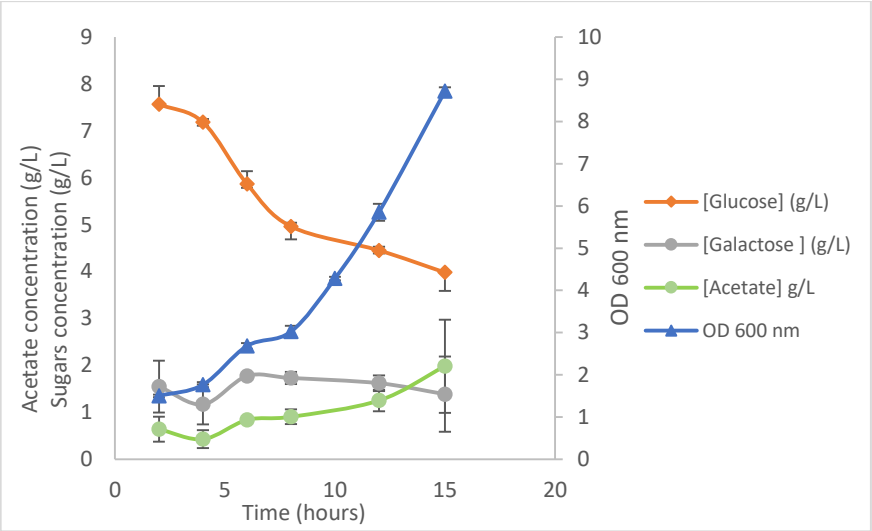


Figure 14 - Profile of optical density (600 nm), glucose, galactose consumption and acetate production in conventional fermentation medium.

However, the profile of the sugar consumption are quite different when the fermentation is performed with agro-industrial extracts, as seen in Figures 15 -21. According to literature, when glucose is limited since the beginning, the alternative sugar catabolic operon promoters is activated in the present of non – PTS sugars (α-lactose, L-arabinose, D-xilose). Then, the lactose become the preferential sugar consumed by *E.coli* [43,133].

The results suggested that *E.coli* VH 35 choose to consume lactose in the presence of an agro-food medium, because this medium do not contains PTS – sugars, such as glucose, as seen in figures 15-21. Lactose concentration decrease over time and reached the minimum level when the optical density stabilizes. This metabolic behavior of *E. coli* occurs in the media composed by agro-industrial extracts analyzed, including media supplemented with amino acids.

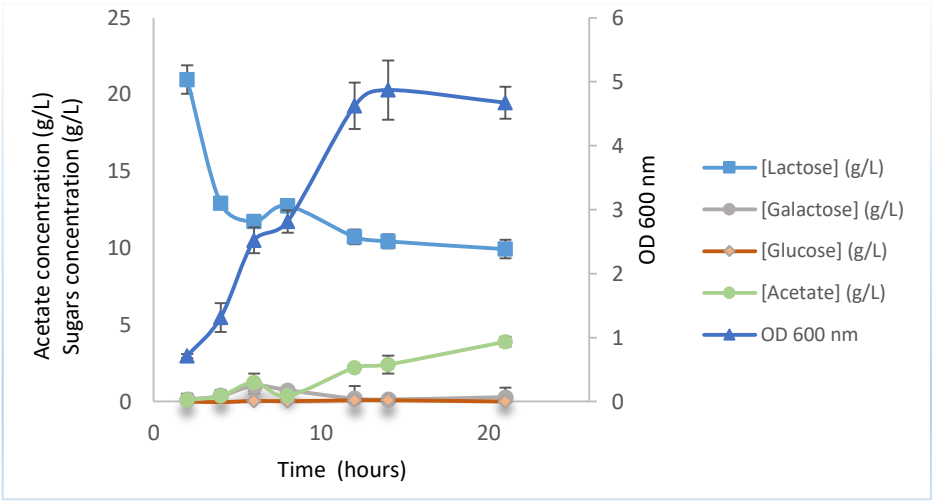


Figure 15 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation A [2% CSL (v/v) + 1.2 % CW (w/v)].

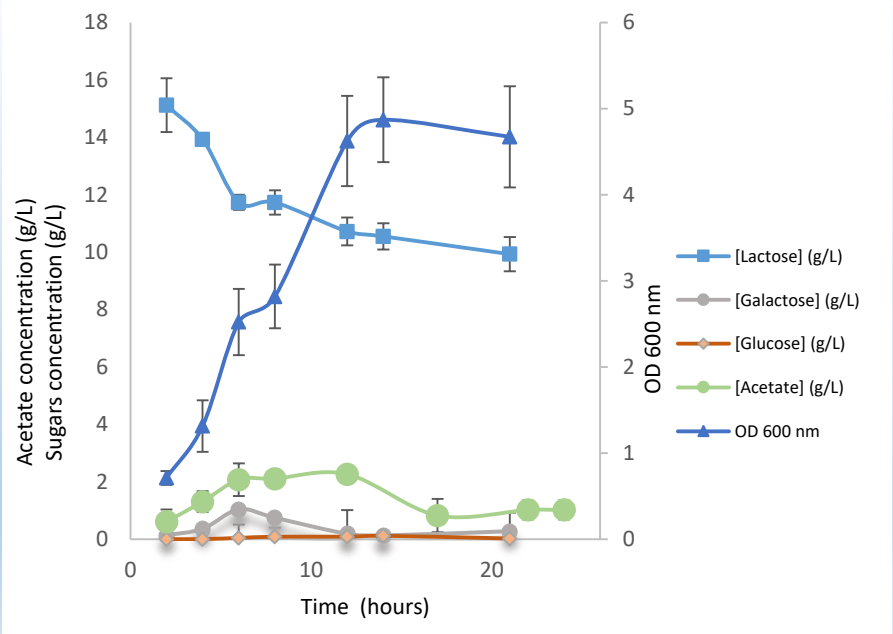


Figure 16 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation A with aromatic amino acids [2% CSL (v/v) + 1.2 % CW (w/v) with aromatic amino acids].

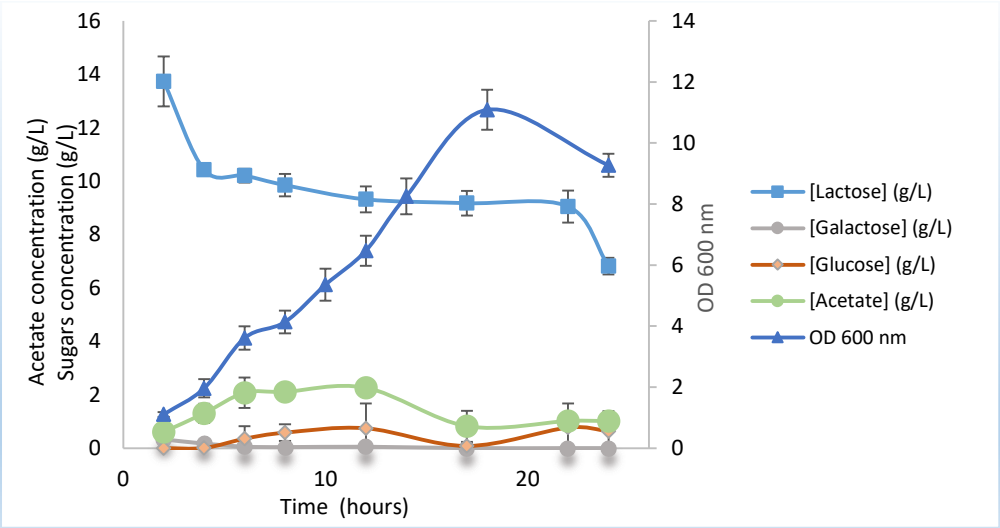


Figure 17 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation B [2% CSL (v/v) + 1.2 % CW (w/v) with 2.5 g/L YE].

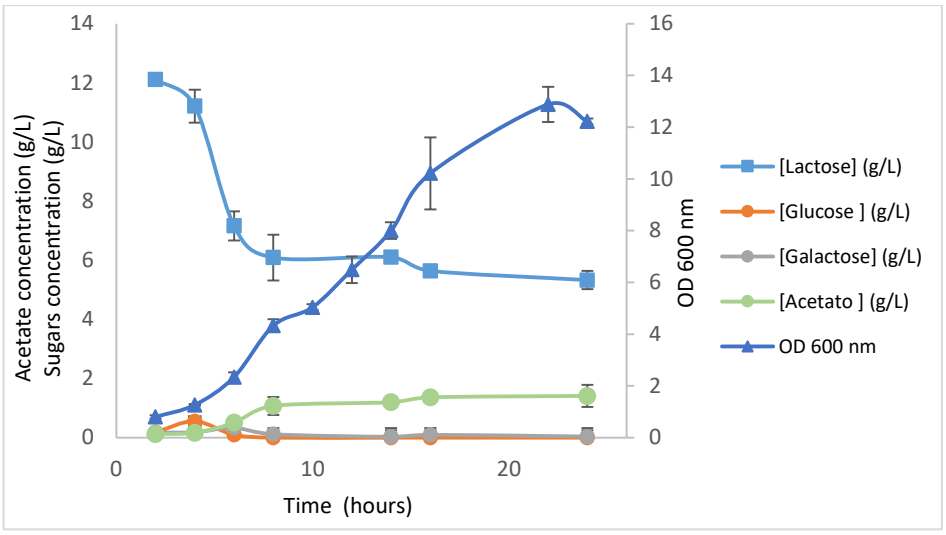


Figure 18- Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation C [2% CSL (v/v) + 1.2 % CW (w/v) with 5 g/L YE].

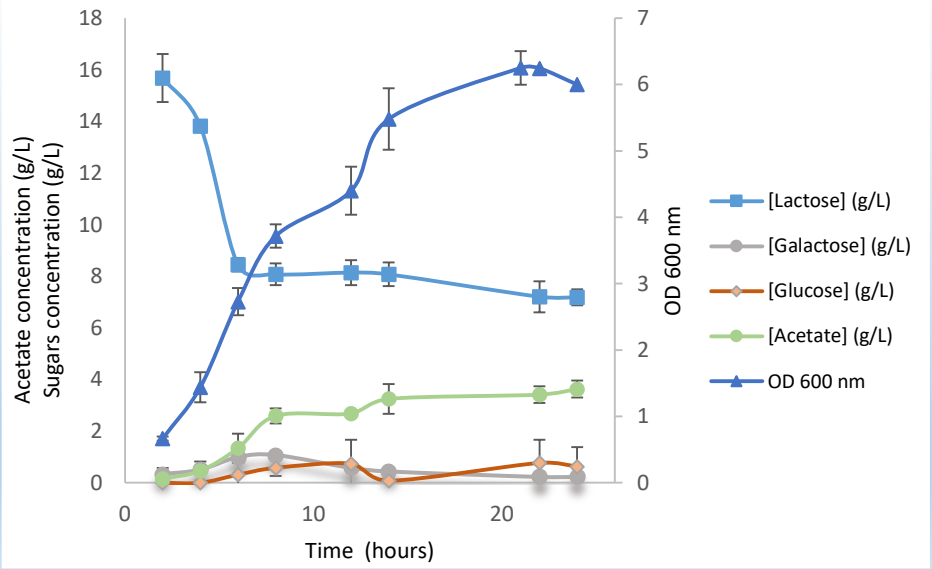


Figure 19 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation C with aromatic amino acids A [2% CSL (v/v) + 1.2 % CW (w/v) with 2.5 g/L YE and aromatic amino acids].

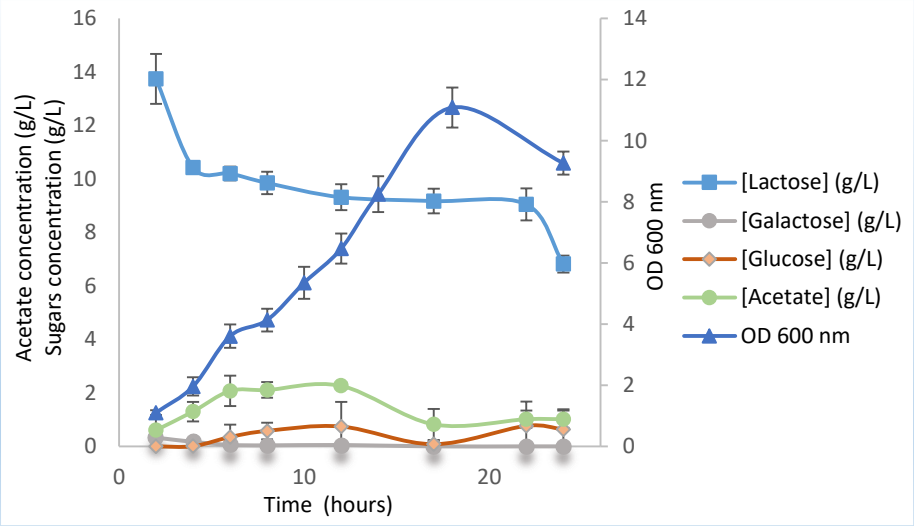


Figure 20 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation D [2% CSL (v/v) + 1.2 % CW (w/v) with 7.5 g/L YE].

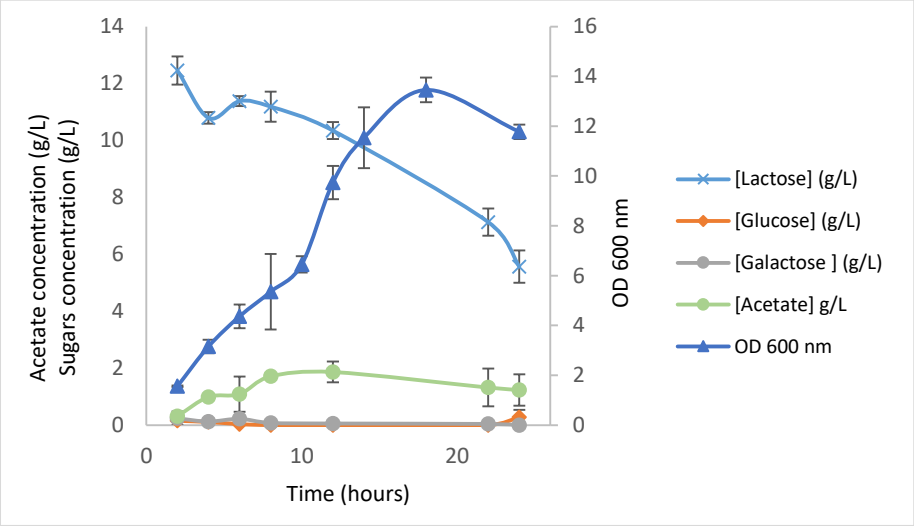


Figure 21 - Profile of optical density (600 nm), glucose, lactose, galactose consumption and acetate production in fermentation E [2% CSL (v/v) + 1.2 % CW (w/v) with 10 g/L YE].

Also, all of lactose molecules should be metabolized to glucose and galactose during fermentation by ß-galactosidase [134] to enhance its consume by the strain VH35, simultaneously with lactose utilization. This occurs because in strain VH35 a higher flux is directed from central metabolism to biomass formation [37]. As can see in the figures 15-21 the lactose concentration decreases over time but the concentration of glucose and galactose does not increase as the consumption of lactose.

Acetate is the by-product most predominantly formed in *E.coli* under aerobic conditions, causing several adverse effects on physiology as inhibition of cell growth and reduction of pDNA production [59,127]. HPLC measurements indicate that the strain VH35 produced a maximum of 3,8 g/L of acetate in the medium C supplemented with amino acids (CSL 2% ; 1,2% CH ; 5,0

(g/L) YE with aaa). So, the level the acetate produced in the presence of this specific agro-food medium is lower than the acetate concentration 5 g/L, that could lead to a reduced growth rate and biomass yields [18].

As can be seen in figure 17, 18, 20, *E.coli* OD_{600}, was very high in the fermentation in the presence of agro-industrial media (2% CSL and 1.2 % CW) supplemented with addiction of 2.5, 5.0, 7.5 and 10 g/L YE. However, these measurements may have been affected by the formation, over fermentation time, of some precipitates in the medium growth, leading to some turbidity.

Sucintly, this results demonstrated that strain VH35 may be an efficient strain to be used in an ecologically and economically sustainable procedure for pDNA production. This work proved that strain VH35 uses agro-food media as an energy source, when insufficient glucose levels are available.

4.3. Proteomic analysis

In *E.coli*, protein expression are result of a compromise between metabolic pathways used for the expression of pDNA and its growth [123]. The aim of proteomics analysis in this work was to provide a comprehensive description of proteome changes of *E.coli* VH35 that occurs in their metabolic adaptation to different conventional and non-conventional growth media. Also a systematic study of the effect of agro-food medium on the proteome E. *coli* were carried out, in order to elucidate some effects of this extracts in its metabolic pathway. Proteomic approaches combining 2DE and MALDI TOF/TOF were applied to identify proteins overexpressed in the proteome of *E. coli* cells cultivated in distinct culture media.

In a first approach, 2DE procedure for the analysis of *E. coli* lysates was optimized in terms of sample preparation and electrophoresis conditions to improve the reproducibility of gel –based proteomics assays. Firstly, the effect of centrifugation of *E. coli* lysates were studied in terms of protein recovery and the number and volume of spots in 2DE gels. So, centrifuged and non-centrifuged *E. coli* lysates were quantified and 250 µg of protein were analyzed by 2DE on a 3-10 pH range strip. Thereafter, using the optimized conditions for sample preparation, the pH ranges 3-10 and 4-7 were tested in order to improve the resolution of gels [78,135].

The results demonstrate that the centrifugation of *E. coli* lysates leads to a decrease in protein concentration, from 5.29 ± 0.37 µg/µL in non-centrifuged samples to 0.64 ± 0,063 µg/µL in centrifuged lysates. By 2DE gel analysis, 419 spots were detected in the gel of non-centrifuged samples, corresponding to a volume of 416.29 cm^3, but in -centrifuged samples only 195 spots were detected, with a total volume of 166.07 cm^3 (figure 22).

Comparing the 3-10 and 4-7 pH range strips, 195 spots with a total volume of 111.04 cm^3 were detected in 3-10 pH range and 305 spots with a total volume of 180, 30 cm^3 were detected in 4-7 pH range, as seen in figure 23.

So, the centrifugation of *E. coli* lysates leads to significant protein losses and this fact is evident when comparing the 2DE gels from non and centrifuged samples. Furthermore, 3-10 pH resulting gels had low resolution and lower number of spots when compared to the 4-7 pH range. Thus, the further experiments were conducted using non-centrifuged lysates and a pH linear gradient of 4–7, providing an improved gel resolution and a higher number of spots.

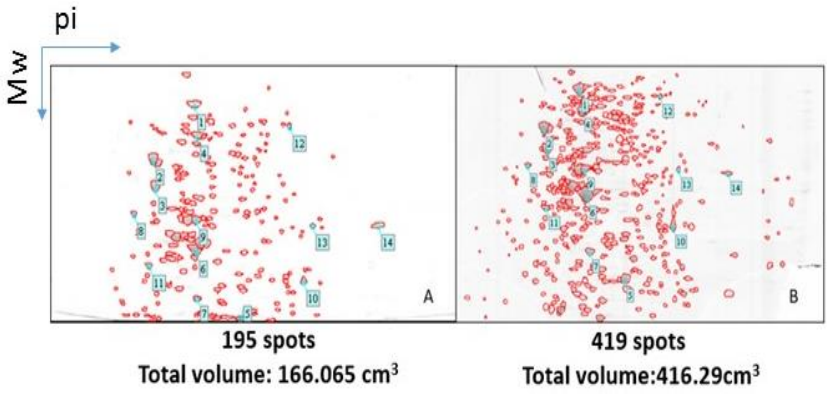


*Figure 22- Centrifuged **(A)** and non-centrifuged **(B)** samples were analyzed by 2DE using pH 3-10 linear IPG strips, 24 cm*

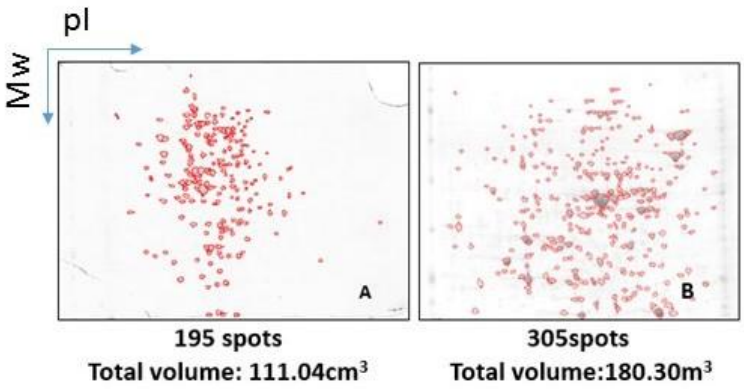


Figure 23 -Samples analyzed by 2DE using (A) 3-10 pH range, 24 cm linear IPG strips and (B) 4-7 pH range, 18 cm linear IPG strips.

After the optimization of 2DE procedure, *E.coli* strain VH35 was growth in different growth media, whose composition is described in table 11. Medium 1 was include for comparison with agro-food media (medium 2; 3; 4; 5). Medium 1 is a result of experimental design optimization performed by our research group for the biosynthesis of plasmid DNA by *E. coli* VH33, where a high production of pDNA was obtained [59].

Table 11 - Growth conditions for proteomic analysis (growth time = 15 hours)

MEDIUM	*MEDIA COMPONENTS*
1	20 g/l *YE*, 20 g/l glucose, 24 g/l triptone and 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan
2	CSL 2% (v/v)+ 1.2% *CH* (w/v)
3	*CSL* 2% (v/v)+ 1.2% *CH* (w/v)+ + 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan
4	*CSL* 2% (v/v)+ 1.2% *CH* (w/v)+5 (g/L) YE
5	*CSL* 2% (v/v)+ 1.2% *CH* (w/v)+5 (g/l) *YE*+ 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan

The success of proteomic is largely based on the appropriate preparation of the protein sample [136]. Protein precipitation is part of the sample preparation procedure and several techniques such as methanol/chloroform, trichloroacetic acid, trichloroacetic acid/acetone, and acetone precipitation, have been applied [81,92,137]. This step is important to remove proteases, cellular material, genomic DNA and interfering lipids and detergents [137,138].

In this work, the *E. coli* lysates were prepared by methanol/chloroform precipitation because of its efficiency in recover a variety of proteins, especially membrane proteins, and remove both lipophilic and hydrophilic interferents [138]. To evaluate the applicability of methanol/chloroform precipitation for *E.coli* VH35 studies, protein concentration in lysates was compared before and after precipitation. Precipitation of protein from the different media with methanol/chloroform procedure provided a recovery of 77.426% in medium 1, 87.675% in medium 2, 73.563% in medium 3, 71.746% in medium 4 and 89.560 % in medium 5. The applied method allows to precipitate the *E. coli* proteins without significant protein losses, since average percentage of recovery is high, 79.994 ± 7.301%.

