UNIVERSITA' DEGLI STUDI
DI MILANO – BICOCCA

SCUOLA DI DOTTORATO DI SCIENZE
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Biotecnologie e Bioscienze

Corso di Dottorato di Ricerca in Biologia e Biotecnologie, XXX ciclo

Transposon based technology in DHFR knockout CHO cell line improves generation of AMH high producing clones for industrial applications

Dott.ssa Fiona Zucchetti

Anno accademico 2016/2017

Dottorato in Biologia e Biotecnologie, XXX ciclo

Dott.ssa Fiona Zucchetti
Matricola: 076262

Tutor: Dott.ssa Renata Tisi

Il lavoro presentato in questa tesi è stato realizzato presso i laboratori di Biologia Molecolare e Cellulare del DiaSorin Research Center, sotto la supervisione del dott. Andrea Dal Corso

Università degli Studi di Milano-Bicocca
Piazza dell'Ateneo Nuovo 1, 20126, Milano

Dipartimento di Biotecnologie e Bioscienze
Piazza della Scienza 2, 20126, Milano

Abstract

The claim of recombinant protein continues to increase for both research and industrial application. More and more resources are invested to develop the most efficient system for high protein expression. Mammalian cells are the best host for quality and reliability of expressed recombinant proteins. A diffuse protein production technology is based on dihydrofolate reductase (DHFR) deficient CHO cell line expression system. DHFR is an essential enzyme involved in purines synthesis and it is used as a selection marker. This technology has the potentiality to strength the selection process by inhibiting DHFR activity with methotrexate (MTX), forcing cells to amplify the DHFR gene and the gene of interest carried by the plasmid used for transfection. Thanks to the continuous progression in recombinant DNA technology, DHFR knockout cell lines are recently developed using different molecular tools. Preliminary data showed that by using the *piggyBac* (PB) transposon system it is possible to enhance the frequency of high producing clones compared to conventional transfections. In order to evaluate if PB transposon potentiality also occur in DHFR knockout cell line and to analyze the effects of amplification process on integrated transgene, research plan involved the comparison of the conventional gene transfer method to the PB transposon system. Our aim was to explore PB transposon in the DHFR knockout CHO expression system to obtain clones producing high yield of anti-Müllerian hormone (AMH) protein for industrial purposes. AMH is a dimeric glycoprotein, member of the TGFβ superfamily that causes regression of Müllerian ducts in male embryos. It has a fundamental role as diagnostic marker in assisted reproduction to predict ovarian reserve. The PB transposon system resulted a more efficient method to obtain high protein expression compared to standard vector. Moreover, it allowed higher performance in gene copy number integration and genetic stability than standard transfection method. In spite of the substantial improvement obtained by this new approach, AMH protein yield in scale-up was unsatisfactory. Therefore, we focused on construct engineering to increase productivity, since a higher MTX gene amplification was expected. Strategies to reduce the DHFR activity acting on its expression pathway were performed to enhance the AMH protein production. Even though encouraging data were obtained with the insertion of a less efficient IRES, achieved AMH protein yield was not comparable to what reported in literature exploiting the same approaches. Moreover, difficulties in culturing cells due to DHFR expression pathway impairments discouraged a further exploration. Thus, we

decided to shift our focus on epigenetic regulatory elements in order to force transcriptional efficiency. Impressive results were obtained in cell pools transfected with vectors carrying the UCOEs, in which high protein expression levels were reached, thus demonstrating the valid impact of these sequences. To further characterize the UCOE we decided to isolate single cell clones from the cell pool derived from the transposon with one UCOE sequence. We aimed to highlight differences between this new vector and our previous PB transposon. We reached the final goal to isolate a high-producing clone, which guaranteed expression and genetic stability over time.

In conclusion, in this work a new generated DHFR knockout cell line was exploited in association with the *piggyBac* transposon system to achieve high expression of recombinant protein. The PB transposon system was confirmed to be a powerful method to enhance the expression of AMH protein allowing higher performances in gene copy number integration and genetic stability compared to the standard transfection method. Employing epigenetic regulatory elements, such as the UCOE, not only a substantial increase in AMH protein yield was achieved, but also expression and genetic stability was strictly conserved over time.

Riassunto

Nell'ambito della ricerca e dell'applicazione industriale la richiesta di proteine ricombinanti è in continua crescita; vengono, quindi, investite sempre più risorse per sviluppare un sistema efficace in grado di ottenere un'elevata espressione proteica. Le cellule di mammifero sono l'ospite di elezione per lo sviluppo di proteine ricombinanti in quanto ne garantiscono la qualità. In particolare, alcune linee cellulari CHO deficitarie dell'enzima diidrofolato reduttasi (DHFR) sono molto diffuse come sistema di espressione. In tali linee cellulari, il DHFR, un enzima essenziale coinvolto nella sintesi delle purine, viene sfruttato come marker di selezione. È, inoltre, possibile aumentare la pressione selettiva grazie all'utilizzo del metotrexate (MTX), molecola che presenta un'attività inibitoria sul DHFR. L'inibizione spinge le cellule ad amplificare il gene DHFR come meccanismo di resistenza, questo porta ad una contestuale amplificazione del gene di interesse presente nella cassetta di espressione del plasmide. Grazie al continuo progresso della tecnologia del DNA ricombinante, sono state recentemente sviluppate, linee cellulari knockout per il gene del DHFR, con l'utilizzo di diversi strumenti molecolari. Dati preliminari, ottenuti nel nostro laboratorio, hanno dimostrato che utilizzando il sistema del trasposone *piggyBac* (PB) è possibile migliorare la frequenza di cloni alto-producenti rispetto al metodo di trasfezione convenzionale. Il progetto di ricerca ha avuto quindi come primo obiettivo quello di valutare se la potenzialità del trasposone si esprimesse anche nella linea cellulare DHFR knockout e di analizzare gli effetti del processo di amplificazione sul transgene integrato, comparandolo con i vettori convenzionali. Lo scopo finale era l'ottenimento di cloni alto-producenti dell'ormone anti-Mülleriano (AMH). L'AMH è una glicoproteina dimerica, membro della superfamiglia TGFβ. Essa causa la regressione dei dotti Mülleriani negli embrioni maschili ed è utilizzata come marker diagnostico nella riproduzione assistita essendo un indicatore della riserva ovarica. I dati ottenuti indicano che il sistema del trasposone risulta essere più efficiente rispetto al vettore standard nell'ottenere un'elevata espressione proteica, consentendo, inoltre, l'integrazione di un numero maggiore di copie del transgene e la stabilità genetica nel tempo. Nonostante il miglioramento ottenuto con questo nuovo approccio, la resa proteica nello scale-up non è stata soddisfacente. Pertanto, ci siamo concentrati sui costrutti per aumentare la produttività e potenziare l'amplificazione genica. Abbiamo quindi sperimentato strategie volte a ridurre l'attività del DHFR agendo sulla sua via di espressione. Anche se l'inserimento di un IRES meno efficiente

ha permesso di ottenere dati incoraggianti, la resa proteica raggiunta non è risultata paragonabile a quella riportata in letteratura. Inoltre, l'indebolimento della via di espressione del DHFR ha reso difficile la crescita delle colture cellulari e pertanto ha scoraggiato un'ulteriore esplorazione. Abbiamo quindi deciso di valutare gli elementi regolatori epigenetici per migliorare l'efficienza trascrizionale. Ottimi risultati sono stati ottenuti in pool di cellule trasfettate con i vettori contenenti gli elementi UCOE. Sono stati infatti raggiunti elevati livelli di espressione proteica dimostrando il sostanziale effetto di queste sequenze. Per caratterizzare ulteriormente gli elementi UCOE, abbiamo deciso di isolare alcuni cloni. Lo scopo era evidenziare le differenze tra questi nuovi vettori ed il trasposone precedentemente utilizzato. L'obiettivo finale è stato raggiunto. Abbiamo, infatti, isolato un clone ad alta produttività caratterizzato sia dalla stabilità genetica sia da quella dell'espressione nel tempo.

In conclusione, l'utilizzo di una nuova linea cellulare knockout per il gene DHFR in associazione con il sistema del trasposone *piggyBac* ha permesso di ottenere una buona espressione della proteina ricombinante AMH. Il trasposone si è confermato un sistema efficiente per incrementare l'espressione proteica; inoltre, l'utilizzo dell'elemento regolatore epigenetico UCOE non solo ha permesso un ulteriore e significativo aumento della resa proteica, ma anche un rigoroso mantenimento della stabilità genetica e dell'espressione nel tempo.

TABLE OF CONTENTS

Introduction

CHAPTER 1: Recombinant protein production

1.1 Recombinant proteins: a brief overview

The history of recombinant proteins began in the early 1970s when Cohen and collaborators discovered recombinant DNA exploiting the bacterium *Escherichia coli*. They synthesized a plasmid vector *in vitro* which was inserted in the bacterium by transformation. Thus, they obtained a powerful tool for introducing foreign DNA into a host cell (Cohen et al., 1973). This work laid the foundation for the development of recombinant protein expression. Inserting exogenous transgenes in hosts and forcing them to express proteins that they naturally could not was successful. Indeed, after only a decade, a recombinant protein entered the market as the first pharmaceutical product for diseases treatment. This recombinant protein, created by Genetech and developed by Eli Lilly, was human insulin (Humulin®) from *Escherichia coli* and it was used to replace the natural protein in patients suffering diabetes because of their insulin-deficiency. It was approved by FDA in 1982 for clinical use (Goeddel et al., 1979). Soon, it became clear that other cells, including mammalian cells, can be exploited as hosts for recombinant proteins and, in some cases, with greater advantages. In this scenario, Genetech in 1986, only a few years after the creation of the recombinant human insulin, received approval for production and marketing of the recombinant human tissue plasminogen activator (tPA, Activase®) from Chinese Hamster Ovary (CHO) cells. Activase® was the first therapeutic protein deriving from recombinant mammalian cells (Wurm, 2004). In a brief period of time, many other recombinant proteins were produced as pharmaceutical products in order to be synthetic substitutes in deficient patients, such as growth hormone, or to take place in particular pathways, such as interferon-α, tissue plasminogen activator and erythropoietin, or to be used as therapeutic antibodies such as infliximab and ritxumiab (Carter, 2011). In 2009, 151 recombinant proteins were approved by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA) as pharmaceuticals. The majority of them, 39%, was obtained in mammalian cells, followed by 29.8% produced in *Escherichia coli*, 18.5% in *Saccharomyces cerevisiae*, 11.2% in hybridoma cells, 0.7% in transgenic goat milk and 0.7% in insect cells (Ferrer-Miralles et al., 2009). Nowadays, the recombinant protein production is a multi-billion dollar business including different branches: therapeutics, diagnostics, industrial processes and research. Just to have an idea, in only therapeutics, antibody and

protein global sales reached $154 billion in 2015 with a growth of 9.2% compared with the previous year (Martin, n.d.) and it is estimated to grow of around 8.6% over the next decade to reach approximately $315.90 billion by 2025 ("$315.9 Million Protein Therapeutics Market Analysis and Trends Report 2016-2025," 2017).

1.2 Recombinant protein expression systems

Recombinant proteins can be obtained using several hosts such as bacteria, filamentous fungi, yeast, insect, plant tissue, transgenic animal and plant, mammalian cells and recently cell-free method are also developed (Ghaderi et al., 2012; Rosenblum and Cooperman, 2014; Schmidt, 2004). Therefore, approaching the recombinant protein production, the first crossroads is the choice of expression system and the decision between prokaryotic and eukaryotic expression systems is a critical juncture. Prokaryotic expression system allows to perform a rapid and simple protein production due to the high cell growth rate and the easiness of process. Production can also be induced using the lactose analogous, Isopropil-β-D-1-tiogalattopiranoside (IPTG). Moreover, the purification step is not complicated, except when proteins become insoluble and consequently inclusion bodies formation occurs. This event is one of the limitations of prokaryotic system because it affects the recovery of functional protein and the purification process. However, the greatest disadvantage using bacteria for recombinant protein production is the lack of most post-translational modifications. Indeed, proteins without correct post-translational modifications might result non functional or might show activity or stability problems (Dumont et al., 2016; Graumann and Premstaller, 2006).

Eukaryotic expression system includes a wide scenario of opportunities from yeast, plant, insect cells to mammalian cells. Yeast cells are a cheap, rapid and easily scalable method to obtain recombinant proteins. They can grow at a very high density with fermentations. The most exploited species in industry are *Saccharomyces cerevisiae* and *Pichia pastoris*. Yeast cells overcome the limitations associated with bacteria system such as protein folding and stereochemistry because they are able to perform complex post-translational modifications. However, alterations can occur in glycosylation patterns that might cause protein activity problems. The main critical feature of yeast cells as host for recombinant protein production is the high mannose insertions during post-translational modification process. This might reduce protein half-life and alter protein activity and immunogenic

properties due to changes in glycosylation pattern (Dumont et al., 2016; Schmidt, 2004). Plant and insects cells are able to produce complex multimeric proteins and glycoproteins. However, exogenous proteins might be produced with a completely different stereoisomeric structure and with a very different glycosylation pattern and glycan structures compare to the native one. Indeed, proteins produced by plant cells lack of many glycosylated residues present in humans, such as sialic acids, and show the presence of α1,3-fructose and β1,2-xylose epitopes, absent in humans. Baculovirus-insect cell expression system is the workhorse of insect cells protein production system. Insect cells infected with the viral vector baculovirus are exploited to produce mainly virus-like particles and vaccines. Proteins produced in insect cells do not develop galactose or sialic acid residues because of the presence of high or pauci-mannose residues. These mannosilation patterns occur because insect cells produce truncated forms of N-glycan precursors (Ghaderi et al., 2012; Kost et al., 2005).

Mammalian expression systems have gained the most significant market share of recombinant protein production due to their ability to carry out proper protein folding, assembly and post-translational modifications. What makes mammalian cells a superior expression system than the others is the capacity to produce proteins harboring human-like complex glycosylations such as N-linked and O-linked glycosylations. The drawbacks of mammalian expression system are the long time required for protein production, the low yield obtained and the high costs of the entire process (Swiech et al., 2012).

A plethora of protein expression platforms, which are able to produce recombinant proteins with high quality at industrial level, are currently available. Many considerations should be taken into account to choose which one fits for the purpose. Prokaryotic or lower eukaryotic systems have the advantageous of being more rapid and cheaper, but mammalian cells guarantee to obtain proteins that closely resemble the native forms. This is essential if functionally active proteins for biological studies or as biopharmaceutical products are required.