Table 12 - Percentage of protein recovery after the precipitation procedure. Values are the mean ± standard of four independent experiments.

Growth Medium	*Protein amount before precipitation (µg)*	*Protein amount after precipitation (µg)*	*Percentage of recovery (%)*
1	2209.72 ± 266.153	1710 .91 ± 65.385	77.426
2	1148.96 ± 102.576	1007.91 ± 86.486	87.675
3	1628.75 ± 217.962	1198.16 ± 154.0	73.563
4	1745.83 ± 145.962	1252.57 ± 23.1618	71.746
5	1442.50 ± 145.893	1231.91 ± 65.325	89.560

This analysis also enables the discovery of the differences in protein concentration of lysates between all the growth media. By analysis of table 12, differences in total mass can be observed in the medium 1 in comparison with other media (agro- food medium), for example, from 2209.72±266.153 µg in media 1 to 1148.96±102.576 µg in media 3.

After 2DE analysis of *E.coli* precipitates, the total number and volume of spots were compared between media using ImageMaster 2D Platinum v7.0, as represented in table 13. The values in the table 13 are representative of the four independent 2DE experiments performed for the analysis of proteome of *E.coli* VH35 cultured in different media (figure 24). Significant differences in total volume was found between media 3 with media 1, 2, 4 and 5. Induced in medium 3, higher number of protein spots were detected in gels, 780.5 ±37.766, corresponding to a total volume of 235.233 ± 73.360 cm^3.

Also, there is a significant difference between the 2DE gels of *E. coli* cells grown on medium B, on the same medium supplemented with aromatic amino acids (medium 3) and on the medium 5 with amino acids and YE. The number of spots increases from 526.5 ± 37.766 spots in medium 2 to 780.5± 86.685 spots in medium 3, to 549.5 ± 84.485 in medium 4 and to 642.75 ± 75.965 in medium 5.

Table 13- Number of protein spots and total volume for the various conditions tested. Values are the mean ± standard deviation of four independent experiments.

Growth Media	***Number of spots***	***Total volume (cm^3)***
1	522.251 ± 66.170	130.073 ± 28.836
2	526.5 ± 37.766	141.933 ± 32.563
3	780.5 ± 86.685	235.233 ± 73.360
4	549.5 ± 84.485	110.973 ± 26.394
5	642.75 ± 75,965	121.9331 ± 2.625

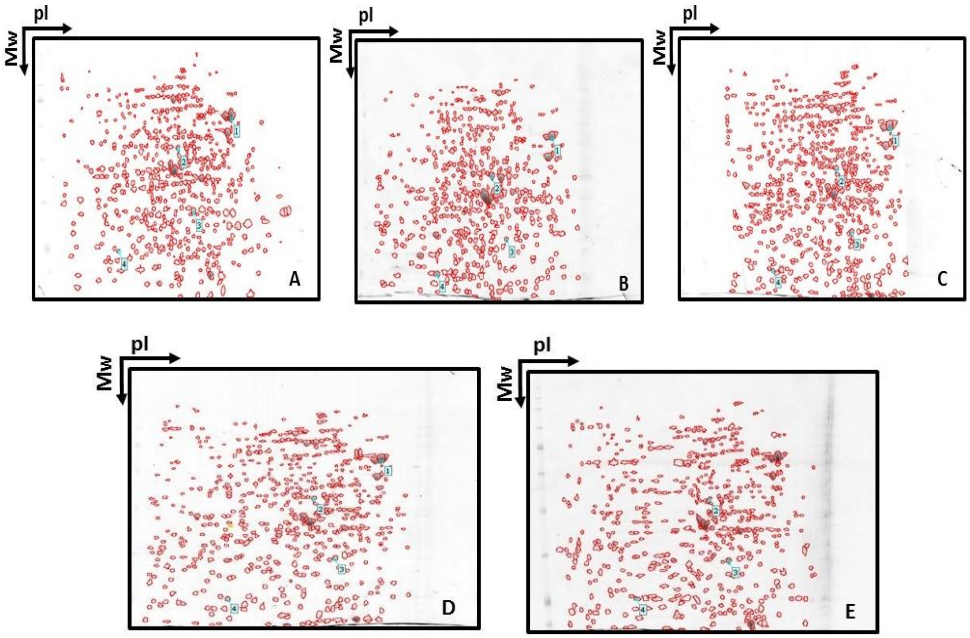


Figure 24 - 2DE analysis of proteins from E.coli VH35 after fermentation with different growth media: (A) Medium 1, (B) Medium 2, (C) Medium 3, (D) Medium4 and (E) medium 5 according to table 12. Samples were analysed using pH 4-7 linear IPG strips, 18 cm.

4.3.1. Proteomic Profile of *E. coli* VH35 grown in different culture media

Under environmental variations such as oxygen limitation, nutrient availability, temperature, pH, and other factors, the organism suffered physiological changes which were triggered by alterations in gene and protein expression. So, the use of different culture media may cause alterations in the central carbon and other metabolic pathways of *E. coli* [78,120].

Protein samples extracted from *E.coli* strain K12, separated by 2DE, and analyzed by MALDI-TOF/TOF for protein identification, revealed the versatile and robust metabolism of *E. coli*, capable to adapt to environmental alterations. Indeed, many enzymes were found to be regulated by different mechanisms [78]. Also, another study demonstrates that the cells utilize alternative carbon and energy sources in response to phosphate limitation [121].

In order to investigate metabolic responses under different growth conditions and related proteome changes, proteomic profile of the *E.coli* strain VH35 grown in different media were compared. Gels obtained from 2DE were subjected to statistical analysis using the software ImageMaster 2D Platinum v7.0, a software which applies t-test (ANOVA), with p-values ≤ 0.05, in order to identify the proteins with significant alterations regarding their expression levels.

Firstly, all images were pooled together according to their classes (culture medium), taking always the same image as reference. Afterwards, the spots presented in all 2DE gels were manually selected as landmarks in order to properly align the images. Then, in order to find proteins with significant expression changes, only spots matching all classes under study were

considered, resulting in the following comparative analysis of the 2DE maps: medium 1 - 5, medium 3 - 5, medium 2 - 3, medium 2 - 4, and medium 3 - 4. This analysis was performed in order to study the effect of using agro-industrial extracts with yeast extract and with aromatic amino acids as medium.

4.3.1.1. Effect of supplementation of agro-industrial extracts with yeast extract in E. coli VH35 proteome by 2DE analysis

To increase pDNA production yields, the agro-industrial extracts were supplemented with YE, and it was observed that the initial growth medium (medium 2) supplemented with YE, increased the production of pDNA from 4.379 ± 0.096 µg/mL to 35.126 ± 1.067 µg/mL. This increase is due to alterations in *E.coli* metabolic pathways, and for this reason, it is important to study its proteome under these conditions. So, the medium 4 supplemented with YE, was compared to medium 2. By 2DE analysis, a total of 386 spots were identified across the gels and, from these 386 proteins, 54 were found differentially expressed. To understand which ones may be involved in the *E. coli* metabolism, the differentially expressed spots were identified by MS. In this condition, the majority of differentially expressed proteins were found in the basic area, representing high and intermediate molecular weight.

After MALDI-TOF/TOF analysis results, MS/MS peaks were generated and crossed with UniProt *Escherichia coli* database, through ProteinPilot 4.5 using Paragon™ Algorithm. Using this strategy, we were able to identify 6 differentially expressed proteins, under a 95% confidence, between the medium 2 and medium 4 which are listed in table 14. In order to obtain more information about the sunergy of these proteins with biological processes, it was performed an analysis regarding these proteins using STRAP 1.5.5 software (Software Tool for Rapid Annotation of Proteins). This program generates several tables and graphs for biological process, cellular component and molecular function of the analyzed proteins, as shown below. Examining this information, it was possible to conclude that from these 6 proteins, 2 are involved in metabolic processes (chaperone protein, galactokinase) and 1 is involved in regulation (Rod shape-determining protein), response to stimulus (chaperone protein) and cellular processes (aspartate aminotransferase). Regarding to molecular functions, 4 proteins (chaperone protein, aspartate aminotransferase, elongation factor G, elongation factor, galactokinase) have catalytic activity and 5 proteins have binding activity, as showed in figure 25.

Table 14 - Identified proteins by MALDI-TOF/TOF between the medium 2 and 4.

Match	*Normalize volume(%) Medium 2*	*Normalize volume(%) Medium 4*	*Protein Description*	*Function*
121	0.0923	0.009	Chaperone protein ClpB P63285	Part of a stress-induced multi-chaperone system, it is involved in the recovery of the cell from heat-induced damage, in cooperation with DnaK, DnaJ and GrpE. Acts before DnaK, in the processing of protein aggregates. Protein binding stimulates the ATPase activity; ATP hydrolysis unfolds the denatured protein aggregates, which probably helps expose new hydrophobic binding sites on the surface of ClpB-bound aggregates, contributing to the solubilization and refolding of denatured protein aggregates by DnaK (By similarity).
138	0.465	0.242	Aspartate aminotransferase P00509	The enzyme catalyses the reaction: L-aspartate + 2-oxoglutarate = oxaloacetate + L-glutamate Aminotransferases share certain mechanistic features with other pyridoxal-phosphate-dependent enzymes, such as the covalent binding of the pyridoxal-phosphate group to a lysine residue.
284	0.612	2.152	Elongation factor G B7UK50	Catalyzes the GTP-dependent ribosomal translocation step during translation elongation. During this step, the ribosome changes from the pre-translocational (PRE) to the post-translocational (POST) state as the newly formed A-site-bound peptidyl-tRNA and P-site-bound deacylated tRNA move to the P and E sites, respectively. Catalyzes the

				coordinated movement of the two tRNA molecules, the mRNA and conformational changes in the ribosome.
115	0.033	0.0758	Elongation factor Tu 1 A7ZSL4	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis.
172	0.992	0.796	Galactokinase A7ZJD2	Catalyzes the transfer of the gamma-phosphate of ATP to D-galactose to form alpha-D-galactose-1-phosphate (Gal-1-P).
107	0.017	0.077	Rod shape-determining protein MreB P0A9X5	Involved in formation of the rod shape of the cell. May act as a negative regulator of FtsI (By similarity).

To perform a simpler revision of the given information about these 6 proteins, STRAP gives us access to pie charts concerning GO terms, as showed in figure 25.

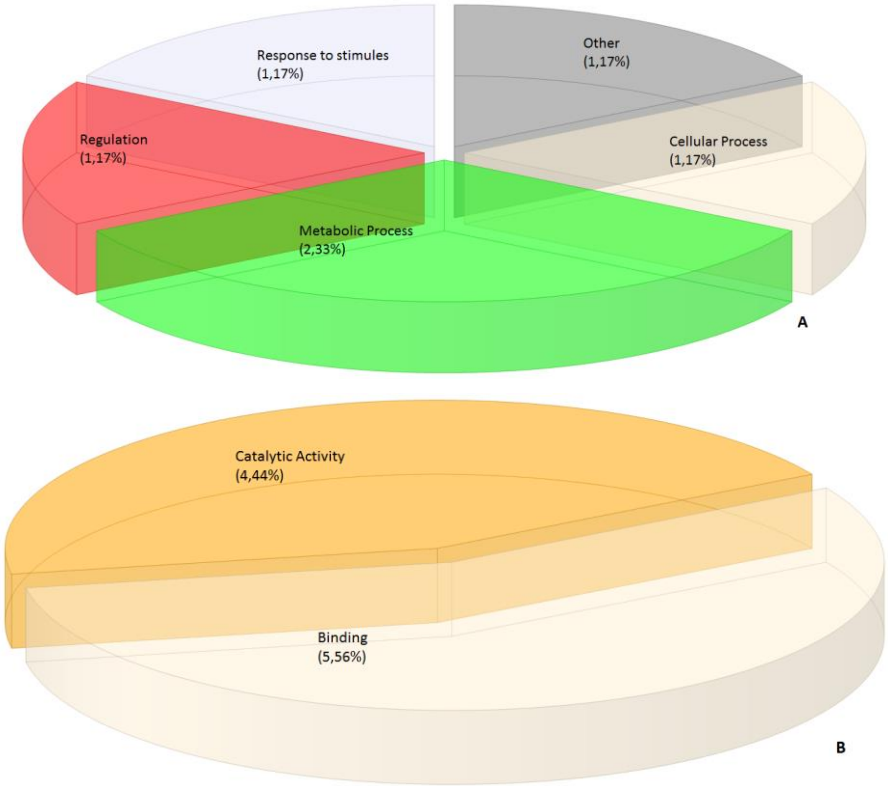


Figure 25 - Pie charts generated by STRAP software showing a resume of the predominance of several factors in the 6 studied proteins. A- Biological Process and B- Molecular Function according to the number of the associated proteins.

For a more profound understanding of these differentially expressed proteins, STRING 10 (Search Tool for the Retrieval of Interacting Genes/Proteins), which is a database of known and predicted protein-protein interactions, was used. This software is a tool that allows the establishment of interactions between different proteins and leads to the conclusion about the function and activity of the analyzed proteins. These proteins were searched through STRING 10 and the protein-protein associations were predicted using medium confidence.

Thus, it was possible to conclude that after 15 hours, the Rod shape-determining protein MreB, a protein associated with cellular division process, is overexpressed in medium 4. On the other hand, the Chaperone protein ClpB, a protein associated to a stress-induced multi-chaperone system, is involved in the recovery of the cell from heat-induced damage, in cooperation with DnaK, DnaJ and GrpE, and is underexpressed in medium 4. It was also possible to identify two proteins (elongation factor G and elongation factor Tu 1) overexpressed in medium 4 comparing to medium 2, which are associated with translation and gene expression processes. The protein aspartate aminotransferase, that was found overexpressed in medium 2, is associated to cellular amino acid metabolic process, so it is possible to conclude that, during that time, the occurrence of amino acids synthesis is more significant in medium 2 that in medium 4. However, it was verified a decrease in pDNA production in medium 2 opposite to medium 4.

For further analysis, the possible effect of YE in a more complex growth medium, supplemented with addition of amino acids, was studied. To achieve this purpose, the growing medium 3 was compared to medium 5, and in this case it was verified an increase in pDNA of 6,019 ± 0,227 µg/mL for 42,061 ± 1,012 µg/mL, respectively. Results demonstrated that when yeast extract was used in the medium containing CW 1,2% (w/v) and CSL 2% (v/v) with aromatic amino acids, 80 differentially expressed protein spots were identified. Using ProteinPilot 4.5 software, 6 differentially expressed proteins were identified under a 95% confidence, between the medium 3 and medium 5, as it can be seen in table 15. These proteins were also analyzed using STRAP 1.5 software, in order to know in which mechanisms they are involved.

Table 15 - Identified proteins by MALDI-TOF/TOF between the medium 3 and 5

Match	***Normalize volume (%) Medium 3***	***Normalize volume (%) Medium 5***	***Protein Description***	***Function***
490	0.023	0.208	Catalase HPII P21179	Decomposes hydrogen peroxide into water and oxygen; serves to protect cells from the toxic effects of hydrogen *per*oxide.
243	0.375	0.194	S-adenosylmethionine synthase B7UHY9	Catalyzes the formation of S-adenosylmethionine from methionine and ATP. The overall synthetic reaction is composed of two sequential steps, AdoMet formation and the subsequent tripolyphosphate hydrolysis, which occurs prior to release of AdoMet from the enzyme.
264	0.043	0.006	Outer membrane protein A P0A911	Required for the action of colicins K and L and for the stabilization of mating aggregates in conjugation. Serves as a receptor for a number of T-even like phages. Also acts as a porin with low permeability that allows slow penetration of small solutes (By similarity).
141	0.009	0.116	Pyruvate dehydrogenase E1 component P0AFG9	Component of the pyruvate dehydrogenase (PDH) complex that catalyzes the overall conversion of pyruvate to acetyl-CoA and CO_2.
328	0.020	0.213	Anthranilate synthase componente 1 P00895	Part of a heterotetrameric complex that catalyzes the two-step biosynthesis of anthranilate, an intermediate in the biosynthesis of L-tryptophan. In the first step, the glutamine-binding beta subunit (TrpG) of anthranilate synthase (AS) provides the glutamine

				amidotransferase activity which generates ammonia as a substrate that, along with chorismate, is used in the second step, catalyzed by the large alpha subunit of AS (TrpE) to produce anthranilate. In the absence of TrpG, TrpE can synthesize anthranilate directly from chorismate and high concentrations of ammonia.
234	0.007	0.606	Elongation factor Tu 1 A7ZSL4	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis.

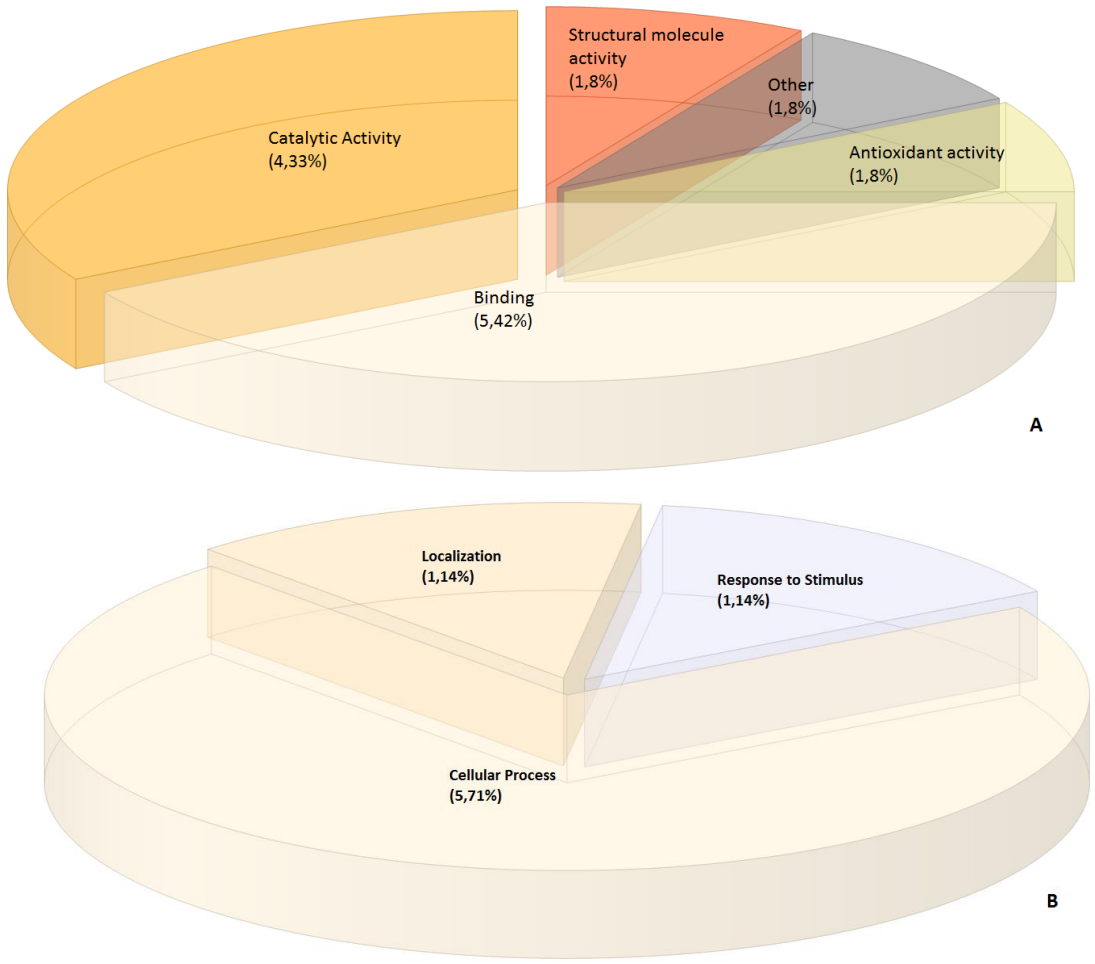


Figure 26 - Pie charts generated by STRAP software showing a resume of the predominance of several factors in the 6 studied proteins (medium 3 and medium 5). A- Biological Process and B- Molecular Function according to the number of the associated proteins.

According to these pie charts we can verify that, regarding to biological processes 5 proteins are involved in cellular processes (catalase HPII, s-adenosylmethionine synthase, outer membrane protein A, pyruvate dehydrogenase E1 component, anthranilate synthase component 1), 1 is involved in response to stimulus and 1 (outer membrane protein A) is involved in localization.

Regarding to molecular functions 4 proteins have catalytic activity (catalase HPII, s-adenosylmethionine synthase, outer membrane protein A, pyruvate dehydrogenase E1 component, anthranilate synthase component 1, elongation factor Tu 1) and 1 protein have structural activity (outer membrane protein A), 5 proteins have a binding activity (catalase HPII, s-adenosylmethionine synthase, pyruvate dehydrogenase E1 component, anthranilate synthase component 1, elongation factor Tu 1).

Using the software STRING 10, it was possible to conclude that the proteins overexpressed in medium 3 are involved in metabolic pathways and in the biosynthesis of secondary metabolism (S-adenosylmethionine synthase), and also, outer membrane protein A is involved in transport across cell membrane associated to ampF and ompC, and it has activity of purine. The proteins catalase HPII, pyruvate dehydrogenase E1 component, anthranilate synthase component 1 and elongation factor Tu 1, are overexpressed in the medium 5. Catalase HPII has the function to response to oxidative stress, and it participates in tryptophan metabolism. Pyruvate dehydrogenase E1 component participates in gluconeogenesis in order to generate glucose. Anthranilate synthase component 1 is involved in biosynthesis of amino acids (phenylalanine, tryptophan and tyrosine), whereas Elongation factor Tu 1 is associated with gene translation and expression processes, thus elucidating the significative differences when it comes to p53 plasmid production.

4.3.1.2. Effect of the use of supplementation of agro-industrial extracts with aromatic amino acids in *E. coli* VH35 proteome by 2DE analysis

The supplementation with amino acids, in the literature, is described to enhance the transport systems, as well as to increase the growth of cells [59,139]. So, aromatic amino acids were added to agro-industrial medium 2 in order to improve pDNA yield and purity. Despite the increase in pDNA production was not significant, we compared the proteome of *E. coli* grown in medium 3 (with amino acids) and the medium 2 (without amino acids) in order to identify differentially expressed proteins that could be involved in the aromatic amino acids metabolism.

Results demonstrate that a total of 386 spots were identified across gels and from these proteins, 23 were found differentially expressed. Identified as differentially expressed proteins

are either in area of high and low molecular weight but in basic area, a protein is notable for meeting in the low molecular weight and acid area.

After MALDI-TOF/TOF analysis, we were able to identify 5 differentially expressed proteins, under 95% confidence, between the medium 2 and medium 3, as it is showed in table 16. Using STRAP 1.5 software it was possible to conclude that regarding to biological processes 1 protein is involved in metabolic process (2,3-bisphosphoglycerate-independent phosphoglycerate mutase), 1 is involved in localization (trigger factor) and 3 are involved in cellular process (trigger factor, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, anthranilate synthase component 1). Regarding to their molecular functions, 4 proteins have a catalytic activity (trigger factor, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, anthranilate synthase component 1) and 3 proteins have a binding activity (elongation factor Tu 1, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase, anthranilate synthase component 1), as it can be seen in figure 27.

Table 16 - Proteins Identified by MALDI-TOF/TOF between the medium 2 and 3.

Match	***Normalize volume (%) Medium 2***	***Normalize volume (%) Medium 3***	***Protein Description***	***Function***
120	0.005	0.076	Trigger factor A1A8A5	Involved in protein export. Acts as a chaperone by maintaining the newly synthesized protein in an open conformation. Functions as a peptidyl-prolyl cis-trans isomerase.
140	0.008	0.075	Elongation factor Tu 1 A7ZSL4	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis.
262	0.013	0.235	2,3-bisphosphoglycerate-independent phosphoglycerate mutase A7ZTG6	Catalyzes the interconversion of 2-phosphoglycerate and 3-phosphoglycerate.

265	0.213	0.048	Anthranilate synthase component 1 P00895	Part of a heterotetrameric complex that catalyzes the two-step biosynthesis of anthranilate, an intermediate in the biosynthesis of L-tryptophan. In the first step, the glutamine-binding beta subunit (TrpG) of anthranilate synthase (AS) provides the glutamine amidotransferase activity which generates ammonia as a substrate that, along with chorismate, is used in the second step, catalyzed by the large alpha subunit of AS (TrpE) to produce anthranilate. In the absence of TrpG, TrpE can synthesize anthranilate directly from chorismate and high concentrations of ammonia.
50	0.0823264	0.020314	Rod shape-determining protein MreB P0A9X5	Involved in formation of the rod shape of the cell. May act as a negative regulator of FtsI (By similarity).

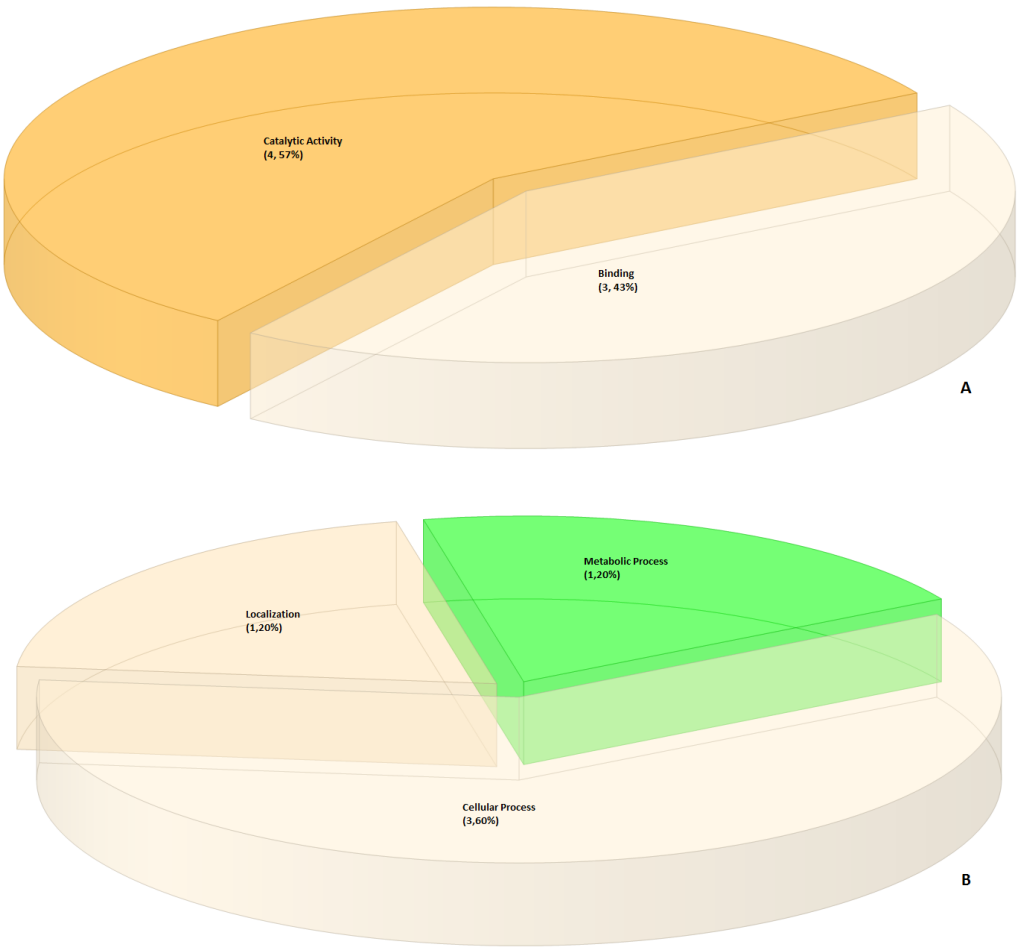


Figure 27- Pie charts generated by STRAP software showing a resume of the predominance of several factors in the 5 studied proteins (medium 2 and medium 3). A- Biological Process and B-Molecular Function according to the number of the associated proteins.

Using the software STRING 10 it was possible to conclude that the proteins overexpressed in medium 3 are involved in translation and in organonitrogen compound biosynthetic process (Trigger factor). They are also associated with translation and gene expression processes (elongation factor Tu 1), and involved in carbohydrate degradation as a cofactor of Mg^{2+} (2,3-bisphosphoglycerate-independent phosphoglycerate mutase). The proteins anthranilate synthase component 1 and rod shape-determining protein MreB after 15 hours of fermentation were found overexpressed in the medium 2, and are related to the biosynthesis of amino acids (phenylalanine, tryptophan and tyrosine) and involved in peptidoglycan biosynthesis, participating in the cell cycle and division, respectively.

The effect of the supplementation with amino acids was also verified by the increase of the production of pDNA from 35.126 ± 1.067 µg/mL, in the medium 4, to 42.061 ± 1.012µg/mL in the medium supplemented with aromatic amino acids (medium 5). So, to identify differentially expressed proteins by adding aromatic amino acids to the medium, the *E. coli* proteome cells

grown in medium 4 was compared with medium 5. 2DE analysis demonstrated that the supplementation of the growth medium with amino acids, resulted in the identification of 23 differentially expressed protein spots, from a total of 251 matching protein spots. The differentially expressed proteins are either in the area of high and low molecular weight and acid and basic area of the gel. Three differentially expressed proteins were found under a 95% confidence, between the medium 4 and medium 5, as it can be seen in table 17.

Table 17 – Proteins identified by MALDI-TOF/TOF between the medium 4 and 5.

Match	***Normalize volume (%) Medium 4***	***Normalize volume (%) Medium 5***	***Protein Description***	***Function***
22	0.203	0.070	Deoxyribose-phosphate aldolase B7NH49	Catalyzes a reversible aldol reaction between acetaldehyde and D-glyceraldehyde 3-phosphate to generate 2-deoxy-D-ribose 5-phosphate.
82	0.016	0.307	Galactokinase A7ZJD2	Catalyzes the transfer of the gamma-phosphate of ATP to D-galactose to form alpha-D-galactose-1-phosphate (Gal-1-P).
119	0.062	0.022	2,3-bisphosphoglycerate-independent phosphoglycerate mutase A7ZTG6	Catalyzes the interconversion of 2-phosphoglycerate and 3-phosphoglycerate.

To perform a simpler revision of the given information about these 3 proteins, STRAP gives us access to pie charts concerning GO terms, as showed in figure 28. According to these pie charts we can verify that, regarding to biological processes, 2 proteins are involved in cellular processes (deoxyribose-phosphate aldolase, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase), and 3 is involved in metabolic process (galactokinase, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase). Regarding to molecular functions, 1 protein has catalytic activity (deoxyribose-phosphate aldolase, galactokinase, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase) and 2 proteins have a binding activity(galactokinase, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase).

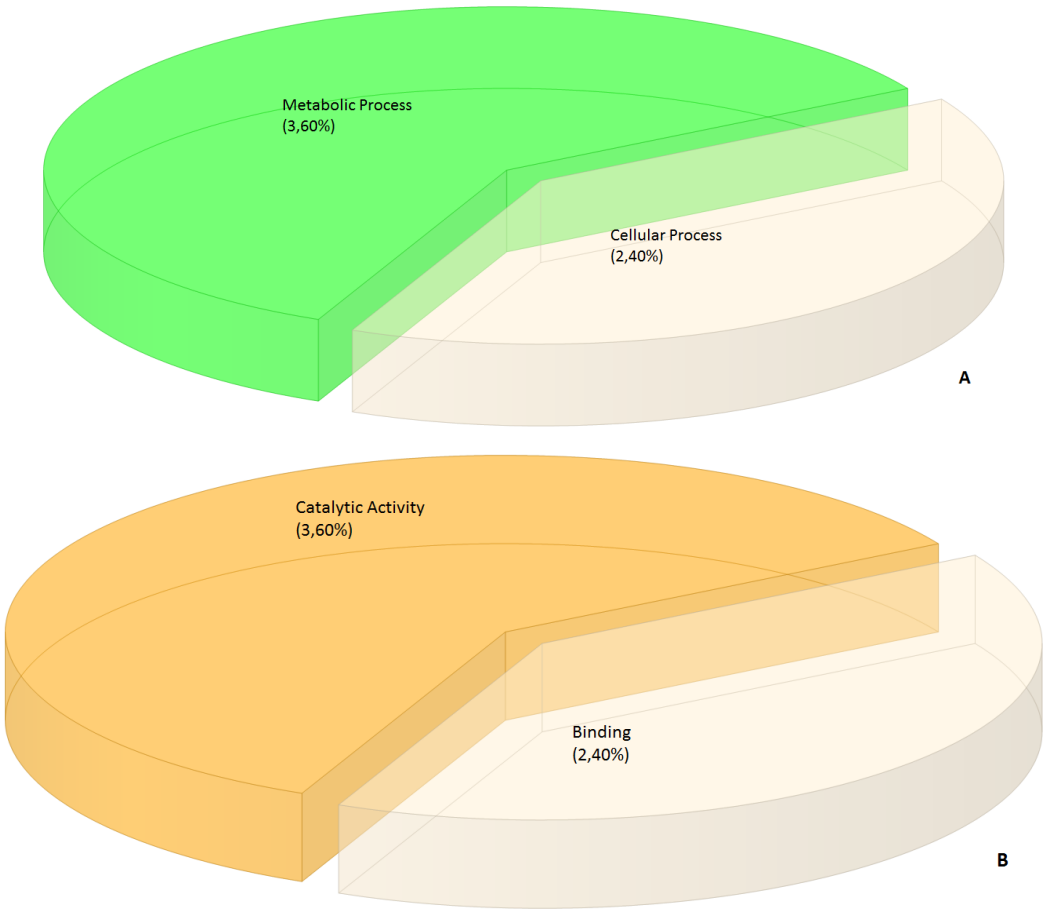


Figure 28 - Pie charts generated by STRAP software showing a resume of the predominance of several factors in the 3 studied proteins (medium 4 and medium 5). A- Biological Process and B-Molecular Function according to the number of the associated proteins.

Using the software STRING 10, it was possible to conclude that the proteins that are overexpressed in medium 4 are involved in carbon metabolism namely in carbohydrate degradation and amino acid biosynthesis (deoxyribose-phosphate and aldolase2,3-bisphosphoglycerate-independent phosphoglycerate mutase). In medium 5, the protein galactokinase is expressed at 15 hours and participates in carbohydrate process as well as indicates the presence of lactose transcription factors.

4.3.1.3. Study of the effect of the use of agro-industrial extracts in *E coli* VH35 proteome by 2DE analysis

Finally, the effect of the use of agro-industrial extracts as culture media on the *E.coli* VH35 proteome was studied. For this purpose, the conventional medium (medium 1), previously optimized by our research group, that presents higher pDNA yields, was compared with the agro-industrial medium 5, which also has a good performance in pDNA production. The medium 5 was described before as the medium with the highest production, when using agro-food substrates as growing medium, obtaining 42,061 ± 1,012 µg/mL. Results demonstrate that when

agro-wastes were used as the sole carbon source, a total of 519 spots were identified across gels, and from these, only 52 proteins were differentially expressed. This comparison reveled that the differentially expressed proteins are within the low molecular weight of the gel. Comparing the mediums 1 and 5, it was not possible to identify differentially expressed proteins under a confidence of over 95%. Table 18 resumes the results obtained from 2-DE analysis when comparing all cell growth medium.

Table 18- Comparative analysis of the 2-DE maps isolated from the different grown media, spots identified and spots differently expressed.

Cell growth medium compared	***Condition tested***	***Spots identified***	***Spots differently expressed***
Medium 1 compared to medium 5	Use of agro-industrial extracts as culture media	215 spots	52 spots
Medium 3 compared to medium 5	Effect of YE in complex growth medium with aaa	519 spots	80 spots
Medium 2 compared to medium 4	Use of the YE in initicial growth medium (2)	386 spots	54 spots
Medium 2 compared to medium 3	Identify proteins involved in metabolism of amino acids	314 spots	25 spots
Medium 4 compared to medium 5	Effect of the addiction with aaa	251 spots	23 spots

Scatter plots and correlations coefficients were calculated in this study, in order to analyze gel similarities and experimental variations. Regression analysis yielded a correlation coefficient greater than 0.9 in every match made, revealing a strong positive correlation. Therefore, figure 29 represents scatter plots and correlation indexes from 2DE gels. The correlation indexes seem to be closer to 1, which indicates that the match between gels was well performed.

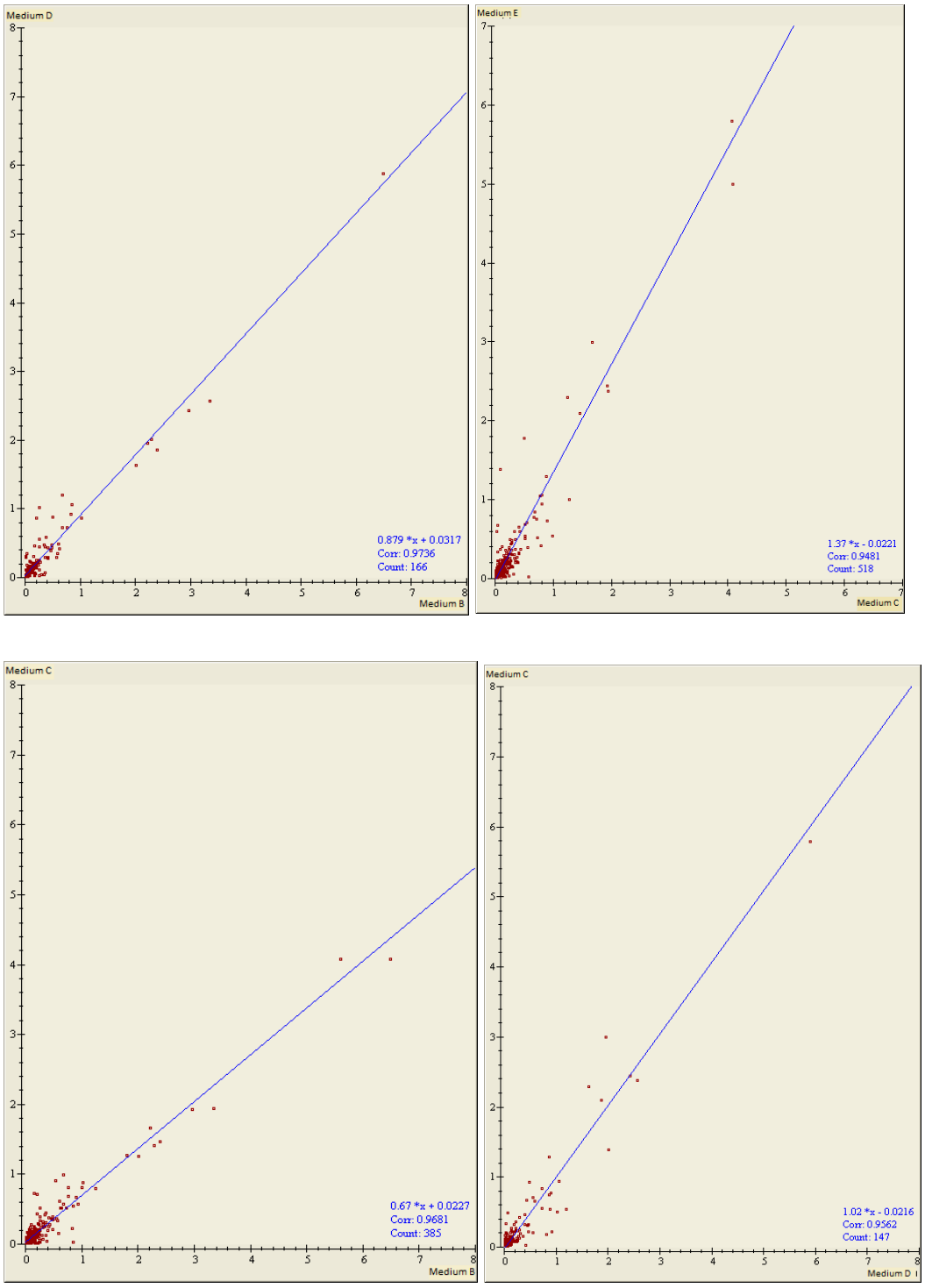


Figure 29 - Scatter plots from 2DE gels. Scatter plots of values (%Vol) of matched spots obtained from representative 2D maps.

Chapter 5 - Conclusions

The present study comprised different specific goals; the optimization of the production of pcDNA-FLAG-p53 using *E. coli* VH35 as host and CW and CSL as sources of carbon and nitrogen, the analysis the *E.coli* VH35 metabolites (sugars and acetate) of growth media by HPLC during fermentation time, and study of the proteome of *E.coli* VH35 using the optimized media.

Regarding the first goal, the results show that the *E.coli* VH35 is able to produce pDNA from agro-industrial wastes, some of them supplemented with YE and amino acids. The medium composed by 2% CSL, 1.2% CW, 5g/L YE and aromatic amino acids presented the best performance in the production of pDNA with a volumetric yield of 42.061 ± 1.012 μg/mL. It proves that is possible to produce pDNA using *E. coli* VH35, with good volumetric yields, under unusual growth conditions, with lactose as a principal carbon source, CSL and YE as nitrogen source and supplemented with aromatic amino acids.

The results with HPLC demonstrate that strain VH35 may be an efficient strain to be used in an economically sustainable procedure for pDNA production. This work proved that strain VH35 uses agro-food media as an energy source, when insufficient glucose levels are available. *E.coli* VH 35 chooses to consume lactose in the presence of an agro-food medium, due to the lack of PTS–sugars in the medium.

In literature, the methodologies available for the analysis of the *E.coli* proteome by 2DE are quite different. So, in this work, the initial optimization of the 2DE protocol was performed in order to improve the number of visualized spots. The initial optimization conducted using non-centrifuged lysates and a pH linear gradient of 4–7, provided an improved gel resolution and a higher number of spots. After optimization, *E. coli* proteome was analyzed by 2DE to provide a comprehensive view of the proteins that are changed due to growing medium alteration.

Results showed that the protein composition of *E.coli* VH35 was different among the various analyzed growth medium. When comparing the glucose medium to the agro-food medium, 52 proteins was found differentially expressed. In addition, when comparing the addition of yeast extract in the initial growth medium to the one supplemented with amino acids, a total 110 proteins spots was found differentially expressed.

Using the ProteinPilot 4.5 software, it was possible to identify under a 95% confidence some proteins in the made comparisons, except between the medium 1 and 5. The proteins identification in this comparison was not possible; however, we can conclude that between the conventional and unconventional mediums, 53 proteins are differentially expressed indicating differences in the cells proteome.The identified proteins were involved in various cellular processes such as metabolic, binding and cellular response processes. It was possible to identify

proteins which are associated with metabolism and that are related to nucleotides synthesis. As it is the case of the rod shape-determining protein MreB (comparison medium 2 with medium 3), anthranilate synthase component 1 (comparison with medium 3 and 5), chaperone protein ClpB (comparison with medium 2 and 4) and deoxyribose-phosphate aldolase (comparison with medium 4 and 5).

Chapter 6 - Future perspectives

In this study it was demonstrated the potential of alternative carbon and nitrogen sources in the therapeutic pDNA biosynthesis. Also, promising grown medium were optimized and selected for a large fermentation scale. However, further investigation may be performed for an overall understanding of the proteomic activity of the *E. coli* strain VH35 in the production of therapeutic plasmids. So, to improve this research:

- To understand the effect of different carbon and nitrogen sources, during the fermentation period using proteomic methodologies;
- Promote s*cale-up* to a mini-reactor platform with a suitable model development using CSL and CW as nitrogen and carbon sources in a growth medium;
- Search new strategies for increasing pDNA yield, for example, use IPTG as an inducer or use other extracts from agro-food industries in the growth media;
- Integrate the process described in this work in a sustainable biotechnological procedure, focusing in the downstream process, especially in purification step (using a monolithic support).

References

[1] R. Soto, L. Caspeta, B. Barrón, G. Gosset, O.T. Ramírez, A.R. Lara, High cell-density cultivation in batch mode for plasmid DNA production by a metabolically engineered E. coli strain with minimized overflow metabolism, Biochem. Eng. J. 56 (2011) 165–171. doi:10.1016/j.bej.2011.06.003.

[2] A. Ghanem, R. Healey, F.G. Adly, Current trends in separation of plasmid DNA vaccines: A review, Anal. Chim. Acta. 760 (2013) 1–15. doi:10.1016/j.aca.2012.11.006.

[3] K.J. Prather, S. Sagar, J. Murphy, M. Chartrain, Industrial scale production of plasmid DNA for vaccine and gene therapy: Plasmid design, production, and purification, Enzyme Microb. Technol. 33 (2003) 865–883. doi:10.1016/S0141-0229(03)00205-9.

[4] A. Xenopoulos, P. Pattnaik, Production and purification of plasmid DNA vaccines: is there scope for further innovation?, Expert Rev. Vaccines. 13 (2014) 1537–51. doi:10.1586/14760584.2014.968556.

[5] Y. Han, S. Liu, J. Ho, M.K. Danquah, G.M. Forde, Using DNA as a drug-Bioprocessing and delivery strategies, Chem. Eng. Res. Des. 87 (2009) 343–348. doi:10.1016/j.cherd.2008.09.010.

[6] M.K. Danquah, G.M. Forde, Growth medium selection and its economic impact on plasmid DNA production., J. Biosci. Bioeng. 104 (2007) 490–497. doi:10.1263/jbb.104.490.

[7] F. Silva, J.A. Queiroz, F.C. Domingues, Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli, Biotechnol. Adv. 30 (2012) 691–708. doi:10.1016/j.biotechadv.2011.12.005.