1.2.1 An industrial point of view on our work

Recombinant proteins are exploited for wide applications in different fields including research, biotechnology and medicine. Medicine represents the main and the best-known application area due to the therapeutic proteins, which provide important treatments for a variety of diseases. Besides recombinant proteins with the function of therapeutics such as

enzymatic or regulatory activity, or delivery function or vaccines, a branch in medicine deals with recombinant proteins for diagnostics. DiaSorin is a diagnostic company that produces immunodiagnostic kits developed on a proprietary platform, the Liaison. Liaison platform is an automated instrument that performs chemiluminescent immunometric analysis of biological matrices (serum, plasma and feces). The kits are based on immune reaction, thus on antigen-antibody recognition. Recombinant proteins and antibodies are fundamental components of the kits being applied as control and calibrators instead of patient serum/plasma. In the early phase of kit development, recombinant proteins can be also used as immunogens for antibody production in animals. In our company, different expression systems are exploited for recombinant protein production. Since recombinant proteins are used for diagnostic, they are not controlled by regulatory affairs. Therefore, considering that specific safety is not required, expression systems such as *E. coli* can be also exploited without any concerns. Moreover, despite a proper folding is strictly necessary, biological activity is not crucial simplifying the condition of the entire expression process. In this scenario, mammalian cells play a pivotal role when protein with a proper folding or difficult to express are required as, for example, proteins with complex post-translational modifications or disulfide bonds. Hence, it is fundamental for us to improve the recombinant protein production in order to satisfy manufacturing demand and to get ready to manage challenging projects on new diagnostic targets.

1.3 An outline on mammalian cell lines

Recombinant proteins and antibodies are mainly produced in mammalian cell lines with a percentage of over 50% among currently approved biologics. It goes without saying that the major focus is on mammalian cells as hosts for recombinant proteins and monoclonal antibodies production (Kantardjieff and Zhou, 2013). Several mammalian cell lines are exploited for this purpose including Chinese Hamster Ovary (CHO), Baby Hamster Kidney (BHK), mouse myeloma cells (NSO and SP2/0), fibroblast-like cell lines derived from monkey kidney tissue (COS), African Green Monkey Kidney cell line (Vero), human cell lines such as Human Embryonic Kidney (HEK), HeLa and human retina derived cells (PER.C6) and hybridoma. Each mammalian expression system has advantages and limitations, thus the selection of the expression host depends on different factors in order to meet specific requirement. Many criteria should be taken into account to perform the best choice: first of

all, the protein to be expressed and its use. Based on protein characteristics, the selection of the expression system can be carried out considering cell growth, extracellular expression capability, post-translational modifications, protein folding and biological activity, regulatory and economic issue, especially if a large-scale production is required, and product safety, in particular for therapeutic proteins (Jayapal et al., 2007; Khan, 2013; Picanco-Castro et al., 2013; Rosenblum and Cooperman, 2014; Swiech et al., 2012; Wurm and Hacker, 2011).

CHO

Chinese Hamster Ovary (CHO) cells are the most frequently used host for the recombinant protein production. Since their first use, CHO cells caught the attention for their potentiality. Nowadays, they command the market with a production of 61% among currently approved biologics derived from mammalian cells (Figure I) (Kantardjieff and Zhou, 2013; Kunert and Reinhart, 2016). Due to its great importance in the scenario of mammalian expression system, the second chapter is fully dedicated to describe CHO cell line.

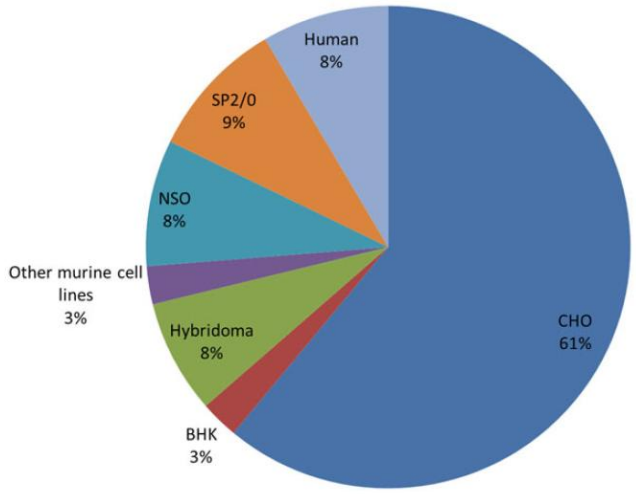


Figure I. Mammalian cell pharmaceutical production: Schematic representation of percentage of protein-based recombinant pharmaceutical produced by mammalian cells (Kantardjieff and Zhou, 2013).

BHK

Baby Hamster Kidney cell line was derived from polyoma transformation of hamster cells. This cell line is exploited for biopharmaceuticals, in particular, for vaccine production. It is suitable for transient and stable expression. This cell line is particularly susceptible to several viral infections. Two important and challenging biotherapeutics, coagulation factor VIIa and VIII, were produced in BHK cell line and approved in United State and Europe. The

strain BHK21, derived from *Mesocricetus auratus* or Syrian golden hamster, is the most widely used ("BHK-21 [C-13] ATCC ® CCL-10™ Mesocricetus auratus kidney nor," n.d.; Dumont et al., 2016).

NS0

Murine myeloma NS0 cells were isolated from the plasmacytoma tumor MOPC21 induced in a Balb/c mouse by peritoneal injection of a mineral oil. NS0 cell line was initially used as a fusion partner for generation of hybridoma cells, but today is mainly used for recombinant protein expression, in particular antibody production. This cell line is auxotrophic for glutamine because contains very low levels of endogenous Glutamine Synthetase (GS). The deficiency in a metabolic enzyme can be exploited as a selection marker. NS0 cells require an exogenous copy of GS to survive in a glutamine-free medium. Therefore, cells positively transfected with a plasmid carrying the GS gene and the gene of a heterologous protein can be selected in a glutamine-free medium. This system allows to obtain cells that stably express GS and the protein of interest. Furthermore, it is possible to strengthen the selective pressure with methionine sulfoximine (MSX), an inhibitor of GS enzyme (Cacciatore et al., 2010; Galfrè and Milstein, 1981).

SP2/0

Sp2/0 cells were derived from the fusion of a Balb/c mouse spleen with mouse myeloma cells. This cell line is mainly used as fusion partner for generation of hybridoma cells ("Sp2/0-Ag14 ATCC ® CRL-1581™ Mus musculus (B cell); Mus muscu," n.d., p. 0).

COS

COS stands for CV-1 Origin, SV40. This name highlights that COS cells were derived from CV-1 cell line, in turn, originated from the kidney of the African green monkey and the use of SV40 virus to perform immortalization. The SV40 virus used for immortalization produces a large tumor antigen (T antigen) that is defective in genomic replication. COS cells highly express SV40 large T antigen which is required to trigger viral DNA replication at the SV40 origin. In this way, it is sufficient to transfect COS cells with vectors carrying the SV40 region of replication and the gene of interest to obtain prolonged transient protein production. Two forms of COS cell lines commonly used are COS-1 and COS-7 (Aruffo, 2001; Gluzman, 1981).

Vero

Vero cell line was derived from the kidney of an African green monkey by Y. Yasumura and Y. Kawakita at the Chiba University in Japan in 1962. Vero cells are susceptible to several viruses. They are used as a vaccine cell substrate and they are also widely applied for virus replication studies and plaque assays. The Inactivated Poliovirus Vaccine (IPV), a very fundamental vaccine widely used in the U.S., was produced in Vero cell line (Sheets, 2000).

HEK

Human Embryonic Kidney cell line, the well-known HEK293, was generated after exposure to sheared fragments of human adenovirus type 5 (Ad5) DNA by Graham and collaborators in 1977. The source of the cells was a healthy aborted foetus. The number in the name HEK293 was referred to the 293rd experiment of Graham (Graham et al., 1977). HEK293 cells are widely diffuse because they are easy to handle and they grow in suspension in serum free medium. These characteristics make them suitable for large-scale production. Despite their epithelial origin, HEK293 cells are able to produce proteins with their proper post-translational modifications and folding. This feature distinguishes HEK293 cells from the most widely used CHO cells which are unable to carry out some types of human glycosylations and post-translational modifications (Durocher and Butler, 2009; Grabenhorst et al., 1999). HEK293 cell line is widely used for both transient and stable expression especially for neurobiology studies (Thomas and Smart, 2005). Engineered cell lines were also developed to improve recombinant protein expression. Therefore, highly transfectable derivatives of HEK293 cells that contain SV40 large T-antigen (HEK293-T) or Epstein-Barr Virus Nuclear Antigen (HEK293-EBNA) were established. In the same way as COS cell line is used, these cells lines are competent to replicate vectors carrying the SV40 or EBV region of replications, thus transgene expression is prolonged after transient transfection (DuBridge et al., 1987; Meissner et al., 2001).

HeLa

HeLa cells were derived from a cancer sample taken from the cervix of a 30-years-old young black woman called Henrietta Lacks. HeLa was the first human immortalized cell line and was established in 1951. Currently, HeLa cells are used for production of viral vaccines such as other human diploid cells (Masters, 2002; Petricciani and Sheets, 2008).

PER.C6

PER.C6 cell line was derived from primary foetal retinoblasts that were immortalized upon transfection with E1 region of adenovirus type 5. This cell line is exploited for vaccine production due to its easiness to propagate many viruses (Fallaux et al., 1998; Pau et al., 2001; Sanders et al., 2013).

Hybridoma

Hybridoma is “an immortal antibody-secreting cell line derived by fusion of a short-lived lymphocyte and a myeloma cell line” as defined by its inventor César Milstein. This technology is exploited to produce high level of monoclonal antibody. The name “hybridoma” was coined by Leonard Herzenberg around 1976-1977 during a conversation among colleagues. This term became popular within the group, even if one of them defined the word “hybridoma” garbled Greek (Milstein, 1999). Monoclonal antibodies production has gained a position of substantial importance in economics. In 2011 a top nine classification of biologic drug sold in US market was published and monoclonal antibodies occupied the first place with US sales of $24.6 billion followed by hormones ((Rob) Aggarwal, 2014).

CHAPTER 2: Chinese Hamster Ovary cells: the flagship of mammalian expression system

2.1 The potentiality of CHO cells

The enhancing interest in CHO cell line as expression system comes from four main advantages.

Firstly, CHO cell culture and safety properties guarantee an easy handling and out of danger product due to serum-free and suspension condition. The serum-free culture condition results in a better reproducibility across different batches and in a higher safety if compared to media containing human or animal derived proteins. The suspension culture allows an easy scale-up for bioreactor production (Kim et al., 2012). In addition, the robustness of CHO cells allows them to withstand pH variations, changing in oxygen level and pressure alterations (Ghaderi et al., 2012). Besides, CHO cell line is considered a safe host. Analysis performed since CHO cells entered the biopharmaceutical market demonstrated its safety which will easily guarantee future FDA approval for new products derived from them (Jayapal et al., 2007). A couple of reasons make CHO cells resistant to viral infection and, thus, a safe host: one is the lack of several key viral entry receptors and the second is that they do not encode for some gene necessary for viral protein synthesis, as happens for cowpox host range factor (CP77) which confers to CHO cells Vaccinia virus resistance (Xu et al., 2011).

Secondly, CHO cells machinery can perform post-translational modifications, including glycosylations with human-like N-glycans. This ability is fundamental for biopharmaceutical industry because it allows to preserve stability and activity of therapeutic glycoproteins. However, CHO cells, lacking of α[2-6] sialyltransferase and α[1-3/4] fucosyltransferases, do not produce some types of human glycosylations. In addition, they generate non-human modifications such as sialic acid N-glycolylneuraminic acid (Neu5Gc) and galactose-α-1,3-galactose (α-Gal) epitope, which can cause immune response (Ghaderi et al., 2012).

Thirdly, the extraordinary plasticity and variability of CHO genome in combination with its feature of being functionally hemizygous for many genes allowed the isolation of mutants with deficiencies in metabolic enzymes. Since mutants required certain specific nutrients to survive due to the lack of the specific enzymes, the deficiency became an ideal selection marker, consenting the generation of producer lines. Mutant cell lines deficient for adenine

phosphoribosyl transferase (APRT) and dihydrofolate reductase (DHFR) are some examples (Jayapal et al., 2007).

Fourthly, since CHO cell line has been increasingly employed in the last thirty years, remarkable advances were implemented which have guaranteed substantial increase in recombinant protein production. In particular, gene amplification strategy exploited in CHO cell line with GS and DHFR systems, deeply explained below, allowed to enhance specific productivity. Moreover, improvements in vector engineering, selection strategies and integration in transcriptional hotspots contributed to reach high expression level of recombinant proteins (Kim et al., 2012).

Recombinant protein quality, stability, activity and safety can be obtained using CHO cells. The drawback is the high cost of the final product considering the entire process from raw material (such as media, transfection reagents, etc.) to the downstream procedures.

2.2 The history of CHO cell line

Cricetulus griseus, well-known as Chinese hamster, is a rodent that lives in the steppes, deserts and fields of Mongolia and China. This species has been used as a laboratory animal for typing pneumococci since the first decade of 1900. Starting from the second decade, it was widely studied for epidemiological research because it was vehicle of the parasite *Leishmania*. In 1948, Chinese hamster landed illegally in the Unites States carried by Dr. C.H. Hu and Dr. R. Watson. This hasty action caused a diplomatic incident between China and Unites States and the two scientists were accused for “war crimes”. Indeed, in the thorny background of the cold war, there was the worry that the Unites States would use Chinese hamster as an agent of biological warfare once infected with microorganisms causing deadly diseases (Jayapal et al., 2007). In the middle of 20th century, in the Dr. George Yerganian's laboratory at the Boston Cancer Research Foundation, the attempts of breeding Chinese hamster caused the development of spontaneous hereditary diseases. This event laid the foundation for genetic research. Dr. G. Yerganian’s laboratory provided the outbred hamster from which the well-known Chinese Hamster Ovary (CHO) cell line was established in 1957 by Dr. T. Puck and collaborators at the University of Colorado. In particular, starting from the ovary tissue, a culture of fibroblast cells with a near diploid karyotype of 22 chromosomes was isolated. This culture underwent a spontaneous immortalization. After a while, cells in culture morphologically changed from a fibroblast to

an epithelioid culture. Probably, after a cloning step of this morphological different culture, the original CHO cell line was finally obtained (Puck et al., 1958; Wurm and Hacker, 2011). This original CHO cell line started to be supplied to many laboratories. These cells were resilient, fast and easy to grow in adherent culture with high cloning efficiency; moreover, karyotype heterogeneity was observed among populations. For these reasons, they soon became a model for cytogenetic studies. It is interesting to highlight that all CHO cell lines on the market require proline in culture media. This means that the deficit in proline synthesis arose in the very first step of CHO cell line history.