[8] A.C. Silva, P.M.R. Guimarães, J.A. Teixeira, L. Domingues, Fermentation of deproteinized cheese whey powder solutions to ethanol by engineered Saccharomyces cerevisiae: Effect of supplementation with corn steep liquor and repeated-batch operation with biomass recycling by flocculation, J. Ind. Microbiol. Biotechnol. 37 (2010) 973–982. doi:10.1007/s10295-010-0748-z.

[9] B. Hong, a P.J. van den Heuvel, V. V Prabhu, S. Zhang, W.S. El-Deiry, Targeting Tumor Suppressor p53 for Cancer Therapy: Strategies, Challenges and Opportunities., Curr. Drug Targets. 15 (2014) 80–9. doi:10.2174/1389450114666140106101412.

[10] D.P. Lane, C.F. Cheok, S. Lain, p53-based Cancer Therapy p53-based Cancer Therapy,

Cold Spring Harb Perspect Biol. 2 (2010) 1–25. doi:10.1101/cshperspect.a001222.

[11] S. Duarte, G. Carle, H. Faneca, M.C.P. de Lima, V. Pierrefite-Carle, Suicide gene therapy in cancer: Where do we stand now?, Cancer Lett. 324 (2012) 160–170. doi:10.1016/j.canlet.2012.05.023.

[12] J. Liu, C. Zhang, Z. Feng, Tumor suppressor p53 and its gain-of-function mutants in cancer The p53 Signaling Pathway Introduction Tumor Suppressive Functions of p53, Acta Biochim. Biophys. …. 46 (2013) 1–10. doi:10.1093/abbs/gmt144.Review.

[13] L. Böhlig, K. Rother, One function - Multiple mechanisms: The manifold activities of p53 as a transcriptional repressor, J. Biomed. Biotechnol. 2011 (2011). doi:10.1155/2011/464916.

[14] N. Vlatković, M.T. Boyd, C.P. Rubbi, Nucleolar control of p53: A cellular achilles' heel and a target for cancer therapy, Cell. Mol. Life Sci. 71 (2014) 771–791. doi:10.1007/s00018-013-1361-x.

[15] O. Toshinori, N. Akira, Role of p53 in Cell Death and Human Cancers, Cancers (Basel). 3 (2011) 994–1013. doi:10.3390/cancers3010994.

[16] K. Lee, T. Wang, A.J. Paszczynski, S.S. Daoud, Expression proteomics to p53 mutation reactivation with PRIMA-1 in breast cancer cells, Biochem. Biophys. Res. Commun. 349 (2006) 1117–1124. doi:10.1016/j.bbrc.2006.08.152.

[17] P.A.J. Muller, K.H. Vousden, Mutant p53 in cancer: New functions and therapeutic opportunities, Cancer Cell. 25 (2014) 304–317. doi:10.1016/j.ccr.2014.01.021.

[18] G.M. Borja, E. Meza Mora, B. Barrón, G. Gosset, O.T. Ramírez, A.R. Lara, Engineering Escherichia coli to increase plasmid DNA production in high cell-density cultivations in batch mode., Microb. Cell Fact. 11 (2012) 132. doi:10.1186/1475-2859-11-132.

[19] A.E. Carnes, J.M. Luke, J.M. Vincent, A. Schukar, S. Anderson, C.P. Hodgson, J.A. Williams, Plasmid DNA fermentation strain and process-specific effects on vector yield, quality, and transgene expression, Biotechnol. Bioeng. 108 (2011) 354–363. doi:10.1002/bit.22936.

[20] I. Van Der Heijden, Production and pharmaceutical formulation of plasmid DNA vaccines, n.d.

[21] A. Carnes, Fermentation design for the manufacture of therapeutic plasmid DNA, Bioprocess Int. (2005) 36–42. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Fermentation+Desig

n+for+the+Manufacture+of+Therapeutic+Plasmid+DNA#0.

[22] T.J. Magliery, L. Regan, A cell-based screen for function of the four-helix bundle protein Rop: A new tool for combinatorial experiments in biophysics, Protein Eng. Des. Sel. 17 (2004) 77–83. doi:10.1093/protein/gzh010.

[23] E.M. Wong, M.A. Muesing, B. Polisky, Temperature-sensitive copy number mutants of ColE1 are located in an untranslated region of the plasmid genome., Proc. Natl. Acad. Sci. U. S. A. 79 (1982) 3570–4. doi:10.1073/pnas.79.11.3570.

[24] D.M. Bower, K.L.J. Prather, Engineering of bacterial strains and vectors for the production of plasmid DNA, Appl. Microbiol. Biotechnol. 82 (2009) 805–813. doi:10.1007/s00253-009-1889-8.

[25] P.H. Oliveira, K.J. Prather, D.M.F. Prazeres, G.A. Monteiro, Structural instability of plasmid biopharmaceuticals: challenges and implications, Trends Biotechnol. 27 (2009) 503–511. doi:10.1016/j.tibtech.2009.06.004.

[26] S.Y. Yau, E. Keshavarz-Moore, J. Ward, Host strain influences on supercoiled plasmid DNA production in Escherichia coli: Implications for efficient design of large-scale processes, Biotechnol. Bioeng. 101 (2008) 529–544. doi:10.1002/bit.21915.

[27] K. Listner, L. Bentley, J. Okonkowski, C. Kistler, R. Wnek, A. Caparoni, B. Junker, D. Robinson, P. Salmon, M. Chartrain, Development of a highly productive and scalable plasmid DNA production platform, Biotechnol. Prog. 22 (2006) 1335–1345. doi:10.1021/bp060046h.

[28] A. Berlec, B. Štrukelj, Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells, J. Ind. Microbiol. Biotechnol. 40 (2013) 257–274. doi:10.1007/s10295-013-1235-0.

[29] M.M. Salem-bekhit, M. Bayomi, F. Al-anazi, H. Radawan, Yield Production of Recombinant Plasmid Dna With Escherichia Coli in Fed-Batch Culture By Pseudo-Exponential, Biotechnol. Biotechnol. (2012) 3709–3715. doi:10.5504/BBEQ.2012.0131.

[30] J. Susnea, B. Bernevic, M. Wicke, S. Liu, K. Schellander, M. Przybylsky, Application of MALDI-TOF-Mass Spectrometry to Proteome Analysis Using Stain-Free Gel Electrophoresis, TripleC. 11 (2013) 13–35. doi:10.1007/128.

[31] G.A.L. Gonçalves, K.L.J. Prather, G.A. Monteiro, D.M.F. Prazeres, Engineering of Escherichia coli strains for plasmid biopharmaceutical production: Scale-up challenges, Vaccine. 32 (2014) 2847–2850. doi:10.1016/j.vaccine.2014.02.023.

[32] K. O'Mahony, R. Freitag, F. Hilbrig, P. Müller, I. Schumacher, Strategies for high titre plasmid DNA production in Escherichia coli DH5α, Process Biochem. 42 (2007) 1039–1049. doi:10.1016/j.procbio.2007.03.014.

[33] P.A. Shamlou, Scaleable processes for the manufacture of therapeutic quantities of plasmid DNA., Biotechnol. Appl. Biochem. 37 (2003) 207–18. doi:10.1042/BA20030011.

[34] M.M. Diogo, J.A. Queiroz, D.M.F. Prazeres, Assessment of purity and quantification of plasmid DNA in process solutions using high-performance hydrophobic interaction chromatography, J. Chromatogr. A. 998 (2003) 109–117. doi:10.1016/S0021-9673(03)00618-6.

[35] F. Silva, L. Passarinha, F. Sousa, J.A. Queiroz, F.C. Domingues, Influence of growth conditions on plasmid DNA production, J. Microbiol. Biotechnol. 19 (2009) 1408–1414. doi:10.4014/jmb.0805.0329.

[36] J. Yang, Y. Yang, Plasmid size can affect the ability of Escherichia coli to produce high-quality plasmids, Biotechnol. Lett. 34 (2012) 2017–2022. doi:10.1007/s10529-012-0994-4.

[37] E. Meza, J. Becker, F. Bolivar, G. Gosset, C. Wittmann, Consequences of phosphoenolpyruvate:sugar phosphotranferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in Escherichia coli, Microb. Cell Fact. 11 (2012) 127. doi:10.1186/1475-2859-11-127.

[38] H. Ingmer, C. Miller, S.N. Cohen, The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response, Mol. Microbiol. 42 (2001) 519–526. doi:10.1046/j.1365-2958.2001.02661.

[39] G. Jing, Z. Huang, Z. Liu, Q. Zou, Plasmid pKKH: an improved vector with higher copy number for expression of foreign genes in Escherichia coli, Biotechnol. Lett. 15 (1993) 439–442. http://link.springer.com/article/10.1007/BF00129314.

[40] R.D. O'Kennedy, C. Baldwin, E. Keshavarz-Moore, Effects of growth medium selection on plasmid DNA production and initial processing steps, J. Biotechnol. 76 (2000) 175–183. doi:10.1016/S0168-1656(99)00187-X.

[41] S. Zheng, K. Friehs, N. He, X. Deng, Q. Li, Z. He, C. Xu, Y. Lu, Optimization of Medium Components for Plasmid Production by Recombinant E.coli DH5α pUK21CMVβ1.2, Biotechnol. Bioprocess Eng. 12 (2007) 213–221. doi:10.1007/BF02931095.

[42] Z.-N. Xu, W.-H. Shen, H. Chen, P.-L. Cen, Effects of medium composition on the production of plasmid DNA vector potentially for human gene therapy., J. Zhejiang Univ.

Sci. B. 6 (2005) 396–400. doi:10.1631/jzus.2005.B0396.

[43] G. Aidelberg, B.D. Towbin, D. Rothschild, E. Dekel, A. Bren, U. Alon, Hierarchy of non-glucose sugars in Escherichia coli., BMC Syst. Biol. 8 (2014) 133. doi:10.1186/s12918-014-0133-z.

[44] A. Bren, J.O. Park, B.D. Towbin, E. Dekel, J.D. Rabinowitz, U. Alon, Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP., Sci. Rep. 6 (2016) 24834. doi:10.1038/srep24834.

[45] L.R. Rodrigues, J.A. Teixeira, R. Oliveira, Low-cost fermentative medium for biosurfactant production by probiotic bacteria, Biochem. Eng. J. 32 (2006) 135–142. doi:10.1016/j.bej.2006.09.012.

[46] F.B. Pereira, P.M.R. Guimarães, J.A. Teixeira, L. Domingues, Optimization of low-cost medium for very high gravity ethanol fermentations by Saccharomyces cerevisiae using statistical experimental designs, Bioresour. Technol. 101 (2010) 7856–7863. doi:10.1016/j.biortech.2010.04.082.

[47] P.M.R. Guimarães, J.A. Teixeira, L. Domingues, Fermentation of lactose to bio-ethanol by yeasts as part of integrated solutions for the valorisation of cheese whey, Biotechnol. Adv. 28 (2010) 375–384. doi:10.1016/j.biotechadv.2010.02.002.

[48] E.J. Gudiña, A.I. Rodrigues, V. de Freitas, Z. Azevedo, J.A. Teixeira, L.R. Rodrigues, Valorization of agro-industrial wastes towards the production of rhamnolipids, Bioresour. Technol. 212 (2016) 144–150. doi:10.1016/j.biortech.2016.04.027.

[49] H. Il Jung, O.M. Lee, J.H. Jeong, Y.D. Jeon, K.H. Park, H.S. Kim, W.G. An, H.J. Son, Production and characterization of cellulose by acetobacter sp. V6 using a cost-effective molasses-corn steep liquor medium, Appl. Biochem. Biotechnol. 162 (2010) 486–497. doi:10.1007/s12010-009-8759-9.

[50] B.L.B. Perini, H.C.M. Souza, M. Kelbert, G.P. Apati, A.P.T. Pezzin, A.L.S. Schneider, Production of ß-Galactosidase from cheese whey using kluyveromyces marxianus CBS 6556, Chem. Eng. Trans. 32 (2013) 991–996. doi:10.3303/CET1332166.

[51] D.R. Palatnik, M.V. Ostermann Porcel, U. González, N. Zaritzky, M.E. Campderrós, Recovery of caprine whey protein and its application in a food protein formulation, LWT - Food Sci. Technol. 63 (2015) 331–338. doi:10.1016/j.lwt.2015.03.027.

[52] C. Spalatelu, Biotechnologival valorisation of whey, Innov. Rom. Food Biotechnol. 10 (2012) 1–8. doi:10.1056/NEJMra1313875.

[53] M.I. Viitanen, A. Vasala, P. Neubauer, T. Alatossava, Cheese whey-induced high-cell-density production of recombinant proteins in Escherichia coli., Microb. Cell Fact. 2 (2003) 2. doi:10.1186/1475-2859-2-2.

[54] M.K. Gouda, a E. Swellam, S.H. Omar, Production of PHB by a Bacillus megaterium strain using sugarcane molasses and corn steep liquor as sole carbon and nitrogen sources., Microbiol. Res. 156 (2001) 201–207. doi:10.1078/0944-5013-00104.

[55] L. Agarwal, K. Dutt, G.K. Meghwanshi, R.K. Saxena, Anaerobic fermentative production of lactic acid using cheese whey and corn steep liquor, Biotechnol. Lett. 30 (2008) 631–635. doi:10.1007/s10529-007-9592-2.

[56] A.E. Carnes, C.P. Hodgson, J.A. Williams, Inducible Escherichia coli fermentation for increased plasmid DNA production., Biotechnol. Appl. Biochem. 45 (2006) 155–166. doi:10.1042/BA20050223.

[57] S.S. Levanon, K.Y. San, G.N. Bennett, Effect of oxygen on the Escherichia coli ArcA and FNR regulation systems and metabolic responses, Biotechnol. Bioeng. 89 (2005) 556–564. doi:10.1002/bit.20381.

[58] L.A. Passarinha, M.M. Diogo, J.A. Queiroz, G.A. Monteiro, L.P. Fonseca, D.M.F. Prazeres, Production of ColE1 type plasmid by Escherichia coli DH5?? cultured under nonselective conditions, J. Microbiol. Biotechnol. 16 (2006) 20–24.

[59] L.M. Martins, A.Q. Pedro, D. Oppolzer, F. Sousa, J.A. Queiroz, L.A. Passarinha, Enhanced biosynthesis of plasmid DNA from Escherichia coli VH33 using Box-Behnken design associated to aromatic amino acids pathway, Biochem. Eng. J. 98 (2015) 117–126. doi:10.1016/j.bej.2015.02.001.

[60] G.W. Luli, W.R. Strohl, Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations, Appl. Environ. Microbiol. 56 (1990) 1004–1011.

[61] J.Y.H. Kim, H.J. Cha, Down-regulation of acetate pathway through antisense strategy in Escherichia coli: Improved foreign protein production, Biotechnol. Bioeng. 83 (2003) 841–853. doi:10.1002/bit.10735.

[62] G.L. Kleman, W.R. Strohl, Acetate metabolism by Escherichia coli in high-cell-density fermentation, Appl. Environ. Microbiol. 60 (1994) 3952–3958.

[63] D.W. Zabriskie, E.J. Arcuri, Factors influencing productivity of fermentations employing recombinant microorganisms, Enzyme Microb. Technol. 8 (1986) 706–717. doi:http://dx.doi.org/10.1016/0141-0229(86)90157-2.

[64] K. Shimizu, Regulation Systems of Bacteria such as Escherichia coli in Response to Nutrient Limitation and Environmental Stresses, Metabolites. 4 (2013) 1–35. doi:10.3390/metabo4010001.

[65] J.A. Williams, J. Luke, S. Langtry, S. Anderson, C.P. Hodgson, A.E. Carnes, Generic plasmid DNA production platform incorporating low metabolic burden seed-stock and fed-batch fermentation processes, Biotechnol. Bioeng. 103 (2009) 1129–1143. doi:10.1002/bit.22347.

[66] J. Shiloach, R. Fass, Growing E. coli to high cell density - A historical perspective on method development, Biotechnol. Adv. 23 (2005) 345–357. doi:10.1016/j.biotechadv.2005.04.004.

[67] P.W. Postma, J.W. Lengeler, G.R. Jacobson, Phosphoenolpyruvate: Carbohydrate phosphotransferase systems of bacteria., Microbiol. Rev. 57 (1993) 543–594.

[68] J. Stülke, W. Hillen, Carbon catabolite repression in bacteria., Curr. Opin. Microbiol. 2 (1999) 195–201. doi:10.1016/S1369-5274(99)80034-4.

[69] S. Huffer, C.M. Roche, H.W. Blanch, D.S. Clark, Escherichia coli for biofuel production: Bridging the gap from promise to practice, Trends Biotechnol. 30 (2012) 538–545. doi:10.1016/j.tibtech.2012.07.002.

[70] J. Plumbridge, Regulation of gene expression in the PTS in Escherichia coli: The role and interactions of MIc, Curr. Opin. Microbiol. 5 (2002) 187–193. doi:10.1016/S1369-5274(02)00296-5.

[71] P. Kotrba, M. Inui, H. Yukawa, Bacterial Phosphotransferase System (PTS) in Carbohydrate Uptake and Control of Carbon Metabolism, 92 (2001) 502–517. doi:10.1016/S1389-1723(01)80308-X.

[72] N. Flores, S. Flores, A. Escalante, R. De Anda, L. Leal, R. Malpica, D. Georgellis, G. Gosset, F. Bolívar, Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an Escherichia coli strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system, Metab. Eng. 7 (2005) 70–87. doi:10.1016/j.ymben.2004.10.002.

[73] A.M. Lewis, N.R. Abu-absi, M.C. Borys, Z.J. Li, The Use of ‘ Omics Technology to Rationally Improve Industrial Mammalian Cell Line Performance, Biotechnol. Bioeng. 113 (2016) 26–38. doi:10.1002/bit.25673.

[74] P. Pomastowski, B. Buszewski, Two-dimensional gel electrophoresis in the light of new developments, TrAC - Trends Anal. Chem. 53 (2014) 167–177.

doi:10.1016/j.trac.2013.09.010.

[75] D.E. Garcia, E.E. Baidoo, P.I. Benke, F. Pingitore, Y.J. Tang, S. Villa, J.D. Keasling, Separation and mass spectrometry in microbial metabolomics, Curr. Opin. Microbiol. 11 (2008) 233–239. doi:10.1016/j.mib.2008.04.002.

[76] A. Zhang, H. Sun, P. Wang, Y. Han, X. Wang, Modern analytical techniques in metabolomics analysis, Analyst. 137 (2012) 293. doi:10.1039/c1an15605e.

[77] S.G. Villas-Boas, Analytical techniques & applications of metabolomics in systems medicine and systems biotechnology., Comput. Struct. Biotechnol. J. 4 (2013) e201301001. doi:10.5936/csbj.201301001.

[78] K. Shimizu, L. Peng, Global metabolic regulation analysis for Escherichia coli K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement, Appl Microbiol Biotechnol. 61 (2003) 163–178. doi:10.1007/s00253-002-1202-6.

[79] S. Carneiro, E.C. Ferreira, I. Rocha, Metabolic responses to recombinant bioprocesses in Escherichia coli, J. Biotechnol. 164 (2013) 396–408. doi:10.1016/j.jbiotec.2012.08.026.

[80] M. Vollmer, E. Nagele, P. Horth, Differential Proteome Analysis: Two-Dimensional Nano-LC/MS of E. Coli Proteome Grown on Different Carbon Sources, J Biomol Tech. 14 (2003) 128–135. doi:10.1016/j.biomaterials.2009.03.056.

[81] E. Fic, S. Kedracka-Krok, U. Jankowska, A. Pirog, M. Dziedzicka-Wasylewska, Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis, Electrophoresis. 31 (2010) 3573–3579. doi:10.1002/elps.201000197.

[82] J. Meade, P. Bartlow, R.N. Trivedi, P. Akhtar, M.M. Ataai, S.A. Khan, M.M. Domach, Effect of plasmid replication deregulation via inc mutations on E . coli proteome & simple flux model analysis, Microb. Cell Fact. 14 (2015) 1–13. doi:10.1186/s12934-015-0212-x.

[83] K. Gevaert, J. Vanderkerckhove, Protein identification methods in proteomics, Electrophoresis. 21 (2000) 1145–1154. doi:10.1002/(SICI)1522-2683(20000401)21:6<1145::AID-ELPS1145>3.0.CO;2-Z.

[84] P. a Haynes, S.P. Gygi, D. Figeys, R. Aebersold, Proteome analysis: biological assay or data archive?, Electrophoresis. 19 (1998) 1862–71. doi:10.1002/elps.1150191104.

[85] S. Sadana, S. Kalia, P. Sharma, H. Singh, Emerging Techniques in Proteomics for Probing Nano–Bio Interactions, BIOMIRROR. 6 (2012) 22–30. doi:10.1021/nn3052499.

[86] B.L. Hood, T.D. Veenstra, T.P. Conrads, Mass spectrometry-based proteomics, Int. Congr. Ser. 1266 (2004) 375–380. doi:10.1016/j.ics.2004.02.087.

[87] M.L. Yarmush, A. Jayaraman, Advances in proteomic technologies., Annu. Rev. Biomed. Eng. 4 (2002) 349–373. doi:10.1146/annurev.bioeng.4.020702.153443.

[88] A. Bodzon-Kulakowska, A. Bierczynska-Krzysik, T. Dylag, A. Drabik, P. Suder, M. Noga, J. Jarzebinska, J. Silberring, Methods for samples preparation in proteomic research, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 849 (2007) 1–31. doi:10.1016/j.jchromb.2006.10.040.

[89] G.E. Healthcare, 2D Platinum 7.0, ImageMaster User Manua. (2008) 1–200.

[90] P.H. O'Farrell, High Resolution of Proteins * Electrophoresis, J. Biol. Chem. 250 (1975) 4007–4021. doi:10.1016/j.bbi.2008.05.010.

[91] A. Gorg, G. Boguth, C. Obermaier, W. Weiss, Two-dimensional electrophoresis of proteins in an immobilized pH 4-12 gradient, Electrophoresis. 19 (1998) 1516–1519. doi:10.1002/elps.1150190850.

[92] A. Gorg, 2-D Electrophoresis. Principles and Methods., Handbook. (2004) 163. doi:10.1186/1755-8794-4-24.

[93] T.E. Sandberg, M. Pedersen, R.A. Lacroix, A. Ebrahim, M. Bonde, M.J. Herrgard, B.O. Palsson, M. Sommer, A.M. Feist, Evolution of Escherichia coli to 42 C and Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations, Mol. Biol. Evol. 31 (2014) 2647–2662. doi:10.1093/molbev/msu209.

[94] R.D. Unwin, C.A. Evans, A.D. Whetton, Relative quantification in proteomics: new approaches for biochemistry, Trends Biochem. Sci. 31 (2006) 473–484. doi:10.1016/j.tibs.2006.06.003.

[95] M. Debald, M. Wolfgarten, G. Walgenbach-Brünagel, W. Kuhn, M. Braun, Non-invasive proteomics-thinking about personalized breast cancer screening and treatment, EPMA J. 1 (2010) 413–420. doi:10.1007/s13167-010-0039-9.

[96] B. Canas, C. Piñeiro, E. Calvo, D. Lopez-Ferrer, J.M. Gallardo, Trends in sample preparation for classical and second generation proteomics, J. Chromatogr. A. 1153 (2007) 235–258. doi:10.1016/j.chroma.2007.01.045.

[97] S.B. Primrose, R.M. Twyman, Principles of Gene Manipulation and Genomics, 2006. doi:10.1136/jmg.23.3.282.

[98] T. Rabilloud, C. Lelong, Two-dimensional gel electrophoresis in proteomics : A tutorial

☆, J. Proteomics. 74 (2011) 1829–1841. doi:10.1016/j.jprot.2011.05.040.

[99] S. Magdeldin, S. Enany, Y. Yoshida, B. Xu, Y. Zhang, Z. Zureena, I. Lokamani, E. Yaoita, T. Yamamoto, Basics and recent advances of two dimensional- polyacrylamide gel electrophoresis., Clin. Proteomics. 11 (2014) 16. doi:10.1186/1559-0275-11-16.

[100] B.D.E. Garfin, D.E. Garfin, Two-dimensional gel electrophoresis : an overview Two-dimensional gel electrophoresis : an overview, Trends Anal. Chem. 22 (2003) 264–272. doi:10.1016/S0165-9936(03)00506-5.

[101] G. Candiano, M. Bruschi, L. Musante, L. Santucci, G.M. Ghiggeri, B. Carnemolla, P. Orecchia, L. Zardi, P.G. Righetti, Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis, Electrophoresis. 25 (2004) 1327–1333. doi:10.1002/elps.200305844.

[102] W.F. Patton, Detection technologies in proteome analysis, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 771 (2002) 3–31. doi:10.1016/S1570-0232(02)00043-0.

[103] F. Chevalier, Standard dyes for total protein staining in gel-based proteomic analysis, Materials (Basel). 3 (2010) 4784–4792. doi:10.3390/ma3104784.

[104] D. Fluor, U. Decyder, D.A. Software, U. Maldi-tof, E. Dige, V.M. Imager, D. Differential, A. Software, E. Dige, E. Dige, E. Dige, Fluorescence 2-D difference gel electrophoresis (2-D DIGE) using Ettan DIGE system and mass spectrometry proteomic analysis of Escherichia coli, (2003) 1–6.

[105] M. Natale, B. Maresca, P. Abrescia, E.M. Bucci, Image analysis workflow for 2-D electrophoresis gels based on imageJ, Proteomics Insights. 4 (2011) 37–49. doi:10.4137/PRI.S7971.

[106] A.W. Dowsey, J. a English, F. Lisacek, J.S. Morris, Z. Yang, M.J. Dunn, NIH Public Access, Proteomics. 10 (2012) 4226–4257. doi:10.1002/pmic.200900635.Image.

[107] C. Pasquali, M.R. Wilkins, G.J. Hughes, R.D. Appep, A. Bairoch, D. Schalle, D.F. Hochstrasser, Two-dimensional gel electrophoresis of Escherichia coli homogenates : The Escherichia coli SWISS-2DPAGE database, Electrophoresis. 17 (1996) 547–555. doi:10.1002/elps.1150170325.

[108] M. Bucknall, K.Y.C. Fung, M.W. Duncan, Practical Quantitative Biomedical Applications of MALDI-TOF Mass Spectrometry, 305 (2002) 1015–1027. doi:0.1016/S1044-0305(02)00426-9.

[109] B. Canas Montalvo, D. Lopez-Ferrer, A. Ramos-Fernandez, E. Camafeita, E. Calvo, Mass

spectrometry technologies for proteomics, Briefings Funct. Genomics Proteomics. 4 (2006) 295–320. doi:10.1093/bfgp/eli002.

[110] C. Jurinke, P. Oeth, D. van den Boom, MALDI-TOF mass spectrometry: a versatile tool for high-performance DNA analysis., Mol. Biotechnol. 26 (2004) 147–164. doi:10.1385/MB:26:2:147.

[111] E. De Hoffmann, V. Stroobant, Mass Spectrometry - Priniples and Applications., 2007. doi:10.1002/mas.20296.

[112] V. Horneffer, M. Gluckmann, R. Kruger, M. Karas, K. Strupat, Hillenkam, F. P, Matrix-analyte-interaction in MALDI-MS: Pellet and nano-electrospray preparations, Int. J. Mass Spectrom. 249–250 (2006) 426–432. doi:10.1016/j.ijms.2005.11.012.

[113] M. Kussmann, E. Nordho, H. Rahbek-nielsen, S. Haebel, M. Rossel-larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A. Kroll-kristensen, Ò.L. Palmp, P. Roepstor, Matrix-assisted Laser Desorption / Ionization Mass Spectrometry Sample Preparation Techniques Designed for Various Peptide and Protein, J. Mass Spectrom. 32 (1997) 593–601. doi:10.1002/(SICI)1096-9888(199706)32:6<593::AID-JMS511>3.0.CO;2-D.

[114] Z. Liu, K.L. Schey, Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis, J. Am. Soc. Mass Spectrom. 16 (2005) 482–490. doi:10.1016/j.jasms.2004.12.018.

[115] M.A. Kuzyk, L.B. Ohlund, M.H. Elliott, D. Smith, H. Qian, A. Delaney, C.L. Hunter, C.H. Borchers, A comparison of MS/MS-based, stable-isotope-labeled, quantitation performance on ESI-quadrupole TOF and MALDI-TOF/TOF mass spectrometers, Proteomics. 9 (2009) 3328–3340. doi:10.1002/pmic.200800412.

[116] B. Thiede, W. Höhenwarter, A. Krah, J. Mattow, M. Schmid, F. Schmidt, P.R. Jungblut, Peptide mass fingerprinting, Methods. 35 (2005) 237–247. doi:10.1016/j.ymeth.2004.08.015.

[117] N.J. Edwards, Protein Identification from Tandem Mass Spectra by Database Searching, in: H.C. Wu, C. Chen (Eds.), Bioinforma. Comp. Proteomics, Humana Press, Totowa, NJ, 2011: pp. 119–138. doi:10.1007/978-1-60761-977-2_9.

[118] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis. 20 (1999) 3551–3567. doi:10.1002/(sici)1522-2683(19991201)20:18<3551::aid-elps3551>3.0.co;2-2.

[119] S.L. Seymour, C. Hunter, ProteinPilot ™ Software Overview, (n.d.) 1–5.

[120] Z. Minic, Proteomic Studies of the Effects of Different Stress Conditions on Central Carbon Metabolism in Microorganisms, J. Proteomics Bioinform. 8 (2015) 80–90. doi:10.4172/jpb.1000355.

[121] K.M. Champion, J.C. Nishihara, J.C. Joly, D. Arnott, Similarity of the Escherichia coli proteome upon completion of different biopharmaceutical, Proteomics. 1 (2001) 1133–1148. doi:10.1002/1615-9861(200109)1:9<1133::AID-PROT1133>3.0.CO;2-S.

[122] B. Soufi, K. Krug, A. Harst, B. Macek, Characterization of the E . coli proteome and its modifications during growth and ethanol stress, Front. Microbiol. 6 (2015) 1–11. doi:10.3389/fmicb.2015.00103.

[123] K. Peebo, K. Valgepea, A. Maser, R. Nahku, K. Adamberg, R. Vilu, Proteome reallocation in Escherichia coli with increasing specific growth rate, Mol. Biosyst. 11 (2015) 1184–1193. doi:10.1039/C4MB00721B.

[124] V.A. Pferdeort, T.K. Wood, K.F. Reardon, Proteomic changes in Escherichia coli TG1 after metabolic engineering for enhanced trichloroethene biodegradation, Proteomics. 3 (2003) 1066–1069. doi:10.1002/pmic.200300418.

[125] V. Luang-in, A. Narbad, F. Cebeci, M. Bennett, Identification of Proteins Possibly Involved in Glucosinolate Metabolism in L . agilis R16 and E . coli VL8, Protein J. 34 (2015) 135–146. doi:10.1007/s10930-015-9607-0.

[126] D. Blankenhorn, J. Phillips, J.L. Slonczewski, Acid- and Base-Induced Proteins during Aerobic and Anaerobic Growth of Escherichia coli Revealed by Two-Dimensional Gel Electrophoresis, J. Bacteriol. 181 (1999) 2209–2216.

[127] M. Wunderlich, H. Taymaz-Nikerel, G. Gosset, O.T. Ramírez, A.R. Lara, Effect of growth rate on plasmid DNA production and metabolic performance of engineered Escherichia coli strains., J. Biosci. Bioeng. 117 (2014) 336–42. doi:10.1016/j.jbiosc.2013.08.007.

[128] J. Stadler, R. Lemmens, T. Nyhammar, Plasmid DNA purification., J. Gene Med. 6 Suppl 1 (2004) 54–66. doi:10.1002/jgm.512.

[129] H. Li, H. Bo, J. Wang, H. Shao, S. Huang, Separation of supercoiled from open circular forms of plasmid DNA, and biological activity detection, Cytotechnology. 63 (2011) 7–12. doi:10.1007/s10616-010-9322-9.

[130] R. Hermsen, H. Okano, C. You, N. Werner, T. Hwa, A growth-rate composition formula for the growth of E. coli on co-utilized carbon substrates., Mol. Syst. Biol. 11 (2015) 801. doi:10.15252/msb.20145537.

[131] M.I. Chávez-béjar, V.E. Balderas-hernandez, A. Gutiérrez-alejandre, A. Martinez, F. Bolívar, G. Gosset, Metabolic engineering of Escherichia coli to optimize melanin synthesis from glucose, Microb. Cell Fact. 12 (2013) 1–12.

[132] A. Vargas-tah, G. Gosset, Production of cinnamic and p -hydroxycinnamic acids in, Front. Bioeng. Biotechnol. 3 (2015) 1–10. doi:10.3389/fbioe.2015.00116.

[133] R. Pérez-Alfaro, M. Santillán, E. Galan-Vasquez, A. Martinez-Antonio, Regulatory switches for hierarchical use of carbon sources in E. coli, Netw. Biol. 4 (2014) 95–108. doi:10.1016/j.socnet.2004.11.008.

[134] D.L. Bissett, R.L. Anderson, Lactose and D-galactose metabolism in group N streptococci: presence of enzymes for both the D-galactose 1-phosphate and D-tagatose 6-phosphate pathways., J. Bacteriol. 117 (1974) 318–20. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=246560&tool=pmcentrez&rendertype=abstract.

[135] A.M. Rodrigues, P.H. Kubitschek-Barreira, G.F. Fernandes, S.R. de Almeida, L.M. Lopes-Bezerra, Z.P. de Camargo, Two-dimensional gel electrophoresis data for proteomic profiling of Sporothrix yeast cells, Data Br. 2 (2015) 32–38. doi:10.1016/j.dib.2014.11.004.

[136] T. Knigge, J. Letendre, T. Monsinjon, Sample preparation for two-dimensional gel electrophoresis: Considering the composition of biological material, Proteomics. 13 (2013) 3106–2108. doi:10.1002/pmic.201300436.

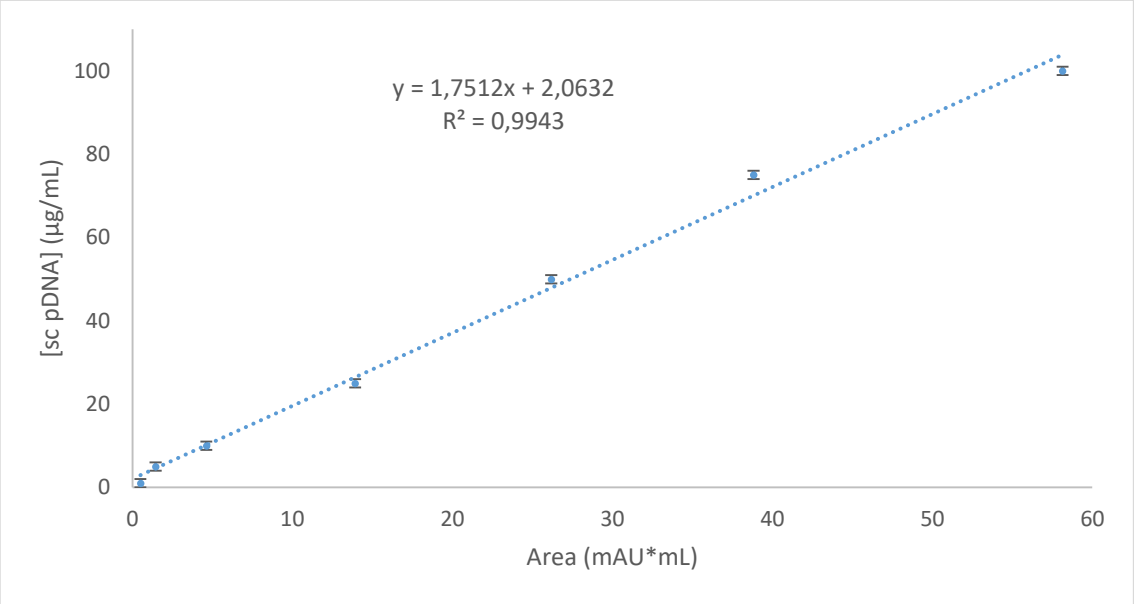
[137] R.R. Burgess, Chapter 20 Protein Precipitation Techniques, 1st ed., Elsevier Inc., 2009. doi:10.1016/S0076-6879(09)63020-2.

[138] P. Feist, A.B. Hummon, Proteomic challenges: Sample preparation techniques for Microgram-Quantity protein analysis from biological samples, Int. J. Mol. Sci. 16 (2015) 3537–3563. doi:10.3390/ijms16023537.

[139] M.J. Whipp, A.J. Pittard, Regulation of Aromatic Amino Acid Transport Systems in Escherichia coli K12, J. Bacteriol. 132 (1977) 453–461.

Appendices

Appendix I - Calibration curve obtained from the correlation between the supercoiled pDNA peak area and the concentration in the range of 1 - 100 µg/mL (y = [sc pDNA] µg/mL, X= peak area mAU* mL)



Appendix II - Effect of different nutrients composition on bacterial growth, peak area of the different tests

	Bacterial Growth		Peak Area		
Medium	**$OD_{600\ nm}$**	**Cell dry mass (CDM) (g/L)**	**pDNA**	**gDNA**	**RNA**
CSL 2% (v/v)+ 1.2% CH(w/v)	9.825 ± 0.643	4.18± 0.047	1.214±0.047	147.69±4.028	475.690± 9.285
CSL 2% (v/v)+ 1.2% CH (w/v)+2.5 (w/v) YE	11.0733± 0.843	4.983 ±0.379	3.763 ±0.243	114. 273± 0.242	492.823 ± 43.233
CSL 2% (v/v)+ 1.2% CH (w/v)+5 (w/v) YE	6.000±0.616	2.700±0.277	18.880 ±0.743	192.849±6.204	461.686±17.475
CSL 2% (v/v)+ 1.2% CH (w/v)+7.5 (w/v) YE	9.267 ±0.379	4.170 ±0.171	11.269 ±1.690	94.381 ±28.127	436.292 ±32.174
CSL 2% (v/v)+ 1.2% CH (w/v)+10 (w/v) YE	11.775 ±0.286	5.299 ±0.129	4.682 ±0.569	124.889 ±32.188	394.151 ±40.258

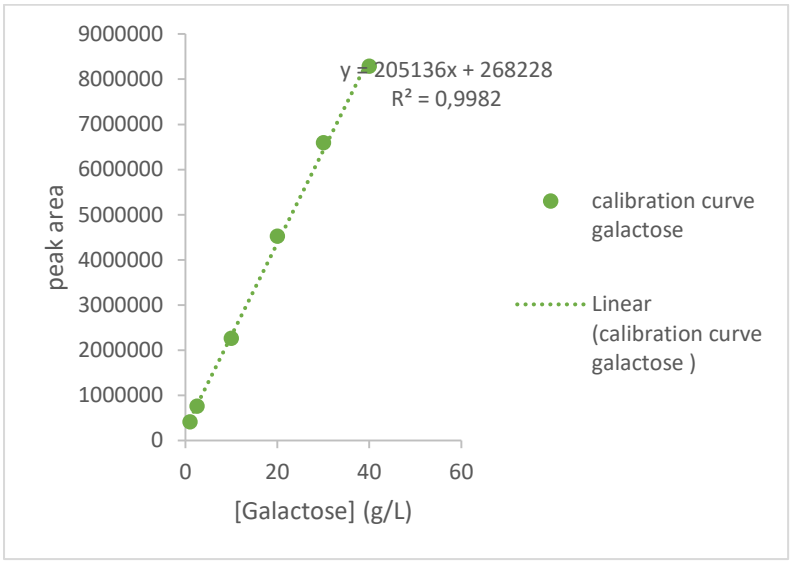
Appendix III - - Effect of manipulating the initial concentration of aromatic amino acids on bacterial growth: peak area of the different test

	Bacterial Growth		Peak Area		
Medium	**$OD_{600\ nm}$**	**Cell dry mass (CDM) (g/L)**	**pDNA**	**gDNA**	**RNA**
CSL 2% (v/v)+ 1.2% CH (w/v)+5 (g/L) YE+ 78.5 (µg/ mL) phenylalanine + 10 (µg/ mL) tyrosine +100(µg/ mL)tryptophan	4.955±0.075	2.230	22.435±0.708	161.020±6.641	507.987±58.350
CSL 2% (v/v)+ 1.2% CH (w/v) + 78.5 (µg/ mL) phenylalanine + 10 (µg/ mL) tyrosine +100(µg/ mL)tryptophan	6.613±0.430	2.972±0.193	16.213±0.201	178.784±10.117	452.823±0.287

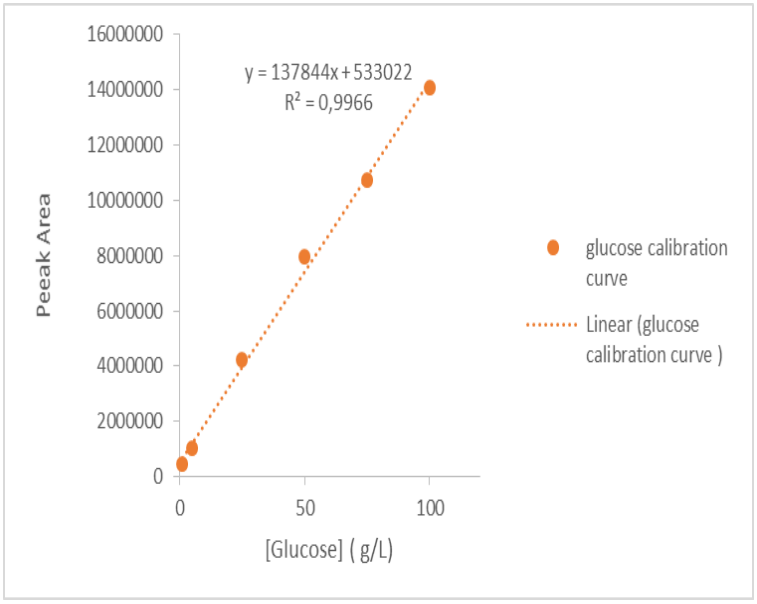
Appendix IV - Result of the best test describe in the literature with greater pDNA concentration.

	Bacterial Growth		Peak Area			PDNA concentration		
Medium	**$OD_{600\ nm}$**	**Cell dry mass (CDM) (g/L)**	**pDNA**	**gDNA**	**RNA**	**pDNA volumetric yield**	**pDNA volumetric yield**	**Purity**
24 g/L yeast extract. 10 g/L glucose. 20 g/L triptone +78.5 (µg/ mL) phenylalanine + 10 (µg/ mL) tyrosine +100(µg/ mL)tryptophan	8.8±0.123	3.96±0.66	38.048±0.554	260.631±15.758	585.371±14.717	68.690±0.793	17.347±0.053	4.300±0.119

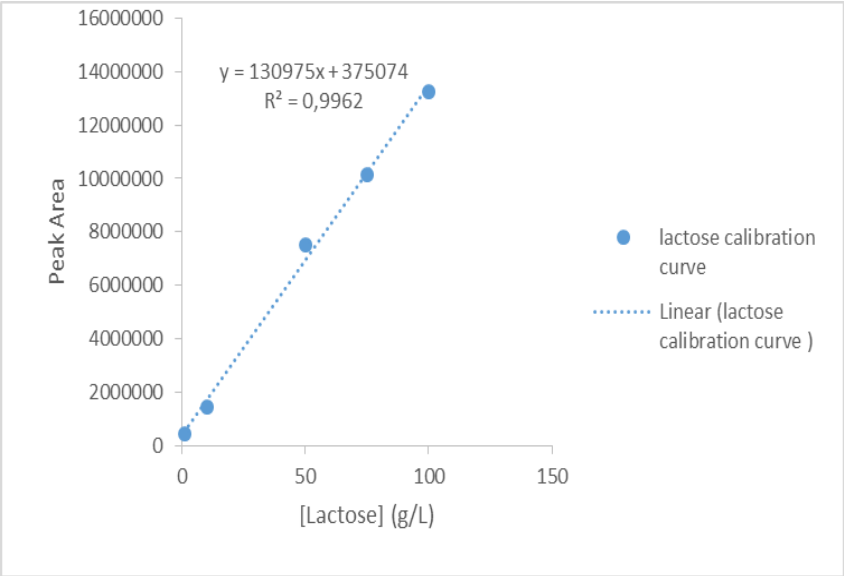
Appendix V - Calibration curve obtained from the correlation between the galactose peak area and the concentration in the range of 1 – 40 g/L.



Appendix VI - Calibration curve obtained from the correlation between the glucose peak area and the concentration in the range of 1 – 100 g/L.

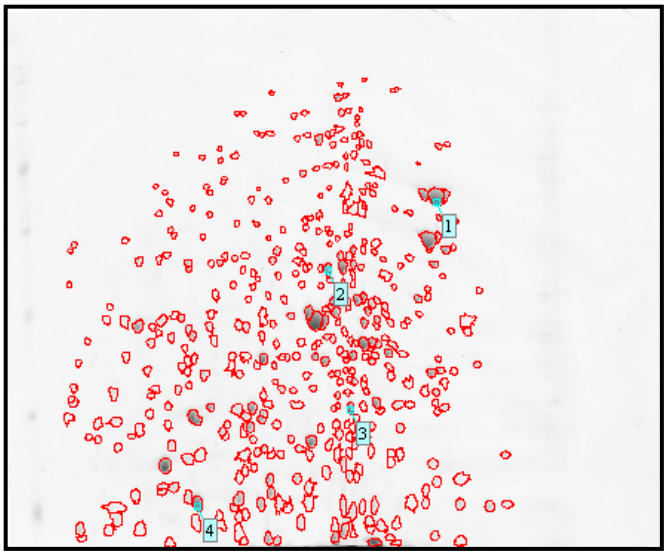


Appendix VII - Calibration curve obtained from the correlation between the lactose peak area and the concentration in the range of 1 – 100 g/L.