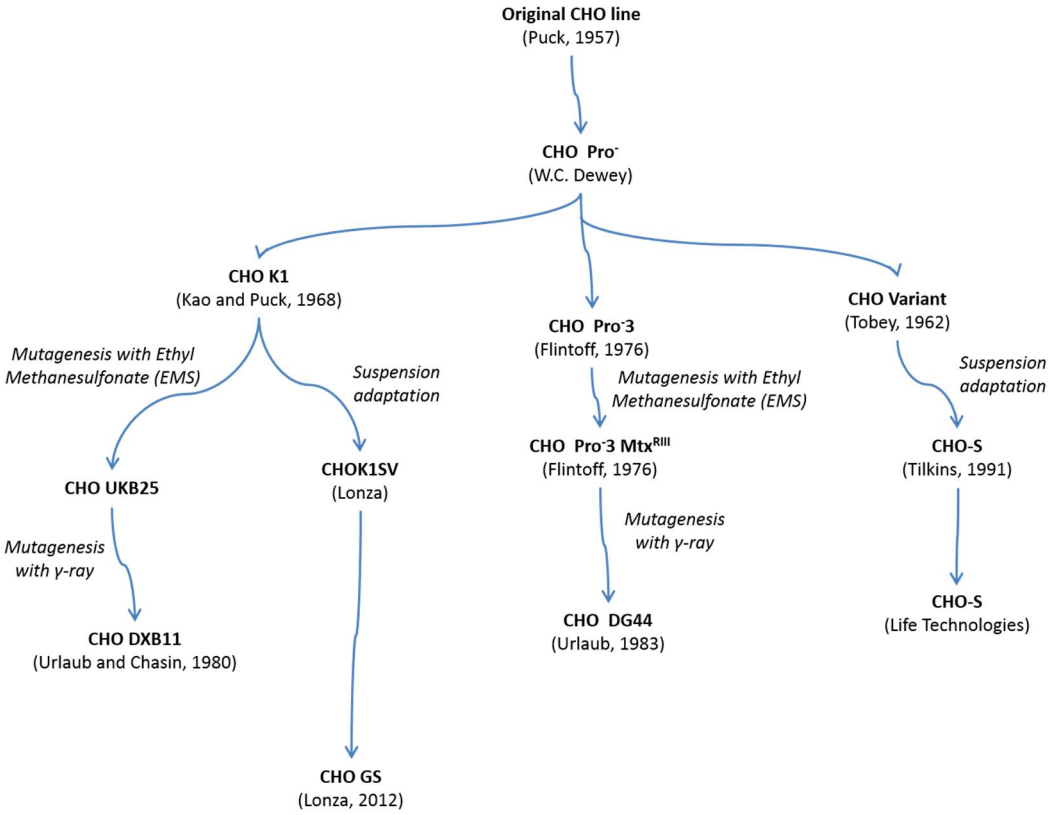


Figure II. The family tree of CHO cell lines: Schematic representation of CHO cell lines origin. The figure is based on multiple sources (Flintoff et al., 1976b, 1976a; Kaas et al., 2015; Kao and Puck, 1968; Kilbey et al., 2012; Lewis et al., 2013; Liu et al., 2010; Puck et al., 1958; Stanley and Siminovitch, 1976; Urlaub et al., 1983; Urlaub and Chasin, 1980; Wurm and Hacker, 2011).

According with literature three main lineages were derived from the original CHO cell line (Figure II):

1) CHO K1, from whom were derived CHO DXB11 and the recent CHO GS cell lines
2) CHO Variant that engendered the CHO-S cells
3) CHO Pro3-, the ancestor of CHO DG44 cell line

(Flintoff et al., 1976b, 1976a; Kaas et al., 2015; Kao and Puck, 1968; Kilbey et al., 2012; Lewis et al., 2013; Liu et al., 2010; Puck et al., 1958; Stanley and Siminovitch, 1976; Urlaub et al., 1983; Urlaub and Chasin, 1980; Wurm and Hacker, 2011).

2.2.1 CHO-K1 lineage

CHO-K1 subclone was isolated from the parental CHO line, CHO Pro^-. It was established in the late 1960s and maintained in Denver laboratory by Dr. Puck and Dr. Kao. CHO K1 cell line spread all over the world and a vial was also deposited at America Type Culture Collection (ATCC). Shortly afterwards, literature was enriched by research exploiting CHO-K1 cell line as host of recombinant protein production. The well-known CHO-K1 lineage is an adherent, epithelial-like cell line suitable as a transfection host. After CHO-K1 subclones spreading, several experiments aimed to isolate auxotrophic mutants were performed and different cell lines were established. The successful idea was to exploit these mutants as selection system to generate stable clones expressing the gene of interest. Based on this conception, in different period of time CHO-DXB11 and CHO-GS arose from CHO-K1.

CHO DXB11

A great goal was reached by Dr. Urlaub and Dr. Chasin in 1980 when they isolated a cell line called CHO DXB11. Starting from CHO-K1, UKB25 cell line was first generated after a round of random chemical mutagenesis using Ethyl MethaneSulfonate (EMS). UKB25 cell line lacks of a copy of dihydrofolate reductase (DHFR) gene. DHFR is an enzyme that catalyzes the conversation of dihydrofolic acid to tetrahydrofolic acid, an essential cofactor carrier involved in purine synthesis. Subsequently, the UKB25 (dhfr+/dhfr-) cell line was exposed to a second round of mutagenesis with γ-radiation and DXB11 (dhfr-/dhfr-) cell line was isolated. The strategy was to transfect DXB11 (dhfr-/dhfr-) cells with a plasmid carrying a functional copy of DHFR in an expression cassette with the gene of interest. In this way, transfected cells could compensate for DHFR deficiency surviving in a selective medium. Moreover, it was possible to amplify gene copy number strengthening the selection with

methotrexate (MTX), a DHFR inhibitor. DXB11 cell line marked the history of recombinant protein production in mammalian cell lines being the first CHO host cell for large-scale production of a protein product (human tissue plasminogen activator). Currently, it is still exploited for numerous protein products on the market (Kaas et al., 2015; Wurm, 2013).

CHO GS

Recently, a new selection system was developed based on the strategy exploited by NS0 cells. CHO-K1 cells were adapted to grow in protein-free suspension and, in 2002, a subline called CHOK1SV was isolated by the Lonza Biologics company and deposited at European Collection of Animal Cell Culture (ECACC). This cell line exploited the glutamine synthetase (GS) gene expression system. Starting from these cells, several CHO GS cell lines were generated. The latest GS Gene expression system was launched in the market in 2012. CHO GS cells are deficient for glutamine synthetase gene. Therefore, cells positively transfected with a plasmid encoding for GS gene linked with the gene of interest can survive in glutamine-free culture condition. This strategy allows to perform a strong selection of cells expressing the protein of interest. In addition, it is possible to strengthen the selection process adding the GS enzyme inhibitor, methionine sulphoximine (MSX). However, the amplification step is usually not necessary because a few copies of the transgene per cell are enough to obtain high protein producing clones. The potentiality of GS system made it the gold standard for mammalian gene expression already counting for 35 licensed pharmaceutical products (Lewis et al., 2013; Wurm, 2013).

2.2.2 CHO Pro$^-$3 lineage

In 1976, Dr. Flintoff and collaborators isolated and characterized three methotrexate-resistant phenotypes from CHO cells. CHO cell line that they used was obtained from Dr. W. C. Dewey (Pro$^-$). Pro$^-$3 and Pro$^-$4 were CHO Pro$^-$ cell line subclones that were developed in their laboratory. Pro$^-$3 MTXRI was MTX-resistant cell line selected from Pro$^-$3 with Class I resistant properties. Pro$^-$3 MTXRII and Pro$^-$3 MTXRIII were MTX-resistant cells of Class II and Class III, respectively. In 1983, Dr. Urlaub and collaborators received Pro$^-$3 MTXRIII cell line from Dr. Flintoff's laboratory and exposed them to y-irradiation mutagenesis aiming to obtain a cell line with full deletion of both DHFR alleles. They succeeded and isolated CHO-DG44 cell line (Flintoff et al., 1976b, 1976a; Urlaub et al., 1983).

CHO DG44

The idea of Dr. Urlaub and collaborators started from the observation of a reversion rate to DHFR activity in CHO-DXB11. In order to overcome this problem, they decided to proceed with a full depletion of DHFR loci. They decided not to operate on CHO-DXB11, but to use $Pro^{-}3$ MTXRIII that proved to be suitable for deletion of both DHFR alleles (Flintoff et al., 1976a; Urlaub et al., 1983). Nowadays, the literature is enriched with hundreds of articles describing experiments on DG44 CHO cells.

2.2.3 CHO Variant lineage

Dr. R.A.Tobey at Los Alamos National Laboratory received another subline of the parental CHO Pro^{-} line and in 1962 generated CHO Variant cells. Starting from this cell line, the well-known commercial line CHO-S (Life Technology), which is widely used for recombinant protein production, arose.

CHO-S

It is not very clear the real origin of CHO-S cell line. In his review, Dr. Wurm described two CHO-S cell lines that, in all probability, differ each other in term of culture conditions and phenotypic and genotypic features. One cell line, called CHO-So (o=original) by Dr. Wurm, was isolated by Dr. Thompson and collaborators starting from CHO-Toronto (a sister cell line of CHO-K1). In the Dr. Wurm's description, CHO-MTXRIII cells were CHO-So derivative cell line and, consequently, the CHO DG44 cells. The other cell line, called CHO-Sc (c=commercial), passed on through Dr. Tobey's laboratory to the company Life Technology that has marketed it since 2002 with the name of CHO-S (Wurm, 2014).

2.3 The "Quasispecies" concept

CHO cell line is a huge family including various cell lines that widely differ genetically. Its early history and its characteristic of genetic instability explain the variability and diversity of genomic structures observed in the members of CHO cell lines. The concept of "Quasispecies" was introduced by Dr. Wurm to identify this constant genomic remodeling within a population (Wurm, 2013). The genomic heterogeneity of CHO cell lines was widely demonstrated in 40 years of cytogenetic studies. A comparison between karyotype of Chinese hamster and the ancestor of CHO-K1 showed that Chinese hamster has 22

chromosomes (10 autosomal pairs and 2 sex chromosomes) in contrast to CHO-K1 cell line that has a peculiar set of 21 chromosomes: 8 corresponding with those of Chinese hamster and 13 named 'Z-chromosomes' due to their rearranged structures (Figure III) (Wurm and Hacker, 2011).

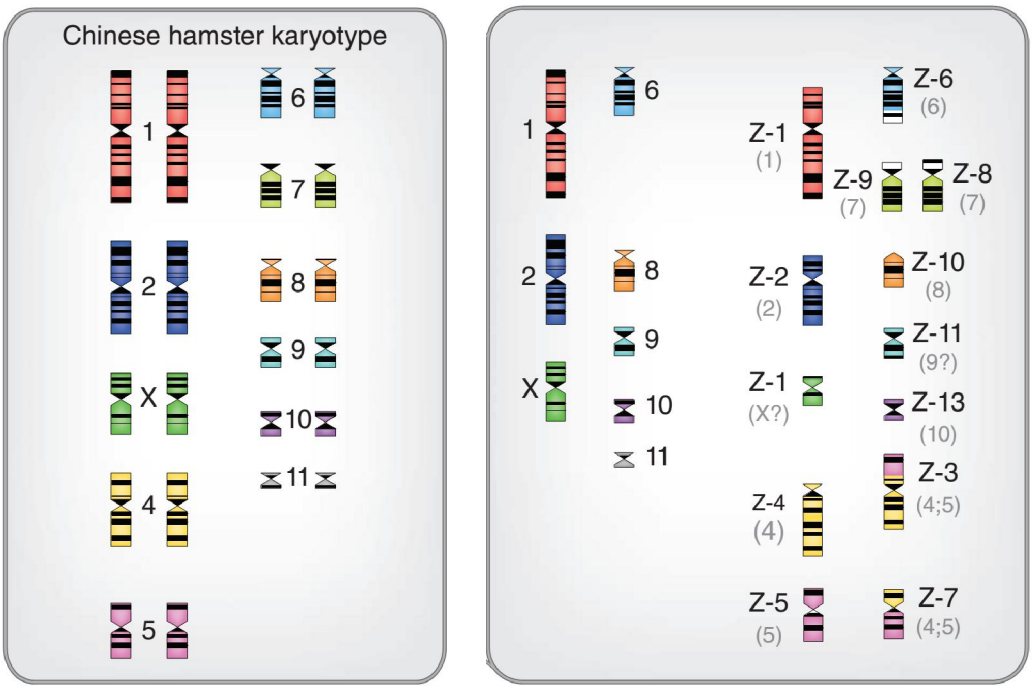


Figure III. Schematic representation of Chinese hamster and CHO-K1 cell line karyotypes: Comparison between karyotypes of Chinese hamster (on the left) and of CHO-K1 cell line (on the right) (Wurm and Hacker, 2011).

Dr. Wurm and Dr. Hacker's work is one of the different studies that describe 'macroscopically' chromosomal rearrangements in CHO cell line using different techniques. In an elegant work presented by Dr. Lewis and collaborators, genomes of seven CHO cell lines were compared with a "whole-genome sequence-level" approach highlighting the heterogeneity among cell lines. The authors analyzed seven CHO cell lines from the CHO-K1, DG44 and CHO-S lineages compared to genome sequence of a female Chinese hamster (*Cricetulus griseus)*. Results showed that mutations including SNPs, indels, CNVs and missing genes were mixed in a distinct set for each cell line. Moreover, it was demonstrated that accumulations of mutations occurred rapidly during the development of production cell lines. Phenotypic modifications seem to be displayed in several ways due to the different set of mutations; anyway, no alterations in proliferation and immortalized processes were present. On the contrary, genes fundamental for bioprocess, such as

apoptosis, were affected by mutations (duplications) (Lewis et al., 2013). The advantage of having available a detailed map of CHO cells genomic landscape is to increase helpful knowledge for improving cell line selection, characterization and engineering, as well as bioprocess and culture condition optimization. This, in turn, means to maximize CHO cells potentiality and increase CHO-based bioprocessing and manufacturing.

2.4 CHO DHFR knockout cell line: from a fortuitous mutation to a engineered state of the art

Research activity was the initial aim of the isolation of CHO mutants with deficiencies in metabolic enzymes. However, it did not take much to understand the great potentiality of these mutants for recombinant protein production. Indeed, the nutritional requirement of mutant cells was exploited as a system to select those transfected with exogenous genes. The first dihydrofolate reductase (DHFR) deficient cell line, named CHO DXB11, was generated with random mutagenesis by Dr. Urlaub and Dr. Chasin in 1980. Therefore, a fortuitous mutation allowed to isolate DHFR deficient CHO cell line that briefly spread as an efficient recombinant protein expression system. DHFR is a small monomeric enzyme that catalyzes the conversion of folic acid to tetrahydrofolate (THF). THF is a cofactor carrier for one-carbon moieties required in various biosynthetic reaction. It is essential for *de novo* synthesis of purines (Figure IV). Purines production is so fundamental for cells that they are equipped with a rescue pathway, the salvage synthesis. Cells can perform this pathway in the presence of purine precursors, hypoxanthine and thymidine (HT) (Anderson et al., 2011; Omasa, 2002). For this reason, cells lacking functional copies of the DHFR gene must be grown in culture medium supplemented with HT to survive. This metabolic selection process can be used to select for cells that have been transfected with a gene of interest for recombinant protein expression along with an exogenous copy of the DHFR gene. Moreover, the addition of methotrexate (MTX), a folic acid analog that blocks DHFR activity, in culture medium strengthens selection process by inhibiting DHFR activity and forces cells to amplify DHFR gene and the gene of interest. Consequently, this technology allows to generate stable and high producing clones (Cacciatore et al., 2010; Kingston et al., 2002). DHFR deficient cell lines currently used are DXB11 and DG44. DXB11 cell line carries a deletion of one DHFR locus and a missense mutation (T137R) of the second DFHR locus making cells incapable of reducing folate.