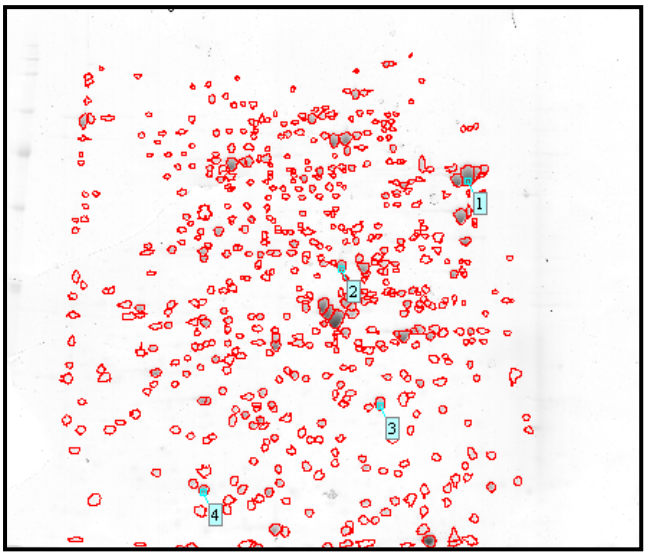


Appendix VIII - 2DE analysis of proteins from *E.coli* VH35 after various growth conditions. Samples were analysed using pH 4-7 linear IPG strips, 18 cm. (second technical duplicate).

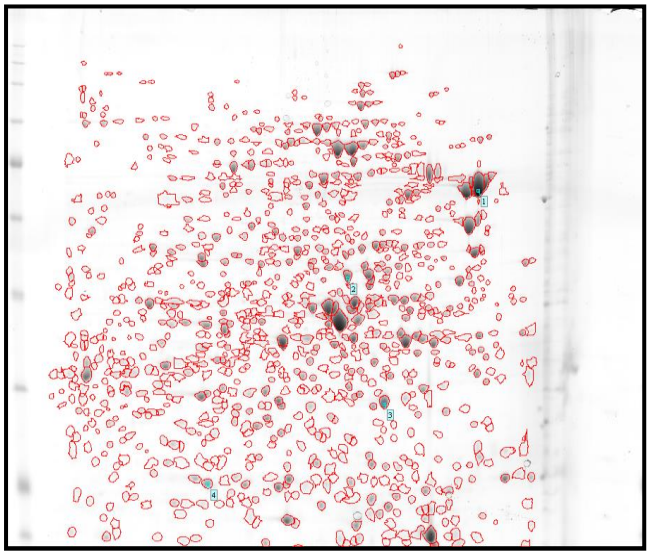
1: 10 g/l yeast extract, glucose, triptone and 78. 5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml) tryptophan;



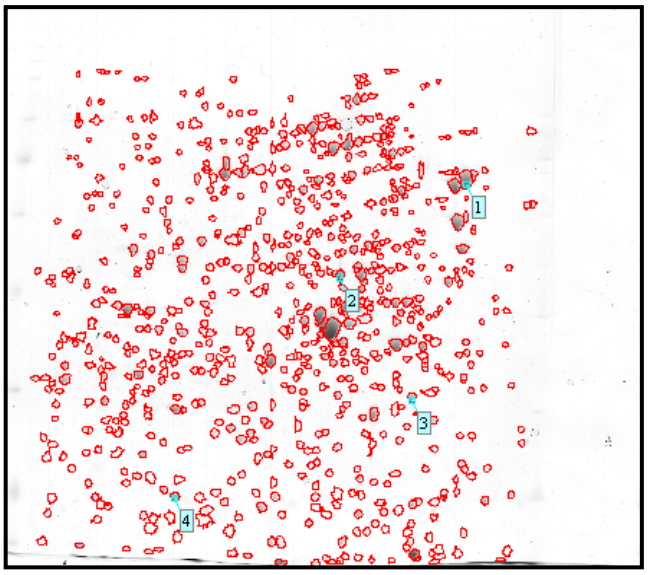
2: CSL 2% (v/v)+ 1,2% CH (w/v)



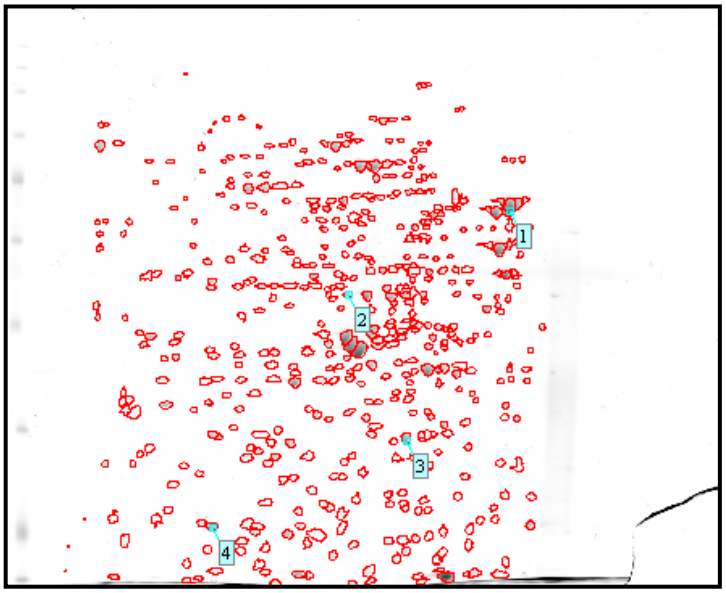
3: CSL 2% (v/v)+ 1.2% CH (w/v)+ + 78,5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan



4: CSL 2% (v/v)+ 1.2% CH (w/v)+5 (g/l) YE



5: CSL 2% (v/v)+ 1.2% CH (w/v)+5 (g/l) YE + 78.5 (µg/ ml) phenylalanine + 10 (µg/ ml) tyrosine +100(µg/ ml)tryptophan.



Appendix IX – ANOVA values and spots overexpressed

- Statistical analysis of the comparison of the growth medium 1 with the growth medium 5

Match ID	*Max*	*Match Count*	*Normalize volume (%) Medium 1*	*Normalize volume (%) Medium 5*	*Anova*
107	0.786039	2	0.033171	0.786039	7.61823e-5
51	0.191445	2	0.01997	0.191445	1.49451e-4
14	1.33305	2	1.33305	0.0781738	2.51919e-4
93	0.203261	2	0.0287876	0.203261	3.0881e-4
133	0.204443	2	0.204443	0.0456968	3.20723e-4
9	0.181293	2	0.181293	0.0153873	6.29955e-4
113	0.227941	2	0.227941	0.0624017	9.18592e-4
75	2.30079	2	2.30079	0.26495	0.00245093
163	0.122814	2	0.0245451	0.122814	0.00256563
130	0.0811792	2	0.0213375	0.0811792	0.00289529
39	0.155308	2	0.155308	0.0224839	0.00290709
103	0.105231	2	0.105231	0.0224088	0.00303066
82	1.22431	2	1.22431	0.270754	0.00407169
77	0.291873	2	0.291873	0.0904653	0.00415418
123	0.629077	2	0.267673	0.629077	0.00485063
61	0.510686	2	0.162288	0.510686	0.00492256
55	0.227914	2	0.0321712	0.227914	0.00529018
23	4.85585	2	4.85585	1.1409	0.00587823
162	0.393045	2	0.139896	0.393045	0.0060058
85	0.217057	2	0.0956645	0.217057	0.00601855
156	0.159785	2	0.159785	0.0772616	0.00672034
205	0.08679	2	0.0030405	0.08679	0.00770979
74	0.395079	2	0.395079	0.202042	0.0093183
201	1.53537	2	0.817512	1.53537	0.00990294
178	1.32061	2	0.510529	1.32061	0.0170923
95	0.107195	2	0.0137228	0.107195	0.0195535
134	0.120002	2	0.0537452	0.120002	0.0199351
159	0.446166	2	0.446166	0.0877714	0.0247657
157	0.396954	2	0.206276	0.396954	0.0249849
184	0.0782094	2	0.0782094	0.023207	0.0267221
176	0.0492081	2	0.0492081	0.0069866	0.0274535
125	2.48727	2	1.49021	2.48727	0.027456

47	0.121234	2	0.0268641	0.121234	0.0291716
68	0.2959	2	0.0611041	0.2959	0.0306461
66	1.47941	2	1.47941	0.909383	0.033769
171	0.311837	2	0.311837	0.0552015	0.0344383
154	0.105389	2	0.0563084	0.105389	0.035708
190	0.211965	2	0.0423534	0.211965	0.0363131
207	0.543395	2	0.0522399	0.543395	0.0383639
12	0.0612912	2	0.0612912	0.0161276	0.0430382
161	0.068504	2	0.0352267	0.068504	0.04449
166	0.164753	2	0.164753	0.0625692	0.0446321
129	0.0944548	2	0.0944548	0.0745297	0.0450803
206	0.0299991	2	0.00265067	0.0299991	0.0461748
170	0.105474	2	0.105474	0.0713926	0.0473861
181	1.11889	2	1.11889	0.184613	0.0479923
196	0.0502335	2	0.0502335	0.0130122	0.0486929
20	0.0565727	2	0.0565727	0.00743748	0.0515931
116	0.198487	2	0.0671921	0.198487	0.0532117
167	0.0299429	2	0.0299429	0.0112381	0.0566943
40	0.226686	2	0.226686	0.0653657	0.0570287
153	0.0653472	2	0.0188969	0.0653472	0.0577048

Statistical analysis of the comparison of the growth medium 3 with the growth medium 5

Match ID	*Max*	*Match Count*	*Normalize volume (%) Medium 3*	*Normalize volume (%) Medium 5*	*Anova*
234	0.605993	2	0.00651352	0.605993	6.60404e-7
277	0.217391	2	0.0394031	0.217391	3.50333e-6
181	0.170629	2	0.170629	0.037467	2.89663e-5
348	0.3417	2	0.0522104	0.3417	3.6807e-5
455	1.53944	2	1.53944	0.0276317	3.82499e-5
20	0.192082	2	0.192082	0.109888	5.31479e-5
98	0.0907396	2	0.0907396	0.0525836	2.0239e-4
191	0.159784	2	0.0203103	0.159784	2.17251e-4
328	0.212984	2	0.0199272	0.212984	3.50618e-4
74	0.220081	2	0.220081	0.106846	5.48941e-4
477	0.208232	2	0.0228207	0.208232	5.68397e-4
226	0.266482	2	0.0325798	0.266482	9.55866e-4
189	0.15809	2	0.0764343	0.15809	0.00100054
261	0.339074	2	0.196578	0.339074	0.00119396
469	0.184117	2	0.0182405	0.184117	0.00141494

262	2.27233	2	2.27233	1.51488	0.00175703
266	0.333083	2	0.175992	0.333083	0.00331275
490	0.115791	2	0.00924298	0.115791	0.00332238
482	0.0334204	2	0.00827762	0.0334204	0.00339005
463	0.0451712	2	0.0228602	0.0451712	0.00346166
243	0.374969	2	0.374969	0.194414	0.00364109
384	0.0118733	2	0.0020718	0.0118733	0.00380246
210	2.49491	2	2.49491	1.23758	0.00468805
264	0.0429731	2	0.0429731	0.00631782	0.00594733
194	0.199046	2	0.199046	0.0554137	0.00657989
451	0.0131296	2	0.00257603	0.0131296	0.00666551
478	0.11039	2	0.0394403	0.11039	0.00696658
72	0.094783	2	0.094783	0.0496935	0.00775502
395	0.0554008	2	0.0554008	0.0195562	0.00802392
383	0.0478518	2	0.0478518	0.00600458	0.00826452
255	1.92292	2	1.92292	0.695886	0.00835198
26	0.119487	2	0.119487	0.0246231	0.00838771
31	0.387723	2	0.387723	0.236148	0.0086027
423	0.132871	2	0.0971284	0.132871	0.00892976
483	0.0722105	2	0.0276474	0.0722105	0.00995723
500	0.0257241	2	0.00502703	0.0257241	0.0103641
322	0.144864	2	0.144864	0.0486741	0.0106441
436	0.125281	2	0.0602508	0.125281	0.01102
95	0.0490451	2	0.00848804	0.0490451	0.01215
67	0.0734305	2	0.0734305	0.044982	0.012163
176	0.0332553	2	0.0332553	0.0103116	0.012454
215	0.47338	2	0.47338	0.1282	0.0126319
256	0.0256273	2	0.00480562	0.0256273	0.0138433
187	0.0497288	2	0.0109584	0.0497288	0.0147648
317	0.229374	2	0.105772	0.229374	0.015145
495	0.0861526	2	0.042925	0.0861526	0.0154315
19	0.912671	2	0.912671	0.258067	0.0162575
101	0.142641	2	0.142641	0.0072164	0.016428
46	0.286998	2	0.119771	0.286998	0.0165048
462	0.0327791	2	0.0130828	0.0327791	0.0176261
69	0.202878	2	0.202878	0.100965	0.0178009
365	0.0708558	2	0.025162	0.0708558	0.0180655
87	0.265924	2	0.265924	0.0747118	0.018242
347	0.141713	2	0.0563971	0.141713	0.0189287
16	0.0775308	2	0.0775308	0.0197867	0.0191563

440	0.0421009	2	0.00940232	0.0421009	0.0205384
270	0.163844	2	0.163844	0.0487552	0.0211116
369	0.204025	2	0.120526	0.204025	0.0217392
325	0.0385483	2	0.00926636	0.0385483	0.0219255
268	0.0887795	2	0.0301552	0.0887795	0.0224735
385	0.0517073	2	0.0124246	0.0517073	0.0239986
461	0.0422423	2	0.0101451	0.0422423	0.0248189
160	0.0561543	2	0.0561543	0.0149077	0.0250985
151	0.127376	2	0.0269615	0.127376	0.0259363
184	0.0608178	2	0.0141892	0.0608178	0.0267015
188	0.0290756	2	0.00730214	0.0290756	0.0272158
489	0.0250391	2	0.0102658	0.0250391	0.0272835
219	0.0480558	2	0.0480558	0.0198166	0.0275799
487	0.0615893	2	0.030096	0.0615893	0.0289327
57	0.512502	2	0.512502	0.387193	0.0293438
453	0.0753565	2	0.0462111	0.0753565	0.0298605
479	0.116767	2	0.0531587	0.116767	0.0303568
82	0.12716	2	0.12716	0.0889144	0.030864
130	0.203866	2	0.203866	0.146604	0.031006
342	0.132117	2	0.132117	0.0389465	0.0311036
77	0.236382	2	0.236382	0.0341368	0.0311393
250	0.0388442	2	0.0388442	0.0268369	0.0312445
302	0.142202	2	0.00912293	0.142202	0.0339038
491	1.07337	2	0.545031	1.07337	0.0342261
367	0.0468399	2	0.0468399	0.0266124	0.0342315
488	0.38749	2	0.166448	0.38749	0.0344051
345	0.110808	2	0.0687347	0.110808	0.0350213
241	0.0708874	2	0.0708874	0.0443871	0.0360004
484	0.110169	2	0.0514428	0.110169	0.0366636
360	0.0677774	2	0.042237	0.0677774	0.0367968
85	0.0678236	2	0.0160552	0.0678236	0.0368924
467	0.0217062	2	0.00513149	0.0217062	0.0370026
336	0.0514873	2	0.0136345	0.0514873	0.0382059
68	0.11833	2	0.11833	0.0928719	0.0397443
314	0.487478	2	0.25537	0.487478	0.0405295
379	0.0237947	2	0.00282457	0.0237947	0.0406501
197	0.0652158	2	0.0652158	0.0324604	0.0407972
349	0.0382803	2	0.0382803	0.0100324	0.0435079
401	0.100146	2	0.100146	0.0363616	0.044558
353	0.0861595	2	0.0423416	0.0861595	0.04471

36	0.0535728	2	0.0535728	0.00397712	0.0482728
282	0.0617422	2	0.0617422	0.0391499	0.0498038
145	0.123675	2	0.123675	0.091134	0.0504033
326	0.0410751	2	0.0115939	0.0410751	0.0524192
249	0.0553802	2	0.0156343	0.0553802	0.0527003
324	0.0923218	2	0.0195415	0.0923218	0.0541958
512	0.0298894	2	0.00895963	0.0298894	0.054775
318	0.0734557	2	0.0734557	0.0501949	0.0548741
138	0.18585	2	0.18585	0.123437	0.054937
29	0.0901046	2	0.0901046	0.0279205	0.056481
229	0.176481	2	0.13184	0.176481	0.0577679
474	0.131862	2	0.0718499	0.131862	0.0590358
354	0.0879584	2	0.0879584	0.0074302	0.0591756
494	0.207301	2	0.107032	0.207301	0.059371
289	0.0248822	2	0.0248822	0.00242228	0.0594585

- Statistical analysis of the comparison of the growth medium 2 with the growth medium 4

Match ID	*Max*	*Match Count*	*Normalize volume (%) Medium 2*	*Normalize volume (%) Medium 4*	*Anova*
258	0.298053	2	0.0156927	0.298053	7.04868e-5
183	1.19115	2	0.0237016	1.19115	1.56774e-4
172	0.992304	2	0.992304	0.795808	0.00404983
203	0.06407	2	0.0107042	0.06407	0.00476951
260	0.167428	2	0.0208562	0.167428	0.00889004
142	0.154171	2	0.154171	0.0277145	0.0117455
284	2.15168	2	0.611858	2.15168	0.0133501
231	0.151785	2	0.0402347	0.151785	0.0237833
95	0.142944	2	0.0290899	0.142944	0.0259895
107	0.0772192	2	0.0171265	0.0772192	0.0267883
174	0.0810536	2	0.0204249	0.0810536	0.031934
17	0.138582	2	0.138582	0.0957871	0.0326017
180	1.04518	2	0.421828	1.04518	0.0330922
48	0.169954	2	0.169954	0.0308216	0.0335705
104	0.0801206	2	0.0200594	0.0801206	0.0345308
8	0.170151	2	0.170151	0.0393	0.0349424
121	0.0926985	2	0.0926985	0.00910955	0.0437891
129	0.0776455	2	0.0776455	0.0305075	0.0439794

267	0.419438	2	0.206475	0.419438	0.0529516
190	0.126203	2	0.0198208	0.126203	0.0529605
158	0.0837688	2	0.0837688	0.0114833	0.057524
192	0.290136	2	0.101914	0.290136	0.0576431
138	0.464964	2	0.464964	0.241597	0.0579386
115	0.0752875	2	0.0325616	0.0752875	0.0587388
291	0.043684	2	0.00863575	0.043684	0.0599276

- Statistical analysis of the comparison of the growth medium 2 with the growth medium 3

Match ID	*Max*	*Match Count*	*Normalize volume (%) Medium 2*	*Normalize volume (%) Medium 3*	*Anova*
234	0.186692	2	0.0288977	0.186692	1.32451e-5
262	0.234537	2	0.0134344	0.234537	2.64523e-5
58	0.220943	2	0.220943	0.0141903	3.3331e-4
208	2.25961	2	1.51201	2.25961	3.35543e-4
265	0.212594	2	0.212594	0.0480059	7.27594e-4
255	0.110583	2	0.110583	0.0297692	8.8835e-4
176	0.135132	2	0.0143176	0.135132	9.75665e-4
120	0.075897	2	0.00485438	0.075897	0.0012937
243	0.0384734	2	0.0384734	0.0055699	0.00153855
20	0.109682	2	0.109682	0.0248633	0.00185487
227	0.0403726	2	0.0280018	0.0403726	0.00209911
140	0.0753026	2	0.00817734	0.0753026	0.00264803
168	1.82793	2	1.2352	1.82793	0.00275667
71	0.106658	2	0.106658	0.0237748	0.00383091
104	0.121845	2	0.00433514	0.121845	0.00406447
198	6.05878	2	4.13279	6.05878	0.00497127
165	0.0976603	2	0.00397649	0.0976603	0.00586592
95	0.245555	2	0.082762	0.245555	0.00623766
46	0.0940664	2	0.0940664	0.0206671	0.00682547
319	2.54961	2	1.82058	2.54961	0.00713511
295	0.0282667	2	0.00544784	0.0282667	0.00744852
35	0.2357	2	0.2357	0.0707375	0.00786289
85	0.0971857	2	0.0971857	0.0185856	0.00824219
352	0.0269661	2	0.0269661	0.00425093	0.0084987
28	0.0478152	2	0.0478152	0.00797031	0.00877873
166	0.467075	2	0.467075	0.160986	0.00906907

90	0.15677	2	0.0428386	0.15677	0.00997999
201	0.347678	2	0.19406	0.347678	0.0103362
94	0.254041	2	0.0602848	0.254041	0.012645
269	0.0391529	2	0.011076	0.0391529	0.0138637
277	0.0265589	2	0.0265589	0.00587278	0.0160614
215	0.0607986	2	0.0607986	0.0133567	0.0167444
236	0.228922	2	0.228922	0.085345	0.0183156
329	0.561651	2	0.561651	0.193762	0.0186393
361	0.0662699	2	0.0662699	0.0301098	0.0212147
53	0.143969	2	0.143969	0.0385127	0.0217316
123	0.0676157	2	0.0467111	0.0676157	0.022812
242	0.0927292	2	0.0927292	0.019685	0.0251941
7	0.132605	2	0.132605	0.0455898	0.0277979
190	0.0443076	2	0.0443076	0.00990175	0.0319709
232	0.486504	2	0.486504	0.284664	0.0347436
362	0.173708	2	0.173708	0.0120912	0.0373943
288	2.74793	2	2.08103	2.74793	0.0402503
16	0.0684282	2	0.00727861	0.0684282	0.0415723
163	0.238275	2	0.238275	0.0926644	0.0441652
149	0.171232	2	0.171232	0.0181233	0.0464177
264	0.102013	2	0.0388944	0.102013	0.0466079
98	0.14791	2	0.14791	0.0515852	0.0467403
50	0.0823264	2	0.0823264	0.020314	0.0476656
37	0.208354	2	0.208354	0.0750793	0.0546644
110	0.0537272	2	0.0321	0.0537272	0.0571426
119	0.353605	2	0.0258717	0.353605	0.0581697
270	0.341056	2	0.341056	0.173472	0.0592357
144	0.171941	2	0.0418221	0.171941	0.0594058

- Statistical analysis of the comparison of the growth medium 4 with the growth medium 5

Match ID	*Max*	*Match Count*	*Normalize volume (%) Medium 4*	*Normalize volume (%) Medium 5*	*Anova*
10	0.18206	2	0.18206	0.0329386	8.49051e-4
206	0.146043	2	0.0971284	0.146043	0.00119079
94	0.823081	2	0.428144	0.823081	0.0103489
124	0.23424	2	0.23424	0.0461457	0.0132819

90	0.071366	2	0.0156343	0.071366	0.0164741
40	0.15421	2	0.0224839	0.15421	0.0178341
192	0.163477	2	0.163477	0.0156503	0.0180614
96	0.13184	2	0.13184	0.0285365	0.018343
115	0.163844	2	0.163844	0.0518453	0.0237596
227	0.18678	2	0.0228207	0.18678	0.0248066
49	0.123675	2	0.123675	0.0422237	0.025596
66	0.789098	2	0.789098	0.278919	0.0296244
178	0.21105	2	0.120526	0.21105	0.032898
22	0.202878	2	0.202878	0.0704831	0.0334073
82	0.307232	2	0.015876	0.307232	0.0342255
18	0.109317	2	0.109317	0.039876	0.0386956
11	0.196474	2	0.11833	0.196474	0.0392745
119	0.0617422	2	0.0617422	0.0219323	0.0448808
143	0.042237	2	0.042237	0.0106241	0.0458064
9	0.204337	2	0.204337	0.0975244	0.0519112
121	0.0523609	2	0.0523609	0.0265683	0.0533641
92	0.0777196	2	0.0398923	0.0777196	0.0596023

Appendix X -List of the total found proteins by MALDI-TOF/TOF using ProteinPilot under a 95% confidence. Protein scores greater than 56 are significant ($p<0.05$).