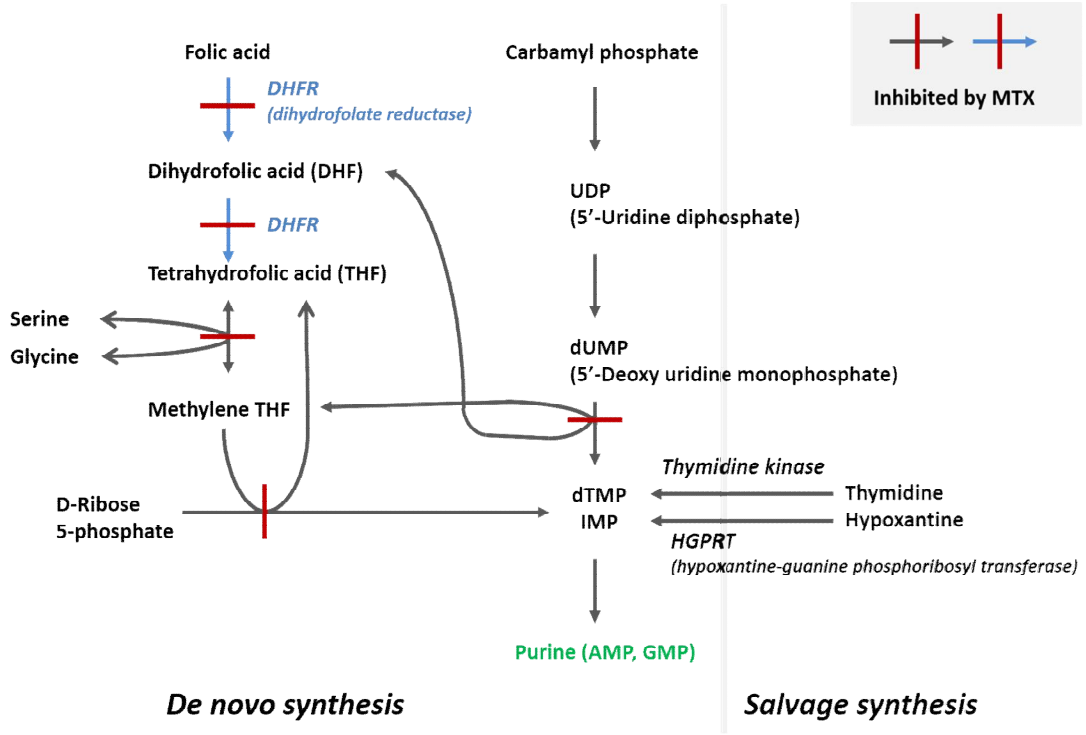


Figure IV. DHFR pathway: *de novo* and salvage biosynthesis pathways for purines (Omasa, 2002).

DG44 cells, instead, have a full deletion of the two DHFR loci, one positioned on chromosome 2 and the other on a shortened marker chromosome variant Z2. DXB11 cells result in a less effective selection strategy due to the rate of reversion to DHFR activity. On the contrary, CHO DG44 cells allow strong metabolic selection of recombinant cell lines (Wurm, 2013). However, these cells are unusable for production at industrial level because they are known to display several problems during cell culture, such as difficulty of growth in serum free medium and clumping aggregation in shaken cultures (Ho et al., 2013). Nowadays, thanks to the continuous progression of mammalian cells engineering, DHFR knockout cell lines are recently developed using different molecular tools (Aga et al., 2015; Brown et al., 2016; Gaj et al., 2013, 2012; Hauschild-Quintern et al., 2013; Santiago et al., 2008). This is one of the cases in which advances in knowledge and biotechnology evolved from a method to a state of the art. There is no literature until now about the new-engineered cell line used in this work; therefore, its performances are unknown. Based on data about DXB11 and DG44 cells, it is possible to hypothesize that this DHFR knockout cell line could guarantee efficient selective pressure and extensive amplification process avoiding difficulties disclosed by DG44 CHO cell line.

CHAPTER 3: Transposon: a powerful approach for efficient transgene delivery

3.1 Transposable elements

Transposable Elements (TEs) are DNA sequences that are able to move from one genetic location to another. First discovered in maize by Dr. Barbara McClintock in the 1940s, TEs came back again and gained the attention they deserved from scientific community only thirty years later. TEs are widely represented in different taxa of organisms, both prokaryotes and eukaryotes. Many hypothesis were formulated trying to explain the function of these particular elements. TEs were considered “selfish DNA”, “parasitic DNA” or “junk DNA” based on the effects they induce on the genome. Indeed, TEs are able to replicate faster than sequences of the host genome, thus, they commonly cause insertional mutations and recombinations. Over the last decades, it has become clear how deeply TEs are involved in the genome evolution. However, the most representative TE type in eukaryotic genome is inactive (Narayanavari et al., 2017; Reilly et al., 2013).

Most of transposable elements can be assigned to two main classes based on their transposition mechanism: Class I and Class II transposable elements. Class I elements, also called retrotransposons, transpose via a RNA intermediate using a copy and paste mechanism. The RNA intermediate is reverse transcribed into DNA by a TE-encoded reverse transcriptase and then integrated into a new genomic site. In this way, TE is maintained in its original position and the additional copy of retrotransposons is moved to another site. Retrotransposons are divided in long terminal repeat (LTR) and non-long terminal repeat (non-LTR) retrotransposons. The latter includes the short interspersed nuclear elements (SINEs) and the long interspersed nuclear elements (LINEs). LINEs function as autonomous elements; instead, SINEs require LINEs machinery for transposition. LTR retrotransposons present long terminal repeats which, flanking the inner coding region, encode for transcriptional regulatory elements (Pritham, 2009). Class II elements, referred as DNA transposons, exploit a DNA intermediate for the movement. This class can be dived in three subclasses: a) *Classic* DNA transposons are excised as dsDNA intermediate and reintegrated elsewhere in the genome by a “cut and paste” mechanism. The element “jumps” from one DNA molecule to another without leaving a copy in the original position; b) *Helitrons*, which represent a second major subclass of DNA transposons, are most likely mobilized as ssDNA

intermediates through a replicative, rolling-circle–like mechanism; c) *Mavericks* encode a viral-like DNA polymerase and a retroviral-like integrase, therefore their transposition cycle involves integration of a dsDNA intermediate (Pritham, 2009).

3.2 *Classic* DNA transposon

DNA transposons are found to be high representative in most of eukaryotic genomes. They comprise about 25% of the genomic DNA of *Xenopus tropicalis* (western clawed frog), about 20% of *Hydra magnipapillata* (hydra) DNA, superior than 19% of *Aedes aegypti* (yellow-fever mosquito) DNA, superior than 13% of *Oryza sativa* ssp. *japonica* (rice) DNA, and superior than 6% of *Phytophthora infestans* (potato late blight) DNA. They are also present in human genome accounting for about four hundred thousand individual elements, i.e. 3% of genomic DNA (Yuan and Wessler, 2011). DNA transposons consist of a transposase gene that is flanked by terminal inverted repeats (TIR) of variable length between different subclasses. The transposase binds TIRs and activate a "cut and paste" mechanism whereby the element is excised from the donor site, causing a double-strand break, and inserted elsewhere in the genome, named target site. A circular pre-excision synaptic complex composed of the transposase in a tetrameric structure brings together the transposon ends. Subsequently, the complex binds the target site and mediates the transposon integration. Breaks in the target DNA are repaired by host enzymes, thereby generating short target–site duplications (TSDs) flanking the inserted transposon. Instead, breaks at the donor site are repaired either by end joining, a process that often leaves 'footprints' at the donor site, reflecting element excision, or homology-dependent gene conversion, using a sister chromatid or homolog as template (Muñoz-López and García-Pérez, 2010). Based on the sequence similarity of the transposase, the length of the TSDs and the terminal nucleotides of the TIRs, eukaryotic transposons can be divided in superfamilies, among which the T*c1/mariner, hAT* and *piggyBac* are the three main ones. The majority of superfamilies, including T*c1/mariner, hAT* and *piggyBac*, presents a catalytic domain composed of a triad of acidic amino acids, called DDE/D motif. This triad consists of two aspartic acid (D) and a glutamic acid (E) residues or three aspartic acid residues forming a RNase H-like structure with α-helices and β-strands. In superfamilies *Tc1/mariner* and *hAT*, the DDE/D triad is folded in a catalytic pocket containing two divalent metal ions that take part in the various nucleophilic reactions during DNA cleavage. In superfamily *piggyBac* the catalytic domain remains unclear (Yuan and Wessler, 2011). The transposition

mechanism works with a first cleavage of one strand at both ends of the element exposing 3'-hydroxyl groups (OH) in both the element and the host DNA ends. The exposition of 3'OH triggers the cleavage of the second strand and the detachment of the transposon from the donor site. The second strand cleavage is performed in three different ways. *Tc1/mariner* transposases make double-strand cuts to release the transposon end from the donor site. *hAT* transposases cleave one strand at NTNNNNAN target site leaving 3'OH in flanking DNA that binds the other strand in a hairpin structure. *Piggybac* transposases cleave one strand leaving 3'OH at the transposon end and generate the hairpin structure with the other strand at the transposon end. This action leaves 5'-TTAA overhangs ends in the flanking DNA. Once the transposons leave the donor site, the 3'OH at their ends attacks simultaneously the target DNA performing the reaction called strand transfer (Figure V) (Skipper et al., 2013).

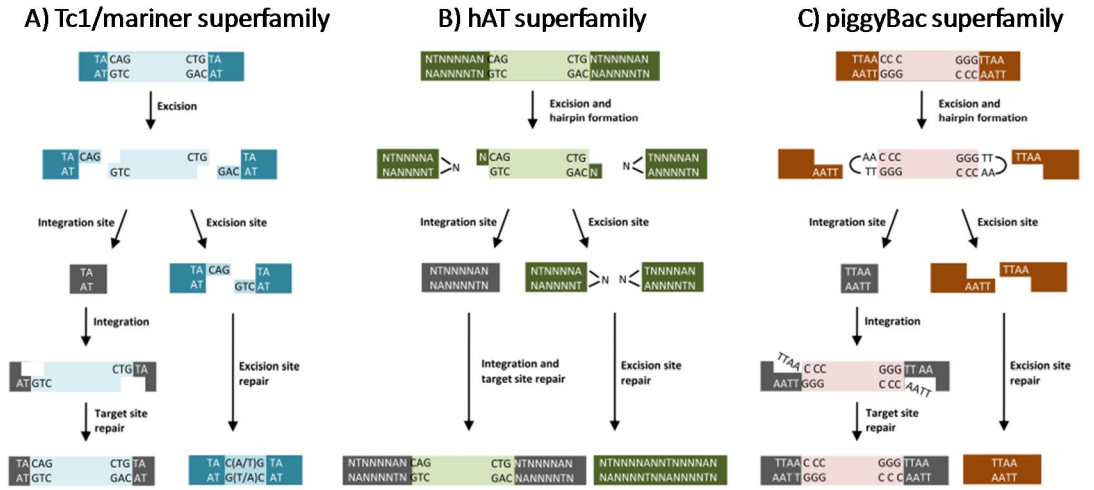


Figure V. Schematic representation of cut and paste transposition: A. Transposition of Tc1/*mariner* elements leads to double-stranded breaks and formation of a 3'-overhang at the excision site. DNA repair by host-encoded enzymes creates a characteristic footprint at the excision site. Integration occurs at TA dinucleotides which are duplicated upon transposition. The single-stranded gaps are repaired by host-encoded enzymes. B. *hAT* transposition creates hairpins at the ends of the flanking donor DNA. Integration is targeted to NTNNNNAN target sites, and creates target-site duplication. Excision site repair leaves a random footprint. C. *PiggyBac*-mediated excision is followed by hairpin-formation at the transposon ends. After integration into TTAA target sites that are consequently duplicated, the single-stranded breaks are repaired by ligation. The 5' TTAA overhangs created at the excision site anneal, thus repairing the double-stranded break without leaving any footprint (Skipper et al., 2013).

3.3 DNA transposons: powerful genomic tools

The native structure of the DNA transposon consists in a sequence, encoding for a transposase, flanked by terminal inverted repeats (TIRs) which act as transposase target sites. A powerful molecular tool, called transposon system, was obtained simply disassembling the transposase sequence from the TIRs binding sites. Thus, these two elements were arranged in two distinct plasmids: a donor plasmid carrying a DNA sequence of interest flanked by the TIRs and a helper vector encoding for the transposase. In this way, a gene of interest can be excised from the donor plasmid and integrated in a chromosomal locus (Figure VI). Therefore, the transposons can be exploited as controllable DNA delivery vehicles (Ivics et al., 1997; Ivics and Izsvák, 2010).

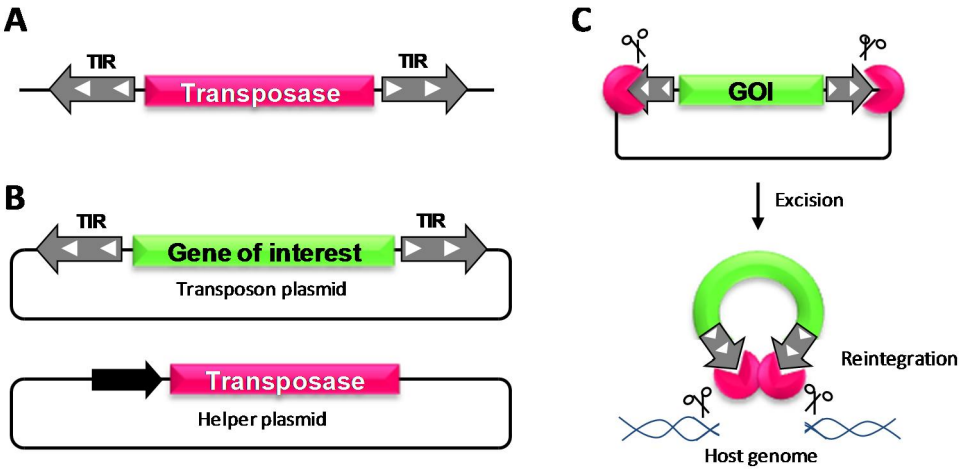


Figure VI. The transposon system: A. Native DNA transposon structure. B. Gene delivery toolbox: the transposase coding region is replaced by a gene of interest within the transposable element and inserted in a plasmid. This transposon is non-autonomous and can be mobilized in trans by a transposase if it is supplied on a separate plasmid (helper). C. The co-transfection of these two plasmids into cells provides the platform for the transposition into the host genome (Ivics and Izsvák, 2010).

Transposons were initially used only in plants and invertebrates for transgenesis and insertional mutagenesis, until the discovery of a fossil transposon in salmonid fish genome. This transposon, belonging to *Tc1/mariner* superfamily, was reactivated thanks to molecular phylogenetic reconstruction. Named *Sleeping Beauty*, due to its re-awaking, it was engineered to be employed in mammalian cells. The *Sleeping Beauty* transposon allowed to overcome the limitations in species compatibility typical of transposons used until its discovery. Indeed, it guarantees highly efficient transposition-mediated gene transfer in most of vertebrate animal models. Since the absence of a transposon analogous to *Sleeping Beauty* in vertebrate genomes, the risk of a potential co-transposition of an

endogenous transposon can be also avoided (Ivics et al., 1997). In a brief period of time, other transposons caught the attention as a potential tool targeting vertebrates, such as *piggyBac*, isolated from the cabbage looper moth, *Trichoplusia ni* (Ding et al., 2005); *Frog prince*, reconstructed from inactive transposons of frog *Rana pipiens* (Miskey et al., 2003); *Hsmar1*, reactivated from an ancient human transposon (Miskey et al., 2007) and *Passport*, isolated from fish *Plauronectes platessa* (Clark et al., 2009). However, among these, only *Sleeping Beauty* and *piggyBac* became the transposon system extensively used as a gene delivery tool.

3.3.1 Features

Over years, research on transposons pointed out several parameters that helped to choose the correct type of transposon in relation to the final application. The features of each transposon system are transposition efficiency, stability, cargo capacity, integration site preference and transposition to linked chromosomal sites ("local hopping").

Efficiency of transposition is perceived as a bottleneck to efficient gene delivery. Aiming to overcome this limitation, hyperactive versions of *Sleeping Beauty* (SB) and *piggyBac* (PB) transposon, called SB100X and hyPBase respectively, were generated. SB100X and hyPBase allowed to obtain higher performances compared to their original versions (Mátés et al., 2009; Yusa et al., 2011). Moreover, since the transposition efficiency decreases if the transposase is in excess, a phenomenon known as "overproduction inhibition", it is important to regulate the ratio between donor plasmid and helper vector concentrations to control the number of transposition events (Narayanavari et al., 2017).

Stability of integration is a great advantage by exploiting transposons since a reduction of about 90% in transgene silencing was observed in clones generated with SB and PB transposon systems compared to standard random integrations (Grabundzija et al., 2010). That is probably because the transposons are less prone to be silenced since they are elements naturally present in genome. Moreover, the transposition mediates the insertion of DNA sequence within the TIRs avoiding the complete integration of whole plasmid backbone (Grabundzija et al., 2010).

Another important aspect of transposition regards the cargo capacity that can change depending on the transposon used. Since the discovery of transposons, viral vectors have offered several attractive properties as stable gene-delivery vehicles. Retroviral and lentiviral vectors have been successfully used for delivering multiple transgenes, but these

systems can carry a limited cargo of up to 8 kb, which is limited by the packaging capacity of their capsid envelope (Thomas et al., 2003). In contrast, SB transposon is able to mobilize around 10 kb fragments in mammalian cells. However, starting from 1.7 kb, that is the most efficient cargo capacity of SB transposon, an exponential decrease in transposition efficiency of about 30% is observed for each extra 1 kb addition (Izsvák et al., 2000). Despite the development of SB systems which helped to improve the transposition efficiency for transgenes longer than 10 kb, such as 'sandwhich SB vector', SB10 and SB100X, the cargo capacity remains a SB transposon limitation (Turchiano et al., 2014). On the other hand, PB system can move larger DNA fragments (9.1-14.3 kb) with a high transposition activity. A 15 kb transposon was successfully transposed in primary human lymphocytes (Huye et al., 2011). Larger DNA fragments like 100 kb were also mobilized using PB system in mouse embryonic stem (ES) cells (Li et al., 2011). Even more longer fragments (207 kb) were successfully introduced into human cells and mice with a modified PB transposon (R. Li et al., 2013). In addition, the high cargo capacity of PB transposon was exploited to perform a multiple transgenes delivery into cells. Indeed, a multiplex PB transposon mediated transfer was successfully pursued in human cells. Stable clones were obtained which expressed three subunits of a functional sodium channel that retained its electro-physiological properties after 35 passages (Kahlig et al., 2010).

Integration site is a peculiar characteristic that should be considered while choosing a transposon system for a given application. Indeed, each transposon superfamily shows a typical integration site preference. The superfamily *hAT* members as *Tol2* are observed to insert within the 5' regions of genes (Kotani et al., 2006). *Sleeping Beauty* transposon, instead, prefers heterochromatin to actively transcribed genes, and if it integrates into genes, it occurs into intronic sequences. It targets into TA dinucleotides with an integration site preference that is dictated by the transposase protein, which also discriminates the secondary structure of DNA (Narayanavari et al., 2017). On the other hand, *piggyBac* transposon integrates into transcription units with a prevalence of insertions in introns, upstream of regulatory sequences and in CpG islands. It targets TTAA sequences probably surrounded by a particular string of 100 nucleotides on either sides, since changes in this interval reduce the transposition efficiency. However, a clear-cut structural pattern, as that of SB, is not recognized (M. A. Li et al., 2013; Meir et al., 2011). According to these considerations, it has become clear that a transposon insertion into genes is more suitable for mutagenesis screening, while this is not the case for gene therapy.

Another important feature is the genomic distance from the transposon donor site. Indeed, a phenomenon known as “local hopping” can occur when already integrated DNA transposons operate a new integration targeting sites that are physically close to the donor locus. This event was first described for *Sleeping Beauty* which showed a transposition interval of local hopping of about 5-15 Mb. However, this feature is not observed in *piggyBac* (Fischer et al., 2001; Liang et al., 2009; Muñoz-López and García-Pérez, 2010). Local hopping can be exploited in mutagenesis to produce insertions in a limited chromosomal region.

3.3.2 Relevance for recombinant protein expression

Once the transposon was engineered as molecular tool, it was applied and exploited in several applications such as insertional mutagenesis, cancer gene discovery, transgenesis and even gene therapy (Ivics and Izsvák, 2010). However, recently, its application in recombinant protein production has also been investigated (Matasci et al., 2011). The production of recombinant protein in mammalian cells is a process mainly based on the selection of a stable cell line. The generation and the characterization of this cell line are challenging, laborious and time-consuming procedures representing one of the most limiting step to achieve a rapid and efficient recombinant protein production. One of the reasons of this limitation is the use of conventional transfection method that relies on random and spontaneous transgene integrations into the host genome. This leads to a low probability of plasmid integration into a highly transcribed region and consequently an insufficient level of protein production, a phenomenon known as “positional effect” (Z. Li et al., 2013). In addition, the transgene integration of a standard vector often takes place as a concatemer insertion due to DNA repair and recombination enzymes activities. This event can cause silencing and instability of the transgene (McBurney et al., 2002). Some strategies can partially overcome the shortcomings encountered using conventional gene transfer methods such as epigenetic regulatory elements whose insertion in vectors enhances the transgene expression (Kwaks and Otte, 2006). As an alternative, a viral transgene delivery can be exploited, but several drawbacks dissuade to apply it; in particular the risks related to its handling and the regulatory concerns (Bouard et al., 2009). In this scenario, the transposon system caught the attention for its potentiality. The powerful of this tool consists in three main advantages. Firstly, the PB transposon system mediates a highly efficient transgene integration in mammalian cells. This is widely demonstrated in CHO and

HEK293 cell lines, which command the market of recombinant protein production (Matasci et al., 2011). Secondly, the PB transposon system mediates an active transposition of a transgene in transcribed regions of the host genomes. This active mechanism allows to insert more copies than traditional plasmid integration. Indeed, the standard vector insertion into the cells genome occurs spontaneously by a classical integration as naturally happens when cells encounter foreign DNA with a low probability to take place in highly transcribed region. Moreover, most of the transfected cells produce low level of recombinant protein. On the contrary, the integration in actively transcribed regions mediated by the PB transposon system allows to enhance the frequency of high-producing clones within the population (Matasci et al., 2011). Thirdly, the PB transposon system mediates spread insertions of the transgene as a single copy into the genome due to the active mechanism of integration. In contrast, the random integration of a standard vector usually occurs in the same genomic locus (Girod et al., 2007). Furthermore, with the PB transposon system, the transgene is transposed without the entire backbone, contrary to the standard vector, avoiding silencing effects. Among the several transposons discovered, *piggyBac* resulted the most suitable system for recombinant protein expression. Indeed, it mediates the transposition into active areas of the mammalian genome resulting in a high protein expression level. Moreover, it is able to move up to 14 kb of DNA without compromising the transposition efficiency. Therefore, using this molecular tool it is possible to achieve high levels of stable protein production (Matasci et al., 2011).

CHAPTER 4: Advances in mammalian protein production

The claim of recombinant proteins continues to increase for both research and industrial applications. In order to meet this demand and be able to carry out studies on proteins, researchers require high amounts of them. A lot of efforts and money are invested in technology advances and strategies to achieve ever more high protein expression level.
Starting from the experimental design until the complex vector engineering, each minimal improvement contributes to discover step by step the fascinating world of the recombinant protein production and to improve the entire process for research and industrial purposes.

4.1 Experimental design and optimization of culture condition

Since research continuously improves each step of the recombinant protein production with also tiny changes to meet market demands, it becomes challenging to explore all different advances. The simplest starting point in large-scale protein expression is the correct definition of the production strategy and the optimization of the culture condition.

4.1.1 Stable and Transient Gene Expression

The transfection introduces foreign nucleic acids into mammalian cells to obtain the expression of heterologous genes. The recombinant DNA is delivered to cells either by mechanical, chemical or biological methods. The transfection can result in Stable or Transient Gene Expression (SGE or TGE) (Figure VII). In SGE a plasmid with the gene of interest and a marker gene for selection is introduced in cells. Generally, selectable markers are genes that confer resistance to various selection drugs or genes that compensate for an essential gene that is defective in the cell line used. When cultured in selective medium, cells that were not transfected will die. On the contrary, transfected cells will survive because they express at sufficient levels a protein that confers antibiotic resistance or compensates for the defect in the essential gene. When the transgene is integrated into the host genome, it is expressed even after the host cells replicate. In contrast, the transiently transfected genes are expressed for a limited period because they are not integrated into the genome. After a week, the presence of the plasmid is no longer detected, it can be lost by environmental factors or diluted by cell division (Kim and Eberwine, 2010).

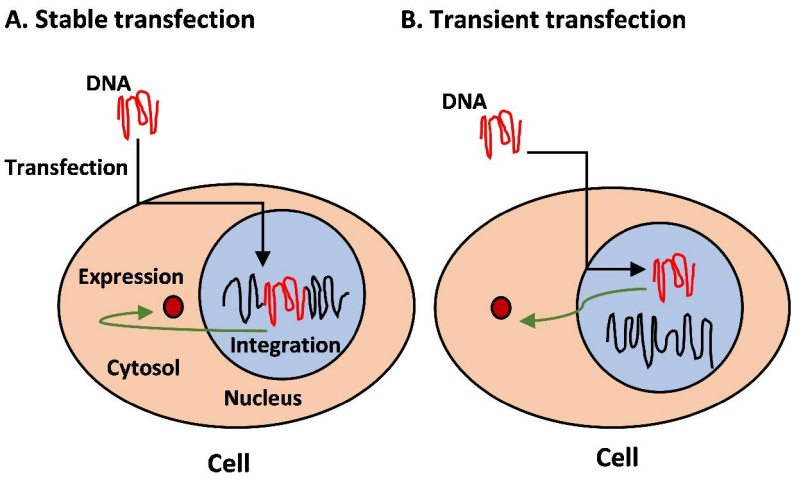


Figure VII. The Transfection system: A. Stable transfection sustains stable gene expression; B. Transient transfection sustains transient gene expression (Adapted from Kim and Eberwine, 2010).

SGE is widely exploited in large-scale industrial production. The process allows to establish stable clonal cell lines that guarantee a long-term consistent recombinant protein production. The starting point to develop a clonal cell line is the transfection of a mother cell line with the plasmid vector carrying the expression cassette and a selection marker gene. There are different methods to deliver the plasmid vector. They can be mainly divided in physical or chemicals methods. Techniques that belong to the first category are electroporation, microinjection and gene gun delivery. Methods of the second category employ positively charged reagents that create a layer around the negatively charged DNA, thus neutralizing its charges. They are widely used and include calcium phosphate co-precipitation, cationic polymers or cationic lipids-based transfection reagents. The transfected cells were then transferred in a selective medium in order to permit the survival of only positive transfectants. The following step is the isolation of single clones that is usually performed by the limiting dilution procedure. A screening allows to select high-producing clones, which are tested for scale-up, purification process, characterization of protein quality and long-term stable expression analysis (Figure VIII) (Ho et al., 2013). In TGE, DNA is introduced into cells usually exploiting chemical methods and it is maintained in cells as an extrachromosomal unit. The transgene expression decreases in about 2 weeks due to its degradation or dilution for cell division. TGE is based on a bulk cell culture transfection with a huge amount of DNA resulting in a high level of protein production thanks to the plasmid copy number per cell (Wulhfard, 2009).

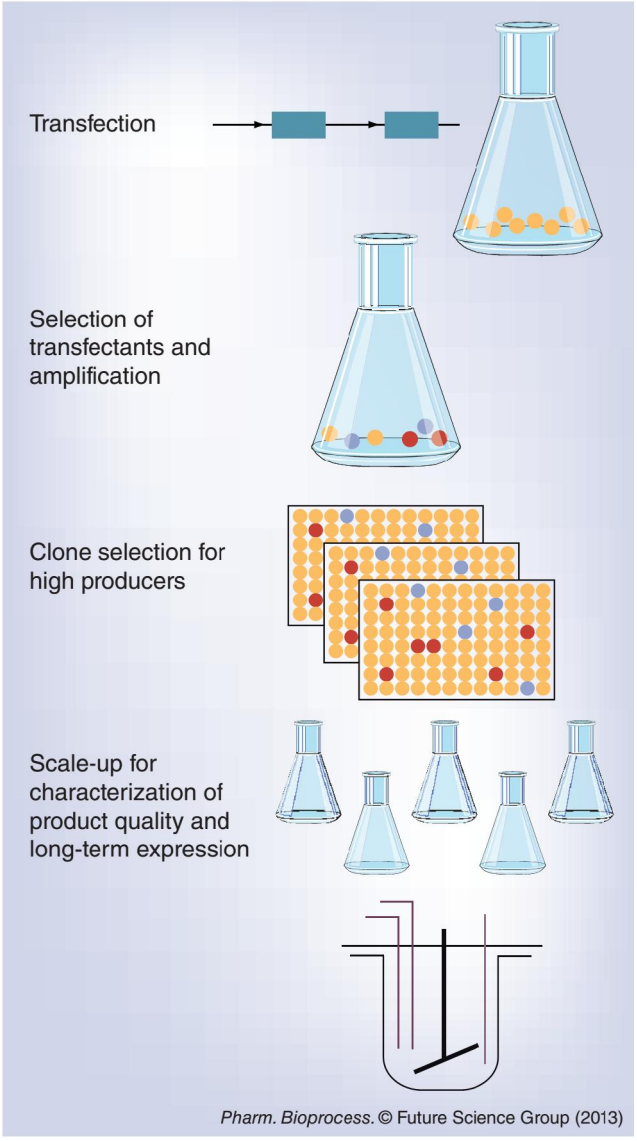


Figure VIII. Stable Gene Expression and scale-up: Schematic representation of the entire stable transfection process and scale-up for production (Ho et al., 2013).

The two transfection methods show advantages and drawbacks, therefore they are exploited based on the goal to be achieved (Table I). SGE guarantees that all cells in the culture express the transgene, which is replicated along with the cellular genome, and it is not lost during cell division. Moreover, SGE generates single clones and allows to obtain higher yields compared to TGE (Liu et al., 2008). Indeed, recombinant protein yields as high as 5 g/L in fed-batch cultures have been achieved. Furthermore, the scalability to 20000 L makes SGE suitable for industrial applications. On the other hand, this technique requires months for the generation and the characterization of cell lines increasing also relative costs (Wulhfard, 2009).