UniProt ID	*Best Protein Mass*	*Best Protein Score*	*Best Protein Description*
B7NH49	27716	142	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O17:K52:H18 (strain UMN026 / ExPEC) GN=deoC PE=3 SV=
P39377	41058	93	Isoaspartyl dipeptidase OS=*Escherichia coli* (strain K12) GN=iadA PE=1 SV=1
A1A8A5	47836	93	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
P21179	84110	92	Catalase HPII OS=*Escherichia coli* (strain K12) GN=katE PE=1 SV=1
B7UHY9	41925	91	S-adenosylmethionine synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=metK PE=3 SV=1
P21179	84110	87	Catalase HPII OS=*Escherichia coli* (strain K12) GN=katE PE=1 SV=1
P0A9X5	36929	86	Rod shape-determining protein MreB OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=mreB PE=3 SV=1
A7ZSL4	43256	85	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A1A8A5	47836	84	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
P63285	95526	84	Chaperone protein ClpB OS=*Escherichia coli* O157:H7 GN=clpB PE=3 SV=1
A7ZSL4	43256	84	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
P21179	84110	82	Catalase HPII OS=*Escherichia coli* (strain K12) GN=katE PE=1 SV=1
A7ZJD2	41416	81	Galactokinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=galK PE=3 SV=1
A7ZSL4	43256	81	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UR09	27730	78	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=deoC PE=3 SV=1
B7UHY9	41925	75	S-adenosylmethionine synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=metK PE=3 SV=1
P0AFG9	99606	73	Pyruvate dehydrogenase E1 component OS=*Escherichia coli* O157:H7 GN=aceE PE=1 SV=2
P21179	84110	73	Catalase HPII OS=*Escherichia coli* (strain K12) GN=katE PE=1 SV=1
B7UHY9	41925	71	S-adenosylmethionine synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=metK PE=3 SV=1
B7UHY9	41925	70	S-adenosylmethionine synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=metK PE=3 SV=1
P0A911	37178	70	Outer membrane protein A OS=*Escherichia coli* O157:H7 GN=ompA PE=3 SV=1
P0A9X5	36929	70	Rod shape-determining protein MreB OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=mreB PE=3 SV=1
P00509	43546	68	Aspartate aminotransferase OS=*Escherichia coli* (strain K12) GN=aspC PE=1 SV=1
A7ZJD2	41416	68	Galactokinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=galK PE=3 SV=1
B7UK50	77532	67	Elongation factor G OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=fusA PE=3 SV=1

A7ZSL4	43256	66	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZTG6	56074	65	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
P0AFG9	99606	64	Pyruvate dehydrogenase E1 component OS=*Escherichia coli* O157:H7 GN=aceE PE=1 SV=2
P0ABZ8	47254	63	Chaperone SurA OS=*Escherichia coli* O157:H7 GN=surA PE=1 SV=1
P0A6F2	41405	63	Carbamoyl-phosphate synthase small chain OS=*Escherichia coli* O157:H7 GN=carA PE=3 SV=1
B7UHY9	41925	63	S-adenosylmethionine synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=metK PE=3 SV=1
P39377	41058	62	Isoaspartyl dipeptidase OS=*Escherichia coli* (strain K12) GN=iadA PE=1 SV=1
P00895	57458	62	Anthranilate synthase component 1 OS=*Escherichia coli* (strain K12) GN=trpE PE=1 SV=2
A7ZTG6	56074	61	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
B7UK50	77532	59	Elongation factor G OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=fusA PE=3 SV=1
A7ZSL4	43256	57	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZSL4	43256	56	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZQ54	21727	55	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
A7ZSL4	43256	55	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZTG6	56074	54	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
P63285	95526	54	Chaperone protein ClpB OS=*Escherichia coli* O157:H7 GN=clpB PE=3 SV=1
P14175	44135	53	Glycine betaine/proline betaine transport system ATP-binding protein ProV OS=*Escherichia coli* (strain K12) GN=proV PE=1 SV=1
P0A6A5	43263	52	Acetate kinase OS=*Escherichia coli* O157:H7 GN=ackA PE=3 SV=1
A7ZQ54	21727	52	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
P14175	44135	50	Glycine betaine/proline betaine transport system ATP-binding protein ProV OS=*Escherichia coli* (strain K12) GN=proV PE=1 SV=1
A7ZJD2	41416	49	Galactokinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=galK PE=3 SV=1
P39377	41058	49	Isoaspartyl dipeptidase OS=*Escherichia coli* (strain K12) GN=iadA PE=1 SV=1
A7ZSL4	43256	49	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UK50	77532	49	Elongation factor G OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=fusA PE=3 SV=1
B7UHJ6	60338	48	CTP synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=pyrG PE=3 SV=1
P14175	44135	47	Glycine betaine/proline betaine transport system ATP-binding protein ProV OS=*Escherichia coli* (strain K12) GN=proV PE=1 SV=1

P0A9X5	36929	47	Rod shape-determining protein MreB OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=mreB PE=3 SV=1
P09169	35540	46	Protease 7 OS=*Escherichia coli* (strain K12) GN=ompT PE=1 SV=1
B7UJQ9	48163	46	Trigger factor OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=tig PE=3 SV=1
P0A911	37178	46	Outer membrane protein A OS=*Escherichia coli* O157:H7 GN=ompA PE=3 SV=1
A7ZSL4	43256	46	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
P0AFN1	8758	45	Phage shock protein B OS=*Escherichia coli* O157:H7 GN=pspB PE=3 SV=1
B7UMJ7	50294	45	ATP synthase subunit beta OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=atpD PE=3 SV=1
B7UR09	27730	45	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=deoC PE=3 SV=1
P21179	84110	45	Catalase HPII OS=*Escherichia coli* (strain K12) GN=katE PE=1 SV=1
A7ZHA4	69072	45	Chaperone protein DnaK OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=dnaK PE=3 SV=1
B7UJ66	49853	44	Argininosuccinate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=argG PE=3 SV=1
P63285	95526	44	Chaperone protein ClpB OS=*Escherichia coli* O157:H7 GN=clpB PE=3 SV=1
A7ZSL4	43256	44	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZQ54	21727	43	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
A7ZSL4	43256	43	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
P00509	43546	42	Aspartate aminotransferase OS=*Escherichia coli* (strain K12) GN=aspC PE=1 SV=1
A7ZHA4	69072	42	Chaperone protein DnaK OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=dnaK PE=3 SV=1
P69911	52634	42	Glutamate decarboxylase beta OS=*Escherichia coli* O157:H7 GN=gadB PE=3 SV=1
A7ZSL4	43256	42	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A1A8A5	47836	42	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
A7ZSL4	43256	42	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZSL4	43256	41	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZSL4	43256	40	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
P58228	52665	40	Glutamate decarboxylase alpha OS=*Escherichia coli* O157:H7 GN=gadA PE=3 SV=1
A7ZTG6	56074	40	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
B7UHV3	40123	40	Aminomethyltransferase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=gcvT PE=3 SV=1
A7ZSL4	43256	39	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UJ66	49853	38	Argininosuccinate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=argG PE=3 SV=1
P63285	95526	38	Chaperone protein ClpB OS=*Escherichia coli* O157:H7 GN=clpB PE=3 SV=1

P0AAI7	43019	38	3-oxoacyl-[acyl-carrier-protein] synthase 2 OS=*Escherichia coli* O157:H7 GN=fabF PE=3 SV=2
A1A8A5	47836	38	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
A7ZTU8	17253	36	ATP synthase subunit b OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=atpF PE=3 SV=1
A7ZR34	41093	36	Phosphoglycerate kinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=pgk PE=3 SV=1
P08200	45727	36	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P08200	45727	35	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
B7UHJ6	60338	35	CTP synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=pyrG PE=3 SV=1
B7NK50	52857	34	NAD/NADP-dependent betaine aldehyde dehydrogenase OS=*Escherichia coli* O7:K1 (strain IAI39 / ExPEC) GN=betB PE=3 SV=1
A7ZTU8	17253	34	ATP synthase subunit b OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=atpF PE=3 SV=1
A7ZSL4	43256	34	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
P08200	45727	34	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
B7UJ66	49853	34	Argininosuccinate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=argG PE=3 SV=1
P37759	40533	34	dTDP-glucose 4,6-dehydratase 1 OS=*Escherichia coli* (strain K12) GN=rfbB PE=3 SV=2
B7UIV5	23325	33	3,4-dihydroxy-2-butanone 4-phosphate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ribB PE=3 SV=1
B7UFF9	31547	33	Cytidine deaminase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=cdd PE=3 SV=1
P77288	19839	32	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
A7ZQ54	21727	32	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
B7UHJ6	60338	32	CTP synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=pyrG PE=3 SV=1
P0AFM5	11679	32	Phosphate starvation-inducible protein PsiF OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=psiF PE=3 SV=1
P77288	19839	32	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P0A9B4	35510	32	Glyceraldehyde-3-phosphate dehydrogenase A OS=*Escherichia coli* O157:H7 GN=gapA PE=3 SV=2
B7UHV3	40123	32	Aminomethyltransferase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=gcvT PE=3 SV=1
P39361	29249	31	Putative sgc region transcriptional regulator OS=*Escherichia coli* (strain K12) GN=sgcR PE=3 SV=1
B7UR09	27730	30	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=deoC PE=3 SV=1
P62584	8761	30	P fimbrial regulatory protein KS71A OS=*Escherichia coli* GN=KS71A PE=1 SV=1
P00350	51449	30	6-phosphogluconate dehydrogenase, decarboxylating OS=*Escherichia coli* (strain K12) GN=gnd PE=1 SV=2
B7UHV3	40123	30	Aminomethyltransferase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=gcvT PE=3 SV=1
Q5JBK6	9381	29	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2

P0AC71	12870	29	Glutaredoxin-4 OS=*Escherichia coli* O157:H7 GN=grxD PE=3 SV=1
P41039	9750	29	Uncharacterized protein YbiI OS=*Escherichia coli* (strain K12) GN=ybiI PE=1 SV=1
A7ZSL4	43256	29	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UIV5	23325	29	3,4-dihydroxy-2-butanone 4-phosphate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ribB PE=3 SV=1
P76180	16308	29	Inner membrane protein YdgK OS=*Escherichia coli* (strain K12) GN=ydgK PE=1 SV=2
P58228	52665	29	Glutamate decarboxylase alpha OS=*Escherichia coli* O157:H7 GN=gadA PE=3 SV=1
A1A8A5	47836	29	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
B7UR09	27730	28	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=deoC PE=3 SV=1
Q5JBK6	9381	28	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
B7UFL0	17687	28	Cytochrome c-type biogenesis protein CcmE OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ccmE PE=3 SV=1
B7UFF9	31547	28	Cytidine deaminase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=cdd PE=3 SV=1
P08200	45727	28	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P21361	5999	28	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P08200	45727	28	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
B7NW39	36110	28	tRNA-modifying protein YgfZ OS=*Escherichia coli* O7:K1 (strain IAI39 / ExPEC) GN=ygfZ PE=3 SV=1
Q8FF44	65552	27	Chaperone protein HscA OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=hscA PE=3 SV=2
Q8FIM4	26202	27	NAD-dependent protein deacylase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=cobB PE=3 SV=2
P09147	37242	27	UDP-glucose 4-epimerase OS=*Escherichia coli* (strain K12) GN=galE PE=1 SV=1
B7UIV5	23325	27	3,4-dihydroxy-2-butanone 4-phosphate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ribB PE=3 SV=1
P16528	29720	27	Acetate operon repressor OS=*Escherichia coli* (strain K12) GN=iclR PE=1 SV=1
P21361	5999	27	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P09126	27780	26	Uroporphyrinogen-III synthase OS=*Escherichia coli* (strain K12) GN=hemD PE=3 SV=1
P0A6A5	43263	26	Acetate kinase OS=*Escherichia coli* O157:H7 GN=ackA PE=3 SV=1
A7ZSL4	43256	26	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UIV5	23325	26	3,4-dihydroxy-2-butanone 4-phosphate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ribB PE=3 SV=1
A7ZR34	41093	26	Phosphoglycerate kinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=pgk PE=3 SV=1
B7UJQ9	48163	26	Trigger factor OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=tig PE=3 SV=1

P08200	45727	26	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P21361	5999	26	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P58228	52665	26	Glutamate decarboxylase alpha OS=*Escherichia coli* O157:H7 GN=gadA PE=3 SV=1
A7ZP09	10687	26	50S ribosomal protein L25 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=rplY PE=3 SV=1
B7UPW3	57293	26	60 kDa chaperonin OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=groL PE=3 SV=1
A7ZTG6	56074	26	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
P77288	19839	25	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
B1J0W4	21801	25	HTH-type transcriptional regulator BetI OS=*Escherichia coli* (strain ATCC 8739 / DSM 1576 / Crooks) GN=betI PE=3 SV=1
B1J0W4	21801	25	HTH-type transcriptional regulator BetI OS=*Escherichia coli* (strain ATCC 8739 / DSM 1576 / Crooks) GN=betI PE=3 SV=1
Q5JBK6	9381	25	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
P30852	40929	25	Protein smf OS=*Escherichia coli* (strain K12) GN=smf PE=3 SV=1
A7ZSL4	43256	25	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
A7ZSL4	43256	25	Elongation factor Tu 1 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=tuf1 PE=3 SV=1
B7UPW3	57293	25	60 kDa chaperonin OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=groL PE=3 SV=1
P0A9T7	13163	25	Uncharacterized HTH-type transcriptional regulator YbaQ OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=ybaQ PE=3 SV=1
P69911	52634	25	Glutamate decarboxylase beta OS=*Escherichia coli* O157:H7 GN=gadB PE=3 SV=1
Q8FG00	12015	25	UPF0339 protein YegP OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=yegP PE=3 SV=2
P75860	18500	24	Uncharacterized fimbrial-like protein YcbV OS=*Escherichia coli* (strain K12) GN=ycbV PE=2 SV=2
P0AGB8	21682	24	ECF RNA polymerase sigma-E factor OS=*Escherichia coli* O157:H7 GN=rpoE PE=3 SV=1
P76160	12083	24	Uncharacterized protein YdfR OS=*Escherichia coli* (strain K12) GN=ydfR PE=4 SV=1
A1A8A5	47836	24	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
P21361	5999	24	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P0AFG9	99606	24	Pyruvate dehydrogenase E1 component OS=*Escherichia coli* O157:H7 GN=aceE PE=1 SV=2
P08200	45727	24	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
A1A8A5	47836	24	Trigger factor OS=*Escherichia coli* O1:K1 / APEC GN=tig PE=3 SV=1
B7UHV1	104227	24	Glycine dehydrogenase (decarboxylating) OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=gcvP PE=3 SV=1
P41072	9121	24	Protein TrbG OS=*Escherichia coli* (strain K12) GN=trbG PE=4 SV=1

A7ZQ54	21727	23	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
P0A6F2	41405	23	Carbamoyl-phosphate synthase small chain OS=*Escherichia coli* O157:H7 GN=carA PE=3 SV=1
Q1R1U5	9322	23	UPF0401 protein YubL OS=*Escherichia coli* (strain UTI89 / UPEC) GN=yubL PE=3 SV=1
Q1R1U5	9322	23	UPF0401 protein YubL OS=*Escherichia coli* (strain UTI89 / UPEC) GN=yubL PE=3 SV=1
P09147	37242	23	UDP-glucose 4-epimerase OS=*Escherichia coli* (strain K12) GN=galE PE=1 SV=1
P07021	17234	23	Putative lipoprotein YfiB OS=*Escherichia coli* (strain K12) GN=yfiB PE=3 SV=2
P76001	13405	23	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
P09883	61550	23	Colicin-E9 OS=*Escherichia coli* GN=col PE=1 SV=4
A7ZMI2	36809	23	Phenylalanine--tRNA ligase alpha subunit OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=pheS PE=3 SV=1
Q8FJ70	10251	22	Acylphosphatase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=yccX PE=3 SV=2
Q59408	17914	22	Dihydrofolate reductase type A13 OS=*Escherichia coli* GN=dfrA13 PE=3 SV=1
Q5JBK6	9381	22	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
Q1R1U5	9322	22	UPF0401 protein YubL OS=*Escherichia coli* (strain UTI89 / UPEC) GN=yubL PE=3 SV=1
Q59408	17914	22	Dihydrofolate reductase type A13 OS=*Escherichia coli* GN=dfrA13 PE=3 SV=1
P0ACW9	5789	22	Uncharacterized protein YdfA OS=*Escherichia coli* O157:H7 GN=ydfA PE=4 SV=1
P16528	29720	22	Acetate operon repressor OS=*Escherichia coli* (strain K12) GN=iclR PE=1 SV=1
P23839	33154	22	UPF0701 protein YicC OS=*Escherichia coli* (strain K12) GN=yicC PE=1 SV=2
P0A9D3	22854	22	Glutathione S-transferase GstA OS=*Escherichia coli* O157:H7 GN=gstA PE=3 SV=1
P77288	19839	22	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P03845	4859	22	Putative uncharacterized protein 1 OS=*Escherichia coli* PE=4 SV=1
B7UPW3	57293	22	60 kDa chaperonin OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=groL PE=3 SV=1
P03853	9208	22	Uncharacterized 9.2 kDa protein OS=*Escherichia coli* PE=4 SV=1
P63884	45606	21	N-acetylmuramoyl-L-alanine amidase AmiC OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=amiC PE=3 SV=1
B7UR09	27730	21	Deoxyribose-phosphate aldolase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=deoC PE=3 SV=1
B7UKF4	23527	21	Adenylate kinase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=adk PE=3 SV=1
P09147	37242	21	UDP-glucose 4-epimerase OS=*Escherichia coli* (strain K12) GN=galE PE=1 SV=1
Q8XEC3	18779	21	Chorismate pyruvate-lyase OS=*Escherichia coli* O157:H7 GN=ubiC PE=3 SV=2
B7UIV5	23325	21	3,4-dihydroxy-2-butanone 4-phosphate synthase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=ribB PE=3 SV=1

Q0TFJ0	22379	21	Lipopolysaccharide core heptose(II)-phosphate phosphatase OS=*Escherichia coli* O6:K15:H31 (strain 536 / UPEC) GN=ais PE=3 SV=1
P77288	19839	21	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P0ACW9	5789	21	Uncharacterized protein YdfA OS=*Escherichia coli* O157:H7 GN=ydfA PE=4 SV=1
P0AC14	30596	21	Dihydropteroate synthase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=folP PE=3 SV=1
A7ZR34	41093	21	Phosphoglycerate kinase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=pgk PE=3 SV=1
P0A9G7	47492	21	Isocitrate lyase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=aceA PE=3 SV=1
P08200	45727	21	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P08200	45727	21	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P64601	19659	21	Uncharacterized protein YhbT OS=*Escherichia coli* O157:H7 GN=yhbT PE=4 SV=1
Q8XDZ4	18159	21	Periplasmic chaperone Spy OS=*Escherichia coli* O157:H7 GN=spy PE=1 SV=1
P08200	45727	21	Isocitrate dehydrogenase [NADP] OS=*Escherichia coli* (strain K12) GN=icd PE=1 SV=1
P77288	19839	21	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P0AD57	35195	21	Octaprenyl-diphosphate synthase OS=*Escherichia coli* (strain K12) GN=ispB PE=1 SV=1
P03853	9208	21	Uncharacterized 9.2 kDa protein OS=*Escherichia coli* PE=4 SV=1
P41072	9121	21	Protein TrbG OS=*Escherichia coli* (strain K12) GN=trbG PE=4 SV=1
P63884	45606	20	N-acetylmuramoyl-L-alanine amidase AmiC OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=amiC PE=3 SV=1
P62573	17244	20	Homoprotocatechuate degradative operon repressor OS=*Escherichia coli* GN=hpcR PE=3 SV=1
Q8FL67	53267	20	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=murE PE=3 SV=3
P76001	13405	20	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
Q5JBK6	9381	20	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
P0ACG4	5498	20	Protein HokC OS=*Escherichia coli* (strain K12) GN=hokC PE=1 SV=1
Q1R1U5	9322	20	UPF0401 protein YubL OS=*Escherichia coli* (strain UTI89 / UPEC) GN=yubL PE=3 SV=1
B7UJ78	9119	20	50S ribosomal protein L27 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpmA PE=3 SV=1
Q52772	11305	20	Protein KleF OS=*Escherichia coli* GN=kleF PE=4 SV=1
Q2EES1	4112	20	Uncharacterized protein YniD OS=*Escherichia coli* (strain K12) GN=yniD PE=1 SV=2
B7UGZ1	45288	20	Serine hydroxymethyltransferase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=glyA PE=3 SV=1
P19778	11139	20	Insertion element IS2 uncharacterized 11.1 kDa protein OS=*Escherichia coli* PE=4 SV=1
P75970	9281	20	Excisionase-like protein from lambdoid prophage 14 OS=*Escherichia coli* (strain K12) GN=xisE PE=4 SV=1