	SGE	TGE
Genetic selection	Yes	No
Time from DNA to product	6-12 months	up to 3 weeks
Specific productivity (pg/cell/day)	up to 50	~ 20
Volumetric productivity (g/L)	up to 5	~ 1
Scalability	up to 20,000 L	up to 100 L
Application	Large scale production of recombinant protein for therapeutics or industry	Small scale production of recombinant protein mainly for research

Table I. Comparison between SGE and TGE: Main differences in SGE and TGE characteristics (Updated from Wulhfard, 2009).

TGE is simpler and faster than SGE because no selection step is required and the period from transfection to product purification is short. Although TGE is an efficient and rapid method to produce recombinant protein with high yield within a few days, it has very high costs that prevent from scaling-up. Moreover, the production is time limited and it is necessary a new transfection for each production. The productivity is moderate compare to that obtained with SGE, even if a yield of 1 g/L was achieved with optimized protocol (Backliwal et al., 2008). For these reasons, by now, therapeutic recombinant proteins are not produced by TGE also because concerns about the heterogeneity of products from batch to batch are to consider. TGE is exploited to produce secreted, intracellular and transmembrane proteins. Therefore, it represents a more versatile transfection method than SGE, in particular for intracellular recombinant proteins because stable cell lines usually do not express them at high levels, unless controlled by an inducible promoter. Nowadays, efficient transient transfectory kits, which comprehend mammalian cell lines, transfection reagents, optimized media and boosts, are available on the market. Thus, TGE represents ever more an attractive and successful approach also for large-scale production reaching up to 100 L in bioreactors. However, since the amount of DNA is about 0,5-1 µg per one million transfected cells, high costs for DNA should be considered (Wulhfard, 2009).

4.1.2 Mild hypothermia culture condition

Advances in cell culture optimization have improved the performance and the robustness of the recombinant protein production also monitoring and controlling the culture parameters. In particular, product quality and potency can be significantly affected by physical (e.g. temperature, gas flow, and impeller speed), chemical (e.g. pH, osmolality, dissolved O2 and CO2, redox potential, and metabolite levels) and biological (e.g. cell concentration, viability, cell cycle, mitochondrial activity, and NADH and LDH levels)

parameters. Optimizing these parameters can significantly enhance the recombinant protein yields while satisfying quality specifications (Kunert and Reinhart, 2016).

The temperature is one of the environmental parameter investigated for improving the recombinant protein production in cells culture. In this regard, a biphasic culture based on a temperature shift is conventionally implemented in many cell culture processes. This procedure is performed on batch cultures consisting of a cell inoculation at a low density, incubated for 24h-72h at 37°C to maximize cell biomass, followed by an exposition to a lower temperature (28-33°C) to encourage increase of the cell specific rate. Even if the molecular mechanisms regulating the response to exposition to low temperature are poorly understood, the impacts on regulation of several key pathways are clearly known. It is reported that the mild hypothermia condition causes several effects that allow to maintain high producing cells with high viability for extended periods: (i) arrest of cell cycle with growth inhibition (ii) decrease of apoptosis (iii) increase of different endogenous proteins expression (iv) reduction of waste metabolites accumulation and nutrient uptake. In this way, the upregulation of the heterologous protein expression is promoted as well as the reduction of intermolecular aggregation of their products (Sunley et al., 2008). The biphasic culture system is a well-established procedure and it allows to increase the recombinant protein or antibodies production of about 2-3 fold in CHO batch culture (Fogolín et al., 2004; Fox et al., 2004; Marchant et al., 2008; Nam et al., 2009; Oguchi et al., 2006; Shi et al., 2005; Yoon et al., 2003). A 38-fold increment of Fab antibody fragments is also reported (Schatz et al., 2003). This technique is exploited in combination with other strategies such as the addition of sodium butyrate or mannose showing the same positive effect (Berrios et al., 2009; Hendrick et al., 2001). It is also demonstrated that the mild hypothermia condition permits to enhance the protein expression without affecting the protein quality and the *in vivo* biological activity (Yoon et al., 2003). It seems that high protein yields obtained with low temperature cultivation are not due to the reduced protease activity, even if there is no universal opinion on that (Clark et al., 2004). A number of transcriptomic and proteomic profiling studies were performed suggesting that mammalian cells actively respond to low temperature by synthesizing specific cold-inducible proteins and by changings in post-translational protein modifications (Kaufmann et al., 1999; Kumar et al., 2008). Anyway, further studies are necessary to better understand the molecular mechanisms that characterize the cold shock response in order to improve even more the process. About ten years ago, the mild hypothermia condition has been shown to enhance

also TGE in CHO cells (Galbraith et al., 2006; Underhill et al., 2006). Shortly after, the highest level of transient antibody production in mammalian cells (corresponding to 60-80 mg/L) was reached in batch culture by 6 days post-transfection exploiting the mild hypothermia culture condition (Wulhfard et al., 2008). Nowadays, the low temperature growth condition in TGE is widely exploited and several efficient transient transfectory kits reporting mild hypothermia protocols are available on the market and allow to obtain g/L of recombinant protein. Even if the cold shock response in mammalian cells is not completely understood, the mild hypothermia condition results a simple strategy to greatly improve the recombinant protein expression in batch or fed-batch culture.

4.2 Vector engineering strategies

In the previous chapter, the major characteristics of the stable expression system were disclosed. In this technique, the isolation of clones that combine high protein expression and good growth condition is the pivotal point. Optimization of cell culture condition can be a valid method to enhance the protein expression, but it cannot help if stability problems occur. Advances in knowledge on molecular mechanisms of the protein expression pathway in mammalian cells allowed to understand which elements could push towards high level of protein production also guaranteeing long-term expression stability. Therefore, elegant strategies are recently implemented focusing on the vector design which greatly improve the protein yield and the stability over time (Ludwig, 2006).

4.2.1 Destabilization in selection markers

Generation of stable producer lines requires the drug selection step to permit integration in the host genome. High producing clones can be obtained by strengthening the selective pressure. However, this is a critical balance because a too strong drug selection could result in a loss of successful clones due to the reduction in their growth rate (Kim et al., 2001; Kim and Lee, 1999; Lai et al., 2013). Selective markers in the mammalian cell system can be divided in two categories based on their activity: they can confer resistance against antibiotics or function as metabolic enzymes. The first category consists of antibiotic resistance genes. The most common is bacterial neomycin-kanamycin phosphotransferase type II enzyme (Neo gene) that confers resistance to G418 drug. Other antibiotic selection markers are available, such as puromycin, zeocin and blasticidin, but they are less

widespread than neomycin one. Metabolic selection markers such as the hamster gene encoding glutamine synthetase (GS) and the mouse dihydrofolate reductase (DHFR) gene belong to the second category (Ludwig, 2006). These selection systems require a proper cell line defective in the corresponding metabolic enzyme. Moreover, in this case, it is also possible to perform the gene amplification process employing metabolic inhibitors for those enzymes (i.e. methionine sulfoximine and methotrexate, respectively). Strategies to increase speed and efficiency of the selection methods have been developed in order to move the pivotal balance toward the achievement of high producing clones. The concept is to attenuate the activity of the selection marker so that: (i) in case of first selective marker category, high producer lines are still efficiently selected at lower doses of the drug reducing its side-effects and (ii) for the second marker category, transfected cells are forced to be more productive in the locus of the selection marker resulting in the high productivity of the adjacent gene of interest. Two approaches are implemented aiming the maintenance of high selection stringency at low drug concentrations. The first strategy consists of the insertion of mutations into the selection marker in order to reduce its activity. It was demonstrated to be a successful method on neomycin phosphotransferase II since its affinity reduction to neomycin led to improve the specific monoclonal antibody productivity of several fold till about 17 (Lai et al., 2013). The second strategy aims to modulate the expression pathway of the selection marker at various steps in order to reduce the enzyme expression. Different methods were implemented and, based on the targeted step, they can be divided in three main groups as follows. Firstly, the transcription efficiency can be impaired acting on DNA level with a codon deoptimization of the selection marker gene. This can be performed employing the least preferred codons of the expression host (Westwood et al., 2010). Secondly, the transcription can be reduced acting on mRNA level by placing the selection marker gene under the control of a weak promoter. Otherwise, it is possible to introduce a low efficient internal ribosomal entry site (IRES) in order to reduce the ribosome-mRNA affinity (Ho et al., 2012; Ng et al., 2012; Niwa et al., 1991). IRESs are sequences that enable the creation of functional bicistronic mRNAs with a high level of cap-independent translation activity. Therefore, both the gene of interest and the selection marker are transcribed into a single mRNA transcript that, in turn, is translated into two proteins. This expedient is widely employed in DHFR deficient cells because the co-localization of the two genes can improve the MTX amplification process while mitigating the gene fragmentation. The most diffuse element in experimental application is the IRES

from encephalomyocarditis virus (EMCV) (Bochkov and Palmenberg, 2006; Ludwig, 2006). An attenuated IRES sequence can be inserted in the construct aiming to destabilize the selection marker and, thus, to strengthen the selection stringency (Ng et al., 2012). An alternative to reduce the transcription acting on mRNA level is the use of adenylate-uridylate (AU) rich elements (ARE) to destabilize the selection marker mRNA. AREs are found in the 3' untranslated region (UTR) of many mRNAs that code for proto-oncogenes, nuclear transcription factors and cytokines. They represent the most common determinant of RNA stability in mammalian cells (Chen and Shyu, 1995). Thirdly, the protein half-life can be reduced acting on protein level with the insertion of murine ornithine decarboxylase (MODC) PEST region (Li et al., 1998; Ng et al., 2007). PEST regions are clusters rich in proline, aspartate, glutamate, serine and threonine. PEST sequences were found in the carboxy-terminus of MODC, an enzyme that catalyses the conversion of ornithine into putrescine in the polyamines biosynthesis of mammalian cells. One of the two PEST regions present in MODC protein is included in a 37 aminoacids sequence that is reported to be responsible for the degradation of MODC enzyme. The intracellular regulation of this enzyme is very fine and its half-life is critical to be short. Thus, PEST sequences became a tool for protein manipulation (Ng et al., 2007).

Each element belonging to the second strategy was successfully explored in CHO DHFR deficient cells using the MTX amplification system. The codon deoptimization was explored on the DHFR selection marker at different deoptimization levels demonstrating a reduction in expression level of the DHFR protein in CHO DG44 transfected pools and derived clones. This event corresponded to an increase in transfection selection efficiency which, in turns, led to enhance the production of the protein of interest. Unexpectedly, the more deoptimized marker resulted less stringent than the less deoptimized one (Westwood et al., 2010). The use of the ARE and MODC PEST sequences, which acted on the DHFR selection marker in CHO DG44 cells, allowed to enhance the IFNg protein productivity of 1,7- and 6,6-fold, respectively, and of 13,3-fold in combination. Moreover, this strategy resulted compatible with the MTX amplification system reaching an increment of protein production until 27-fold after treatment with 50 nM drug concentration (Ng et al., 2007). The PEST sequence in association with an attenuated IRES acting on the DHFR transgene was also explored for the Dectin-1 protein production resulting in the highest reported titer for small recombinant protein productions (<50 kDa) in a stable CHO DG44 cell pool (titers of 246

mg/L in shake flask batch culture and 598 mg/L in bioreactor fed-batch culture). The protein maintained also its characteristic and functional properties (Ng et al., 2012).

Strategies acting on selection marker emerged as attractive alternatives to enhance the recombinant protein production maintaining a strength selective pressure without undesirable side effects of high drug concentrations.

4.2.2 Epigenetic regulatory elements

High expression level and long-term expression stability are essential requirements for industrial recombinant protein expression in mammalian cells. However, a key issue that affects protein production is the expression instability of the clonal cell line over time. Gene silencing often occurs with consequent reduction or even suppression of gene expression. Transcription unit engineering within plasmid vectors can have a significant impact in achieving stable high-level production. The stability of a transgene is influenced by the presence of specific DNA elements and the structure of surrounding chromatin at the site of integration. Moreover, independently from the transgene site integration, transcriptional silencing can also occur via a DNA methylation-mediate mechanism. Strategies based on the insertion of cis-acting DNA elements in plasmid vectors are successfully implemented to reduce or eliminate epigenetic modifications that negatively affect the transgene expression. These elements are known as epigenetic regulatory elements. Some epigenetic regulatory elements show essential characteristics that make them worth for industrial application. In particular: (i) the length of the element not exceeding 4-6 kb is important to avoid construct instability or difficulties on cloning; (ii) universal applicability is another key point because they should be exploited in several commercial cell lines and in combination with different promoters; finally, (iii) conferring stable and high expression levels of the transgene and in copy-dependent manner is the ideal requirement (Kwaks and Otte, 2006). Insights into epigenetic gene regulation have resulted in the development of tools to achieve higher and more reliable protein production. The most common epigenetic regulatory elements used in molecular biology are locus control regions (LCR), insulator such as DNaseI hypersensitive site 4 of the chicken β-globin locus control region (cHS4), ubiquitous chromatin opening element (UCOE), matrix associated region (MAR) and stabilizing anti-repressor element (STAR).

LCR

Locus Control Region (LCR) is the most prominent distal regulatory element in the human β-globin locus control region. It is composed of five domains that exhibit high sensitivity to DNaseI in erythroid cells (hypersensitive sites – HR). Its length reaches about 16 kb. It is required for high-level globin gene expression at all developmental stages. LCR mediates copy-dependent and integration-site independent expression of globin or heterologous gene when linked to them. However, its features make it less suitable as tool for increasing the recombinant protein production because it is too large and it shows a high tissue-specificity, working properly only in erytropoietic cell lines. Exactly the same drawbacks are observed in other LCRs such as the immunoglobulin γ5-V_{preB1}LCR and the human growth hormone LCR. Therefore, the applicability of LCR is mainly limited to the creation of transgenic animals (Kwaks and Otte, 2006; Levings and Bungert, 2002).

cHS4

Based on their functions, chromatin insulators can be divided in barrier insulators, which protect chromosomal domains from heterochromatinization, and enhancer-blocking insulators, which prevent interaction between regulatory elements of different chromatin domains. Some insulators display both activities. The best-studied mammalian insulator is the DNaseI hypersensitive site 4 of the chicken β-globin locus control region (cHS4), a 1.2-kb sequence located at the 59 end of the chicken β-globin domain. cHS4 combines the enhancer blocking and the barrier functions. The enhancer-blocking activity is mediated by binding the transcriptional factor CTCF (CCCTC-binding factor). For the barrier activity, instead, three of five protein factor binding sites located in a 'core' region of 250 bp are involved. However, given that the 250 bp 'core' can be influenced by distance, an extended core element is suggested to confer the full barrier activity (Aker et al., 2007; Liu et al., 2015; Saunders et al., 2015; Sharma et al., 2012). Since cHS4 is demonstrated to protect transcribed regions from chromosomal position effects and from the spreading of inactive chromatin, cHS4 caught the attention as a possible molecular tool. Indeed, once experimented it was shown to decrease the probability of transgene silencing thanks to its barrier function and to reduce the activation of regulatory elements thanks to its blocking activity (Liu et al., 2015; Sharma et al., 2012). Upon its insertion into retroviral vector, cHS4 improved the expression of the integrated transgene cassettes and overcame their silencing suggesting that it can be useful in gene therapy applications. Moreover, cHS4

displayed its activity also when incorporated in lentiviral vectors increasing the transgene expression (Rivella et al., 2000; Sharma et al., 2012). cHS4 was successfully exploited as insulator with beneficial effects in different hosts such as *Drosophilia*, human K562 erythroleukemia tissue culture cells, transgenic mice and oncoretroviral vectors, suggesting its universal applicability. These evidence are in contrast with the mild effect in insulating CMV promoter from position effects obtained in CHO cells, which could infer an hypothesis of erythropoietic cell line-specificity restricting its applicability (Kwaks and Otte, 2006; Saunders et al., 2015). cHS4 was shown to increase the stability of the transgene expression in embryonic cells when inserted in *Sleeping Beauty* transposon (Dalsgaard et al., 2009). Recently, cHS4 was confirmed to act positively when incorporated in both *Sleeping Beauty* (SB) and *piggyBac* (PB) transposons and transfected in retinal pigment epithelium cells. Upon cHS4 insertion in a vector, an increase in transgene expression levels was obtained with both SB and PB transposon transfections. However, a long term protective effect of cHS4 against progressive transgene silencing effects was not observed (Sharma et al., 2012).