A7ZI30	5463	20	50S ribosomal protein L36 2 OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=rpmJ2 PE=3 SV=1
P67664	34494	20	HTH-type transcriptional activator AaeR OS=*Escherichia coli* O157:H7 GN=aaeR PE=3 SV=1
P08400	49598	20	Phosphate regulon sensor protein PhoR OS=*Escherichia coli* (strain K12) GN=phoR PE=1 SV=1
P0A7B0	19691	20	Inorganic pyrophosphatase OS=*Escherichia coli* O157:H7 GN=ppa PE=3 SV=2
P42615	14163	20	Modulator protein MzrA OS=*Escherichia coli* (strain K12) GN=mzrA PE=1 SV=1
Q8L0V4	79207	20	Chondroitin synthase OS=*Escherichia coli* GN=kfoC PE=1 SV=1
P17776	1541	20	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P03845	4859	20	Putative uncharacterized protein 1 OS=*Escherichia coli* PE=4 SV=1
P52604	42669	20	Protein KlcB OS=*Escherichia coli* GN=klcB PE=4 SV=2
B7M9H2	24883	19	Iron-sulfur cluster repair protein YtfE OS=*Escherichia coli* O8 (strain IAI1) GN=ytfE PE=3 SV=1
P19767	9918	19	Insertion element IS1 7 protein InsA OS=*Escherichia coli* (strain K12) GN=insA7 PE=3 SV=1
P10739	3801	19	23S rRNA methylase leader peptide OS=*Escherichia coli* GN=ermC PE=4 SV=1
Q9JMR5	19497	19	Uncharacterized protein YubH OS=*Escherichia coli* (strain K12) GN=yubH PE=3 SV=2
B1ITA4	36201	19	tRNA-modifying protein YgfZ OS=*Escherichia coli* (strain ATCC 8739 / DSM 1576 / Crooks) GN=ygfZ PE=3 SV=1
P76001	13405	19	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
Q8L0V4	79207	19	Chondroitin synthase OS=*Escherichia coli* GN=kfoC PE=1 SV=1
Q8XA44	18723	19	tRNA-specific adenosine deaminase OS=*Escherichia coli* O157:H7 GN=tadA PE=3 SV=2
P0A436	10188	19	Galactitol-specific phosphotransferase enzyme IIB component OS=*Escherichia coli* O157:H7 GN=gatB PE=3 SV=1
P0ACG7	5733	19	Protein HokD OS=*Escherichia coli* O157:H7 GN=hokD PE=3 SV=1
B7UIX3	8495	19	30S ribosomal protein S21 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsU PE=3 SV=1
B7UIX3	8495	19	30S ribosomal protein S21 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsU PE=3 SV=1
B7UK35	9698	19	30S ribosomal protein S17 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsQ PE=3 SV=1
Q9JMT7	16130	19	Uncharacterized protein YuaC OS=*Escherichia coli* (strain K12) GN=yuaC PE=4 SV=1
Q9JMT7	16130	19	Uncharacterized protein YuaC OS=*Escherichia coli* (strain K12) GN=yuaC PE=4 SV=1
P42615	14163	19	Modulator protein MzrA OS=*Escherichia coli* (strain K12) GN=mzrA PE=1 SV=1
P75692	8885	19	Uncharacterized protein YahM OS=*Escherichia coli* (strain K12) GN=yahM PE=4 SV=2
P75692	8885	19	Uncharacterized protein YahM OS=*Escherichia coli* (strain K12) GN=yahM PE=4 SV=2
P0A845	7020	19	Probable Sec-independent protein translocase protein TatE OS=*Escherichia coli* O157:H7 GN=tatE PE=3 SV=1
P13478	9600	19	Colicin-E8 immunity protein in ColE6 OS=*Escherichia coli* GN=imm PE=3 SV=1
P15043	68320	19	ATP-dependent DNA helicase RecQ OS=*Escherichia coli* (strain K12) GN=recQ PE=1 SV=5

P0A845	7020	19	Probable Sec-independent protein translocase protein TatE OS=*Escherichia coli* O157:H7 GN=tatE PE=3 SV=1
Q00192	82145	19	Protein TraI OS=*Escherichia coli* GN=traI PE=4 SV=1
A7ZTG6	56074	19	2,3-bisphosphoglycerate-independent phosphoglycerate mutase OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=gpmI PE=3 SV=1
P0ACW6	8873	19	Uncharacterized protein YdcH OS=*Escherichia coli* (strain K12) GN=ydcH PE=4 SV=2
C1P614	4788	18	Uncharacterized protein YqfG OS=*Escherichia coli* (strain K12) GN=yqfG PE=1 SV=1
P0AAM9	8802	18	Hydrogenase-2 operon protein HybG OS=*Escherichia coli* O157:H7 GN=hybG PE=3 SV=1
P76001	13405	18	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
P08407	18557	18	Fimbrial protein PapE OS=*Escherichia coli* GN=papE PE=1 SV=1
P08407	18557	18	Fimbrial protein PapE OS=*Escherichia coli* GN=papE PE=1 SV=1
P0AE10	20748	18	Alkyl hydroperoxide reductase subunit C OS=*Escherichia coli* O157:H7 GN=ahpC PE=3 SV=2
P0ADX0	7594	18	Uncharacterized protein YheV OS=*Escherichia coli* O157:H7 GN=yheV PE=4 SV=1
P56614	5880	18	Uncharacterized protein YmdF OS=*Escherichia coli* (strain K12) GN=ymdF PE=3 SV=1
P09551	27974	18	Lysine/arginine/ornithine-binding periplasmic protein OS=*Escherichia coli* (strain K12) GN=argT PE=1 SV=3
B7UIX3	8495	18	30S ribosomal protein S21 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsU PE=3 SV=1
B7UIX3	8495	18	30S ribosomal protein S21 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsU PE=3 SV=1
P18353	12579	18	Protein TrbJ OS=*Escherichia coli* (strain K12) GN=trbJ PE=4 SV=3
Q2EES1	4112	18	Uncharacterized protein YniD OS=*Escherichia coli* (strain K12) GN=yniD PE=1 SV=2
P76001	13405	18	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
Q52772	11305	18	Protein KleF OS=*Escherichia coli* GN=kleF PE=4 SV=1
P16528	29720	18	Acetate operon repressor OS=*Escherichia coli* (strain K12) GN=iclR PE=1 SV=1
P0AFW3	6503	18	Ribosome modulation factor OS=*Escherichia coli* O157:H7 GN=rmf PE=3 SV=1
P0ADX0	7594	18	Uncharacterized protein YheV OS=*Escherichia coli* O157:H7 GN=yheV PE=4 SV=1
Q2EES1	4112	18	Uncharacterized protein YniD OS=*Escherichia coli* (strain K12) GN=yniD PE=1 SV=2
Q5JBK6	9381	18	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
B7UPW3	57293	18	60 kDa chaperonin OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=groL PE=3 SV=1
P0AAQ8	13309	18	Uncharacterized protein YbaA OS=*Escherichia coli* O157:H7 GN=ybaA PE=4 SV=1
P75692	8885	18	Uncharacterized protein YahM OS=*Escherichia coli* (strain K12) GN=yahM PE=4 SV=2
P77288	19839	18	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
O82900	21713	18	Uncharacterized protein YubH OS=*Escherichia coli* O157:H7 GN=yubH PE=4 SV=2
P0AEL5	8366	18	Ferrous iron transport protein A OS=*Escherichia coli* O157:H7 GN=feoA PE=3 SV=1

Q9S4X2	25430	18	Putative methylase YubD OS=*Escherichia coli* (strain K12) GN=yubD PE=3 SV=1
P17776	1541	18	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P41072	9121	18	Protein TrbG OS=*Escherichia coli* (strain K12) GN=trbG PE=4 SV=1
P0AB63	9380	18	Protein YciN OS=*Escherichia coli* O157:H7 GN=yciN PE=4 SV=1
P68207	8320	17	UPF0337 protein YjbJ OS=*Escherichia coli* O157:H7 GN=yjbJ PE=3 SV=1
P0AE65	8939	17	Cation transport regulator ChaB OS=*Escherichia coli* O157:H7 GN=chaB PE=4 SV=1
A7ZQ54	21727	17	Protein GrpE OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=grpE PE=3 SV=1
A7ZVT5	12347	17	Trp operon repressor OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=trpR PE=3 SV=1
B1IS55	17053	17	UPF0442 protein YjjB OS=*Escherichia coli* (strain ATCC 8739 / DSM 1576 / Crooks) GN=yjjB PE=3 SV=1
P69793	12739	17	N,N'-diacetylchitobiose-specific phosphotransferase enzyme IIA component OS=*Escherichia coli* O157:H7 GN=chbA PE=3 SV=1
P56614	5880	17	Uncharacterized protein YmdF OS=*Escherichia coli* (strain K12) GN=ymdF PE=3 SV=1
P10739	3801	17	23S rRNA methylase leader peptide OS=*Escherichia coli* GN=ermC PE=4 SV=1
P10739	3801	17	23S rRNA methylase leader peptide OS=*Escherichia coli* GN=ermC PE=4 SV=1
P58217	9194	17	DinI-like protein Z2083/ECs2153 OS=*Escherichia coli* O157:H7 GN=Z2083 PE=4 SV=1
P0AC14	30596	17	Dihydropteroate synthase OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=folP PE=3 SV=1
P05834	6009	17	Bacteriocin microcin B17 OS=*Escherichia coli* GN=mcbA PE=1 SV=1
P76001	13405	17	Uncharacterized protein YcgJ OS=*Escherichia coli* (strain K12) GN=ycgJ PE=3 SV=1
B7UIX3	8495	17	30S ribosomal protein S21 OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=rpsU PE=3 SV=1
P0ABY0	17210	17	Flagellar protein FliL OS=*Escherichia coli* O157:H7 GN=fliL PE=3 SV=1
Q8X616	25035	17	Uncharacterized ferredoxin-like protein YdhX OS=*Escherichia coli* O157:H7 GN=ydhX PE=3 SV=3
P28697	21211	16	Putative uncharacterized protein HtgA OS=*Escherichia coli* (strain K12) GN=htgA PE=5 SV=2
Q5JBK6	9381	16	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
P21361	5999	16	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P18353	12579	16	Protein TrbJ OS=*Escherichia coli* (strain K12) GN=trbJ PE=4 SV=3
P05835	8537	16	Relaxosome protein TraY OS=*Escherichia coli* GN=traY PE=3 SV=1
P05835	8537	16	Relaxosome protein TraY OS=*Escherichia coli* GN=traY PE=3 SV=1
P75992	10338	16	Probable two-component-system connector protein YmgA OS=*Escherichia coli* (strain K12) GN=ymgA PE=2 SV=1
P17776	1541	16	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1

P21361	5999	16	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
Q5JBK6	9381	16	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
B7UFF9	31547	16	Cytidine deaminase OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=cdd PE=3 SV=1
P0AFR5	23197	16	Uncharacterized protein YciO OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=yciO PE=3 SV=1
Q9JMR5	19497	16	Uncharacterized protein YubH OS=*Escherichia coli* (strain K12) GN=yubH PE=3 SV=2
P05834	6009	16	Bacteriocin microcin B17 OS=*Escherichia coli* GN=mcbA PE=1 SV=1
P07024	60786	16	Protein UshA OS=*Escherichia coli* (strain K12) GN=ushA PE=1 SV=2
P77288	19839	16	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P77288	19839	16	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P77288	19839	16	Uncharacterized fimbrial-like protein YfcV OS=*Escherichia coli* (strain K12) GN=yfcV PE=2 SV=1
P0A9T7	13163	16	Uncharacterized HTH-type transcriptional regulator YbaQ OS=*Escherichia coli* O6:H1 (strain CFT073 / ATCC 700928 / UPEC) GN=ybaQ PE=3 SV=1
Q5JBK6	9381	15	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
P64576	10938	15	Uncharacterized protein YghW OS=*Escherichia coli* O157:H7 GN=yghW PE=4 SV=1
P0ACW6	8873	15	Uncharacterized protein YdcH OS=*Escherichia coli* (strain K12) GN=ydcH PE=4 SV=2
Q9JMR5	19497	15	Uncharacterized protein YubH OS=*Escherichia coli* (strain K12) GN=yubH PE=3 SV=2
P64477	7262	15	Uncharacterized protein YdiH OS=*Escherichia coli* O157:H7 GN=ydiH PE=4 SV=2
P05835	8537	15	Relaxosome protein TraY OS=*Escherichia coli* GN=traY PE=3 SV=1
P05835	8537	15	Relaxosome protein TraY OS=*Escherichia coli* GN=traY PE=3 SV=1
A7ZHZ2	39464	15	DNA polymerase IV OS=*Escherichia coli* O139:H28 (strain E24377A / ETEC) GN=dinB PE=3 SV=1
P21361	5999	15	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	15	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P75679	15240	15	Transposase InsN for insertion sequence element IS911A OS=*Escherichia coli* (strain K12) GN=insN1 PE=3 SV=1
P75679	15240	15	Transposase InsN for insertion sequence element IS911A OS=*Escherichia coli* (strain K12) GN=insN1 PE=3 SV=1
P17776	1541	15	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P21361	5999	15	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	15	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P10739	3801	14	23S rRNA methylase leader peptide OS=*Escherichia coli* GN=ermC PE=4 SV=1
P62584	8761	14	P fimbrial regulatory protein KS71A OS=*Escherichia coli* GN=KS71A PE=1 SV=1
P45848	71244	14	Uncharacterized protein YciQ OS=*Escherichia coli* (strain K12) GN=yciQ PE=4 SV=2

B7ULB9	76765	14	Glycine--tRNA ligase beta subunit OS=*Escherichia coli* O127:H6 (strain E2348/69 / EPEC) GN=glyS PE=3 SV=1
P0ABY0	17210	14	Flagellar protein FliL OS=*Escherichia coli* O157:H7 GN=fliL PE=3 SV=1
P17776	1541	14	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P21361	5999	14	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P17776	1541	14	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P04743	8849	14	Major pilus subunit operon regulatory protein OS=*Escherichia coli* GN=papI PE=4 SV=1
P21361	5999	14	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	14	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P62671	6104	14	Stable plasmid inheritance protein OS=*Escherichia coli* O157:H7 GN=flmA PE=3 SV=1
P62671	6104	14	Stable plasmid inheritance protein OS=*Escherichia coli* O157:H7 GN=flmA PE=3 SV=1
P17776	1541	13	Putative PyrE leader peptide OS=*Escherichia coli* GN=pyrL PE=5 SV=1
P21361	5999	13	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
A1AAH4	7098	13	Uncharacterized protein YciY OS=*Escherichia coli* O1:K1 / APEC GN=yciY PE=3 SV=2
P21361	5999	13	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	12	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P75961	28846	12	Inner membrane protein YcfZ OS=*Escherichia coli* (strain K12) GN=ycfZ PE=1 SV=1
Q5JBK6	9381	12	UPF0401 protein YubL OS=*Escherichia coli* GN=yubL PE=3 SV=2
P21361	5999	12	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	12	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	12	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P21361	5999	11	Uncharacterized protein YciG OS=*Escherichia coli* (strain K12) GN=yciG PE=3 SV=1
P0ABY0	17210	11	Flagellar protein FliL OS=*Escherichia coli* O157:H7 GN=fliL PE=3 SV=1
Q8L0V4	79207	11	Chondroitin synthase OS=*Escherichia coli* GN=kfoC PE=1 SV=1
B7M9D0	58278	8	Glutamate--cysteine ligase OS=*Escherichia coli* O8 (strain IAI1) GN=gshA PE=3 SV=1
B7M9D0	58278	7	Glutamate--cysteine ligase OS=*Escherichia coli* O8 (strain IAI1) GN=gshA PE=3 SV=1

Appendix XI - Poster presented at XI Annual CICS-UBI symposium (30 June and 1 July, Covilhã, Portugal)

CICS-UBI
Centro de Investigação em Ciências da Saúde
Health Sciences Research Centre

FACULDADE CIÊNCIAS DA SAÚDE

IMPROVEMENT OF TWO-DIMENSIONAL GEL ELECTROPHORESIS DATA FOR PROTEOMIC PROFILING OF *ESCHERICHIA COLI* CELLS

Amaral A.G., Santos F.M., Gaspar L.M., Espírito Santo M., Queiroz J.A., Passarinha L.A.
CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Covilhã, Portugal;

INTRODUCTION

- The proteome is very dynamic and is largely affected by the growing conditions of the cell [1]. Indeed, the growth medium directly affects both the protein expression and DNA synthesis, so proteomics techniques may allow studying significant alterations that occurs in cell metabolism during the production of plasmid DNA (pDNA).
- *Escherichia coli* (*E. coli*) is an organism extensively characterized and thus several fermentation methods have been studied, optimized and implemented for the production of pDNA using this host. In particular, the study of the proteome may allow understanding the adaptations of *E. coli* metabolism to several growth media [2].
- Two-dimensional gel electrophoresis (2DE) is a method used for separation and fractionation of protein complex mixtures according isoelectric point and molecular weight [1] .

In this study, the main goal was the optimization of sample preparation and electrophoresis conditions in order to improve the fractionation and analysis of *E. coli* lysates by 2DE.

METHODS

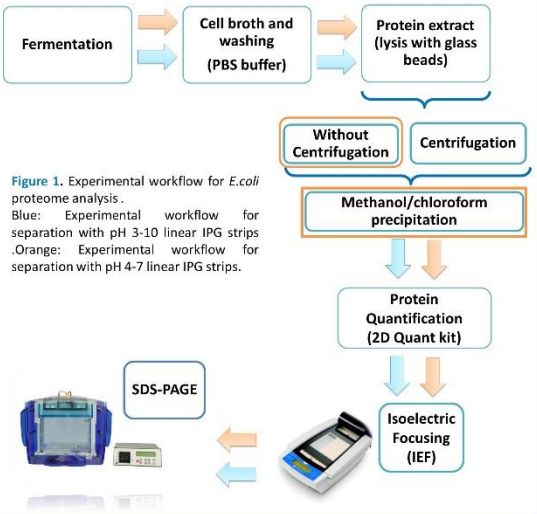
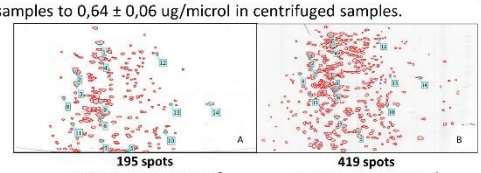


Figure 1. Experimental workflow for *E.coli* proteome analysis .
Blue: Experimental workflow for separation with pH 3-10 linear IPG strips .Orange: Experimental workflow for separation with pH 4-7 linear IPG strips.

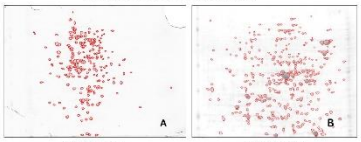
Results and Discussion

- Protein concentration: 5,29 ± 0.37ug/microL in non-centrifuged samples to 0,64 ± 0,06 ug/microl in centrifuged samples.



195 spots — **Total volume: 166.065 cm³**
419 spots — **Total volume:416.29cm³**

Figure 2. Centrifuged **(A)** and non-centrifuged **(B)** samples were analyzed by 2DE using pH 3-10 linear IPG strips, 24 cm.



195 spots — **Total volume: 111.04cm³**
305spots — **Total volume:180.30m³**

Figure 3: Samples analyzed by 2DE using **(A)** 3-10 pH range, 24 cm linear IPG strips and **(B)** 4-7 pH range, 18 cm linear IPG strips.

- Centrifugation of *E. coli* lysates leads to significant protein losses and this fact is evident when comparing the 2-DE gels from non and centrifuged samples.
- 3-10 pH resulting gels had low resolution and lower number of spots when compared to the 4-7 pH range.

CONCLUSIONS

- The experiments conducted using non-centrifuged lysates and a 4-7 pH linear gradient provided a great improvement in gel resolution.
- The application of this methodology will have several advantages for the establishment of *E. coli* proteomic profile during therapeutic pDNA biosynthesis.

REFERENCES

[1] T. Rabilloud and C. Lelong, "Two-dimensional gel electrophoresis in proteomics: A tutorial", J. Proteomics, vol. 74, no. 10, pp. 1829–1841, 2011.
[2] A. Berlec and B. Štrukelj, "Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells", J. Ind. Microbiol. Biotechnol., vol. 40, no. 3–4, pp. 257–274, 2013.

ACKNOWLEDGMENTS

This project was supported by University of Beira Interior – Health Sciences Research Centre (CICS). Santos FM acknowledges a doctoral fellowship (SFRH/BD/112526/2015) from FCT and Gaspar LM acknowledges to a fellowship from Novartis Farma-Produtos Farmacêuticos.

NOVARTIS FCT COMPETE 2020 PORTUGAL 2020

Appendix XII - Participation in the contest " Cultiva o teu Futuro – Inovação no Setor de Leite e Laticinios" sponsored by " Confederação dos Agriculturos de Portugal", 12 October, Lisbon, Portugal.

CAP AGRICULTORES DE PORTUGAL

CULTIVA O TEU FUTURO 5º CONCURSO UNIVERSITÁRIO

SINTESE DO PROJETO

Tema:	Valorização económica do lactosoro através do seu aproveitamento
Autores:	Ana Martinho e Andreia Amaral
Docentes:	Helena Alves e Luís Passarinha
Universidade(s)/Politécnico(s):	Universidade da Beira Interior (UBI)
Objetivos:	Com este projeto, pretende-se criar a AgroBioSynthesis, Lda. a qual pretende desenvolver meios de cultura para crescimento de microrganismos (bactérias) e/ou produção de vetores biológicos, otimizados para cada tipo de célula (estirpe bacteriana) e vetor biológico, a partir de subprodutos lácteos, em particular o lactosoro, para aplicação em I&D, rotina laboratorial, clínica e na indústria farmacêutica. Pretende-se ainda, numa fase posterior, prestar o serviço de otimização de produção de DNA plasmídico (pDNA) em células bacterianas, com vista a aplicação terapêutica, utilizando o lactosoro como matéria-prima principal da formulação dos meios de cultura.
Inovação do projeto:	Trata-se de uma abordagem pioneira de valorização económica do lactosoro, através do seu aproveitamento e utilização na formulação de meios de cultura para crescimento de células bacterianas para utilização em outros sectores. Paralelamente, pretende-se solucionar um problema ambiental atualmente enfrentado pelo sector do leite e laticínios através da diminuição do descarte de lactosoro para o meio ambiente e, por outro lado, possibilitar ainda o seu aproveitamento para utilização em outos sectores e indústrias, conferindo-lhe valor acrescentado. Além disso, numa segunda fase de implementação do projeto, pretende-se ainda prestar um serviço inovador (OptiDNAgro - otimização da produção de DNA plasmídico em determinada célula bacteriana) com base no lactosoro enquanto base para a formulação de meios de cultura otimizados a cada estirpe bacteriana e vetor biológico e, com isso, maximizar a sua cadeia de valor.
Principais conclusões:	Este projeto constitui a uma abordagem nova e inovadora, que servirá para conferir vantagem competitiva ao sector e acrescentar valor a alguns dos seus subprodutos, em particular ao lactosoro, um subproduto subvalorizado e considerado ambientalmente desfavorável.