UCOE

Ubiquitous chromatin opening elements (UCOEs) are promoter regions containing CpG islands of ubiquitously expressed housekeeping genes. It is demonstrated that UCOEs maintain a favorable chromatin conformation and protect the genes in their proximity from silencing, thus resulting in a high and consistent gene expression (Antoniou et al., 2003; Lindahl Allen and Antoniou, 2007; Williams et al., 2005). The best two studied UCOEs are sequences that contain a pair of divergent gene promoters from the Human TATA binding protein (TBP) - proteasome component-B1 (PSMB1) gene loci and from the heterogeneous nuclear ribonucleoprotein A2/B1 (HNRPA2B1) - heterochromatin protein 1Hs-γ (chromobox homolog 3, CBX3) gene loci, which are actively transcribed in all cells of an organism. At the beginning, large UCOEs of 12-16 kb were investigated. Exploiting these elements, high level and long-term stable transgene expression were achieved (Antoniou et al., 2003). These properties were maintained when a short fragment of 8 kb was used in association with CMV promoter which was known to be susceptible to inhibition by RNA interference and methylation of CpG regions (Benton et al., 2002; Williams et al., 2005). Moreover, the protein expression resulted in a substantial increase of 20- to 40-fold (Williams et al., 2005) and a rapid scalability for antibody production was achieved from the isolation of CHO high

producing clone that led to obtain 200 mg/L of antibody in bioreactor (Benton et al., 2002). Nowadays, even compact fragments are employed, ranging from 1.5 to 3 kb, in order to avoid stability and cloning problems and to be incorporated into most viral vectors. Despite their smaller size, these UCOE fragments completely retained their activity in high producing clones generated (Nair et al., 2011; Williams et al., 2005). UCOEs are suggested to display their activity due to the extended large region of unmethylated DNA in combination to a distinct histone modification pattern. These properties guarantee a reproducible and stable transgene expression also within centromeric heterochromatin (Lindahl Allen and Antoniou, 2007). The significant efficacy of UCOEs is reported in several articles, which also demonstrated the versatility of these elements. Indeed, they show their activity in different cell lines, unlike other regulatory elements such as LCR and enhancers, and in various vectors including lentiviral ones encouraging the retroviral vector-based gene therapy (Benton et al., 2002; Boscolo et al., 2012; Hou et al., 2014; Lindahl Allen and Antoniou, 2007; Majocchi et al., 2014; Nair et al., 2011; Neville et al., 2017; Williams et al., 2005; Zhang et al., 2010). UCOEs were also successfully applied in multipotent and pluripotent stem cells, preventing epigenetic silencing and variegation in cell lines. These results strongly reinforce the UCOEs application to stabilize the transgene expression in PSC-based cell and gene therapy approaches (Ackermann et al., 2014; Müller-Kuller et al., 2015). A comparative analysis among cHS4, MAR, STAR and UCOE elements in CHO cells demonstrated that UCOE was the most effective. It allowed to increase antibody yields in both stable cell pools and clones. In addition, UCOE counteracted the effects of progressively reduction of antibody expression and transgene promoter DNA methylation (Saunders et al., 2015). On the other hand, UCOE failures in enhancing the transgene expression in CHO cells are also reported (Otte et al., 2007). Moreover, UCOE is suggested to be promoter specific due to its interaction with promoter regulatory elements (Nair et al., 2011).

MAR

Matrix Attachment Regions (MARs) are DNA elements that anchor the chromatin structure to the nuclear matrix during interphase of the cell cycle. It is supposed they act creating a structural order in the chromatin protecting against position effects and enhancing the gene expression. MARs contain a topoisomerase consensus sequence and other AT-rich sequence motifs, their size is about 3 kb. They are reported to operate independently from

the promoter (Kim et al., 2004; Wang et al., 2012; Zahn-Zabal et al., 2001), even if a recent study showed that using two different MARs in combination, the protein expression level changed based on the promoter inserted (Ho et al., 2015). It is also reported that, unfortunately, MAR elements do not convey copy-number dependent gene expression (Kwaks and Otte, 2006). These elements are widely employed to improve the recombinant protein and antibody production. Since the incorporation of these elements allows to augment the number of high-producing clones, decreasing the number of silenced cells and clonal heterogeneity, efforts on the clones screening and, in turn, the production timeline are reduced. MAR sequences are hypothesized to promote the homologous recombination, but the molecular mechanism of action is not clearly understood (Grandjean et al., 2011). In a recent work, a deletion analysis of the human β-globin MAR sequence was performed to disclose which regions are essential to enhance the transgene expression in mammalian cells. Results showed that AT-rich sequence and specific transcription factor-binding motifs mediated the transcriptional activation of transgenes (Sun et al., 2016). Several MARs were investigated, such as MARs derived from β-globin or β-interferon (IFN) human loci or from chicken lysozyme locus, which had a positive effect on the transgene expression (Girod et al., 2007, 2005; Kim et al., 2004; Zahn-Zabal et al., 2001). Wang and collaborators showed that two β-globin MARs that flank the expression cassette resulted in a higher expression level of a reporter gene in stably transfected CHO cells compared to one β-globin MAR insertion (Wang et al., 2010). A couple of years later they evaluated the expression level of the same reporter gene transfected with constructs containing two different MARs (β-globin and β-IFN MAR) compared to two identical β-globin MARs. They demonstrated that the transfection with the construct carrying the two different MARs effectively increased the protein level of about 4.5-fold compared to that with the two identical β-globin MAR and of about 46.4-fold compared to the control (a plasmid without MARs) (Wang et al., 2012). More recently, the same combinations of MAR sequences were investigated in association with different promoters. Results demonstrated that the level of stably eGFP expression was higher under the CMV promoter control when the two MARs differed, but not under the SV40 one. For SV40 promoter the combination of two identical β-globin MARs was more advantageous (Zhao et al., 2017). About ten years ago, more potent MAR sequences, named human MAR 1-68 and MAR x-29, were computationally identified and successfully tested confirming their validity. This approach paved the way for ever more powerful tools to enhance the recombinant protein production (Girod et al., 2007). MARs

were also studied in transient protein expression where it was demonstrated to positively display their activity, even if this evaluation was not universally accepted. In a recent work, a new MAR from the mouse genome was isolated with in silico approach and it was assessed to increase the IgG productivity in transient transfections (Harraghy et al., 2011). Furthermore, MAR elements were studied with the *piggyBac* transposon system highlighting that the positive effects of MAR sequences acted also in association with transposons. In this case, it is necessary to evaluate the position in which MAR elements are inserted. Indeed, if a MAR element is positioned at the transposon edges, the transposition can be impaired due to topological constrains while it does not occur if it is added at a central position. However, this event seemed to be specific for human MAR 1-68 because no impairment were observed with human MAR X-29. A couple of relevant remarks are that a transposon containing MAR coupled with a strong promoter resulted in the highest protein expression level and that a few transposed transgene copies correspond to high expression level (Ley et al., 2013).

STAR

Stabilizing anti-repressor elements (STARs) are suggested to contrast chromatin associated repression effects. They were identified by a screening of a library of human genomic DNA fragments where only cells containing the genetic elements capable of blocking the spreading of heterochromatin-related proteins, such as HP1, could survive. Comparative analysis demonstrated that STARs are elements highly conserved between human and mouse. Incorporation of these elements into the expression cassette allowed to obtain more colonies expressing the transgene at a high level. Anyway, the mechanism of action is not clearly understood. They are universal applicable since they showed cell line and promoter-independent activity. They also convey copy number-dependent expression and permit to preserve long-term stable expression. Among a series of regulatory elements ranging from 0.5 to 2 kb in length, STAR element 7 and 40 were noticed as the most promising. Compared to various regulatory elements, STARs activity resulted the most effective in generating high-producing CHO cell lines in works of Kwaks and collaborators (Kwaks et al., 2003; Kwaks and Otte, 2006). On the contrary, another comparison study contradicted this consideration (Saunders et al., 2015). Moreover, STARs did not show their robust role in preventing silencing of telomeric genes in HeLa cell line (Majocchi et al., 2014).

CHAPTER 5: Model protein

5.1 AMH gene and protein features

In 1830, Johannes Peter Müller, a physiologist, published *De glandularum secernetium and Bildungsgeschichte der Genitalien*. In his compendium, he described the paramesonephric ducts (after the fact named Müllerian ducts), a pair of embryo ducts that are the precursor of female reproductive organs ("Müller, Johannes Peter nell'Enciclopedia Treccani," n.d.). Since Müllerian ducts are present in both male and female embryos, during male embryonic development, a fine regulation, which involves the Anti-Müllerian Hormone (AMH), occurs to arrest the differentiation in female sex. Therefore, in male embryos, AMH is secreted by Sertoli cells of the testis and the interaction with its receptor, located on the surface of Müllerian ducts cells, causes Müllerian ducts regression and prevents the "default" development of the female plumbing. In females, who do not produce the AMH protein during foetal development, the Müllerian ducts become the uterus and Fallopian tubes (Rey et al., 2016). AMH, also named Müllerian inhibiting substance (MIS), is a member of the transforming growth factor-β (TGFβ) superfamily. TGFβ superfamily is a group of growth factors that regulate several cell growth and differentiation pathways. AMH is a glycosylated disulfide linked homodimer and it is synthesized as 560 amino acids precursor (pre-proAMH) (Figure IX).

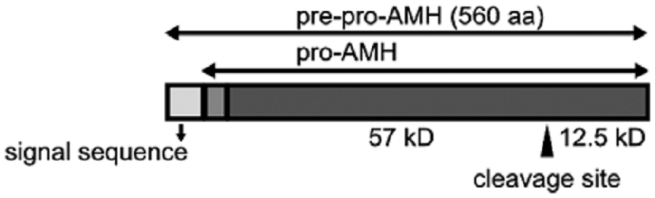


Figure IX. AMH protein maturation: Schematic representation of AMH processing (Rey, 2005).

It has a 24 amino acids signal sequence that is removed during synthesis obtaining the prohormone (proAMH). Full bioactivity is achieved upon proteolytic cleavage at a kex-like site characterized by $R^{-4}XXR^{-1}$ motif between amino acids 451 and 452 resulting in a 25 kDa C-terminal dimer and 110 kDa N-terminal dimer. Those dimers remain associated in a non-covalent complex that is the active form of the protein. The receptor-activating function of total AMH resides in the C-terminal fragment, which shares homology with members of TGFβ family. The N-terminal domain helps folding, stability and solubility of the C-terminal

one. AMH can be experimentally cleaved using plasmin. Also two members of the proprotein convertase family, PC5 and furin, have been reported to be able to process AMH in embryonic rat testes (di Clemente et al., 2010; Nachtigal and Ingraham, 1996; Papakostas et al., 2010; Pepinsky et al., 1988; Rey, 2005).

AMH is encoded by a 2.75 kb gene located in the short arm of the chromosome 19 at position 13.3 (19p13.3). The gene is divided into 5 exons characterized by a high GC content (Figure X). The consensus TATA or CCAAT sequences are not present in human AMH promoter contrary to mouse and rat promoters that contain TATA box. Initially, a functional initiator (Inr) element, which is specifically recognized by the transcription factor TFII-I, was identified in the transcription of the human AMH gene. Subsequently, a major transcription initiation site and three minor sites, a putative oestrogen response element at –1772 and a Sp1 binding site at –303 were characterized at the 5' flanking region. Binding elements for SOX/SRY proteins, SF-1 and GATA factors were also identified as proximal promoter sequences. However, the presence of the SAP62/SF3A2 housekeeping gene at –789 of the human AMH ATG codon, which encodes for a spliceosome protein, suggests that the promoter could be longer.

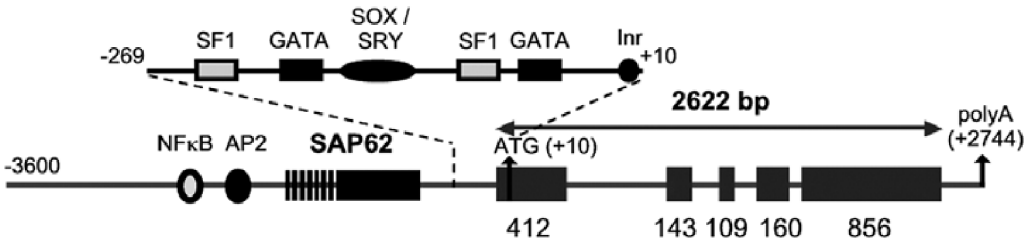


Figure X. AMH gene: Schematic representation of AMH gene and its promoter region (Rey, 2005).

Mutations in AMH gene can cause the persistent Müllerian duct syndrome (PMDS), a disorder that affects the sexual development in males. At least 38 mutations, for the most part missense, are associated to PMDS; these mutations cause an alteration in the correct release or bioactivity of the protein. As a result, the Müllerian ducts are developed in female organs in addition to normal male ones, since AMH protein is not involved in the formation of male reproductive organs. AMH role in human ovarian function does not appear crucial. Women carrying the same mutations of PMDS patients are fertile. It is only supposed a correlation between AMH mutations in female and an early menopause (Belville et al., 2004; Josso et al., 2005, 1997; Rey, 2005).

5.2 AMH signaling

AMH protein was first isolated and purified in 1984 and its gene sequenced in 1986. It was not until 1994 that its receptor was sequenced and cloned allowing the AMH signaling pathway to be unravelled. This receptor is located on the surface of Müllerian ducts cells and it is responsible for the ligand binding. Pro-AMH is cleaved at the kex-like site by the furin or a related prohormone convertase, PC5, in N-terminal and C-terminal dimers, which remain associated in a non-covalent complex (Figure XI). The complex binds to the AMH receptor type II through its C-term domain and this interaction promotes the dissociation of the N-term domain. The recruitment of the AMH receptor type I is necessary to transduce the signal, however, the interaction is not clearly understood. It is supposed that AMH binds the AMH receptor type I once the dissociation of the N-term occurs.

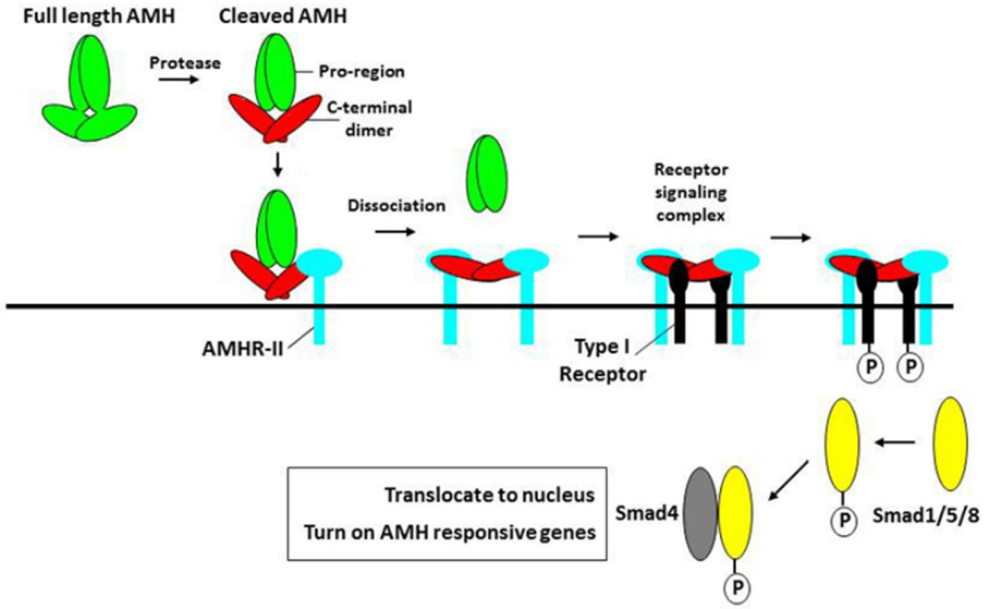


Figure XI. Model of AMH signaling pathway: Schematic representation of the AMH processing and signaling transduction (di Clemente et al., 2010; Rey et al., 2016).

In addition, the intracellular transduction pathways after the type I receptor recruitment seem to vary according to the target cell. Three different type I receptors are considered to be involved: ALK6 (BMPRI-B), ALK2 (ActRI), and ALK3 (BMPRI-A). The AMH receptor type II and type I are single transmembrane serine/threonine kinase receptors. Upon its activation, the AMH receptor type I phosphorylates the cytoplasmatic Smad 1, 5 e 8. The activated Smads form heteromeric complexes with a common partner, Smad4, and move to the nucleus, where they control the gene expression. AMH receptors are also present in granulosa cells of the ovary and in Sertoli and Leydig cells of the testis. It is reported that

AMH acts inhibiting the Leydig cell differentiation and steroidogenesis and it is involved in the granulosa cell response to LH and FSH. However, it remains unclear if AMH plays any essential role in the gonadal physiology (di Clemente et al., 2010; Dumont et al., 2015; Nachtigal and Ingraham, 1996; Papakostas et al., 2010; Rey, 2005).

5.3 AMH function and regulation

AMH has been predominantly known for its role in sexual differentiation. During its development, embryo shows bipotential primordial of external genitalia and two pairs of unipotential internal ducts: the mesonephric (Wolffian) and the paramesonephric ducts (Müllerian). During the 7th week of foetal life, internal and external genitalia develop under control of testis differentiation (Figure XII).

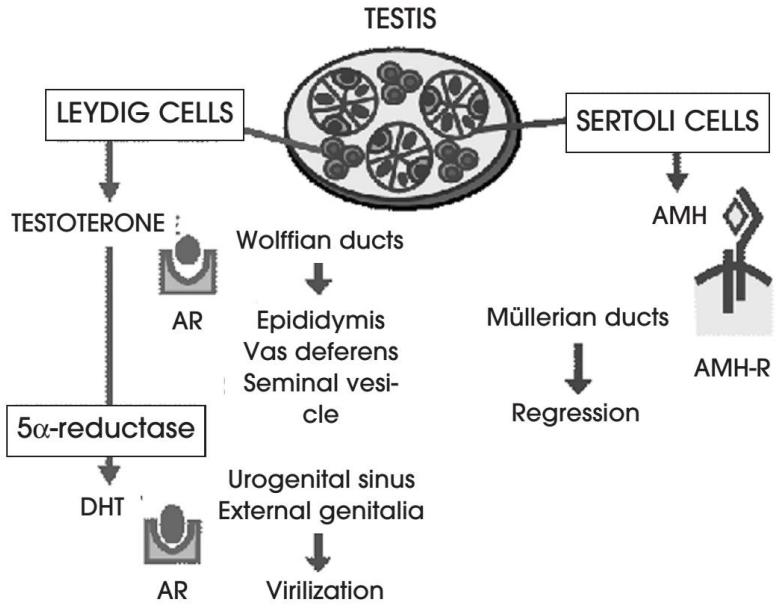


Figure XII. The hormonal control of male foetal sex differentiation: Leydig cells secrete testosterone, which drives Wolffian duct differentiation into the epididymes, vasa deferentia and seminal vesicles acting through the androgen receptor. In the ontogenesis of the external genitalia, testosterone is transformed by 5α-reductase into dihydrotestosterone (DHT), a more potent androgen that binds the androgen receptor to induce external virilisation. Sertoli cells secrete anti-Müllerian hormone (AMH), which binds to a membrane receptor in Müllerian ducts and provokes their regression (Rey, 2005).

Leydig cells secrete androgens which determinate masculinization of Wolffian ducts, the urogenital sinus and external genitalia, whereas Sertoli cells secrete AMH which drives the regression of Müllerian ducts (Matzuk and Lamb, 2008; Rey, 2005; Rey et al., 2016).

The first theorization of the presence of AMH was formulated by the French scientist Alfred Jost in 1940. He conducted experiments on castrated foetuses grafted with testicular tissue

or treated with testosterone. He observed that castrated fetuses grafted with testicular tissue allowed to develop Wolffian ducts in epididymes, vasa deferentia and seminal vesicles while Müllerian ducts regressed. When he treated the castrated fetuses with testosterone, Wolffian ducts correctly proceeded with their male differentiation, but also Müllerian ducts completely developed in Fallopian tubes, the uterus and the upper third of the vagina. Dr. Jost demonstrated that testosterone is not involved in Müllerian ducts regression, which is consequently ascribable to a different testicular hormone that he named "hormone inhibitrice" or "Müllerian inhibitor". However, only in 1978 the AMH protein was identify.

During female sex development, AMH is not detected in the ovary. It starts to be expressed at about the 36th week of human foetus gestation and continues until menopause. AMH is expressed at all steps of folliculogenesis (Figure XIII).

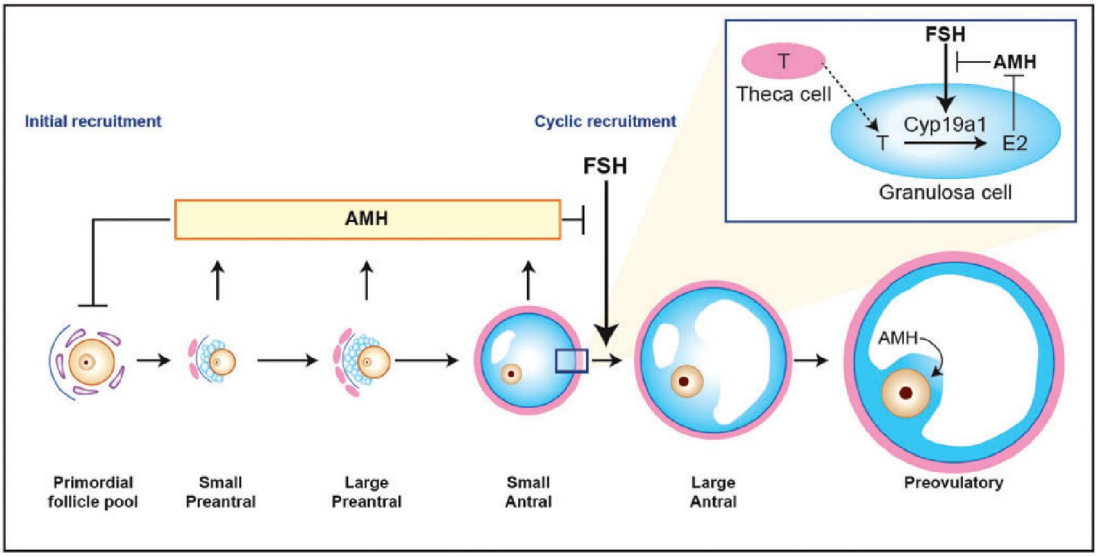


Figure XIII. Schematic model of AMH actions in the ovary: AMH, produced by the granulosa cells of small growing follicles, inhibits initial follicle recruitment and FSH-dependent growth and selection of pre-antral and small antral follicles. In addition, AMH remains highly expressed in cumulus cells of mature follicles. The inset shows in more detail the inhibitory effect of AMH on FSH-induced CYP19a1 expression leading to reduced estradiol (E2) levels, and the inhibitory effect of E2 itself on AMH expression. T, testosterone; Cyp19a1, aromatase (Dewailly et al., 2014).

Granulose cells of primary follicles started to express AMH and continue to do it during the maturation of follicles, even though a gradual reduction in AMH expression occurs as the diameter of the follicle increases. Upon growing follicles reach a diameter of around 8 mm, AMH expression rapidly decreases due to the selection for dominance by the action of pituitary follicle-stimulating hormone (FSH). This is a state of transition from a low-estrogen production to a high-estrogen one. The inhibitory effect on AMH is mediated by estradiol

(E2) through the binding with its receptor b. In sum, AMH is expressed in follicles recruited from the primordial pool but not yet selected by dominance. This underline the pivotal role of AMH as follicular gatekeeper. AMH ensures a little E2 production in small antral follicles prior to selection and plays a fundamental role in ovarian/pituitary balance regulating the number of growing follicles and the development of the selected follicle that will undergo ovulation (Dewailly et al., 2014; Dumont et al., 2015; Visser and Themmen, 2005).

5.4 AMH as biomarker

Since AMH plays a fundamental role in gender differentiation and sexual development, its serum levels determination resulted attractive for assessment of gonadal status and their relative disorders in both males and females. In male, AMH reaches the lowest levels in the first days after birth. After the first week, testicular AMH production starts to increase due to the proliferation of Sertoli cells and rises rapidly during the first month with a peak at about 6 months of age. During childhood AMH levels drastically decline until become very poor in puberty (Figure XIV). This event occurs in conjunction with increasing levels of testosterone. However, AMH decrement is not ascribable only to testosterone increment, because AMH is not influenced by the presence of the androgen hormone during foetal development. The explanation is the absence of androgen receptor in Sertoli cells of foetus. The lack of Sertoli cells response to the androgen hormone consents the maintenance of the high AMH level (Rey, 2005; Rey et al., 2016).

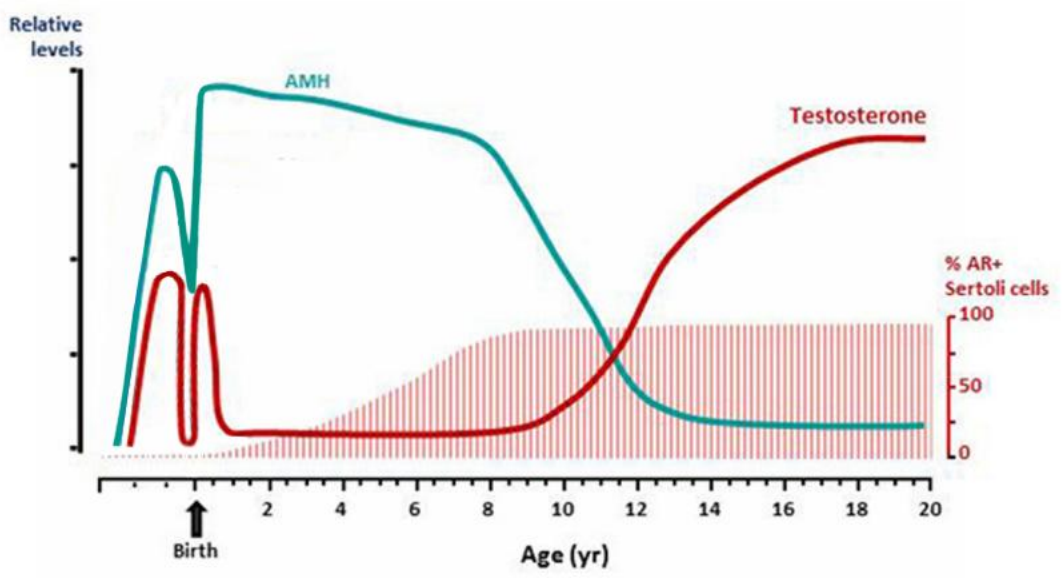


Figure XIV. Profiles of serum levels of testosterone and AMH in the male (Adapted from Rey et al., 2016).

Determination of serum AMH levels in male results indicative in evaluating intersex conditions such as ambiguous genitalia and testicular function in infants and children. In particular, measurable AMH concentration in young male with clinical evidence of bilateral cryptorchidism can be predictive of undescended testes; on the other hand, undetectable values are suggestive for anorchia or functional failure of the abnormally sited gonad. Thus, it can be helpful in presence of a suspected diagnosis of persistent Müllerian duct syndrome (PMDS). Moreover, in prepubertal patients with varicocele AMH levels are higher compared to the control suggesting for dysregulation in seminiferous epthitelian function. Finally, the measurement of AMH levels in basal or FSH-stimulated condition can be helpful to predict spermatogenic response to gonadotropic treatment in young patients with hypogonadotropic hypogonadism. According to these considerations, serum AMH concentration is considered a valid marker in assessing gonadal function, especially in presence of children pathological conditions, such as agonadal or abdominal testes, in which determination based on gonadotropin and testosterone serum levels fails (Josso et al., 2013; Matuszczak et al., 2013; Rey et al., 2016).

In female, the AMH expression begins in perinatal period with a subsequently rise, similar to the increment observed in male neonate. After that, it continues to slowly increase through childhood and adolescence, reaching a peak in the early twenties and, then, it declines progressively becoming extremely low after menopause (Figure XV). However, a wide variability in the AMH expression is observed among women, which reflects the differences in range of primordial and non-growing follicle number and in age at menopause.

Serum AMH levels are measured to evaluate functional state of gonads and related disorders including: (I) menopausal status and premature ovarian failure; (II) ovarian reserve and ovarian responsiveness for infertility and assisted reproduction; (III) ovarian function in patients with polycystic ovarian syndrome; (IV) monitoring of patients with AMH-secreting ovarian granulosa cell tumors. The most established clinical utility of serum AMH detection refers to the determination of ovarian status because AMH levels are correlated to the decline of oocyte/follicle pool with age and thus ovarian aging and menopausal transition. This also implies the applicability in assisted reproduction. The term “ovarian reserve” is used to indicate the number and quality of the remaining oocytes in the ovaries; instead, “ovarian aging” refers to age-related decline in the ovarian reserve.

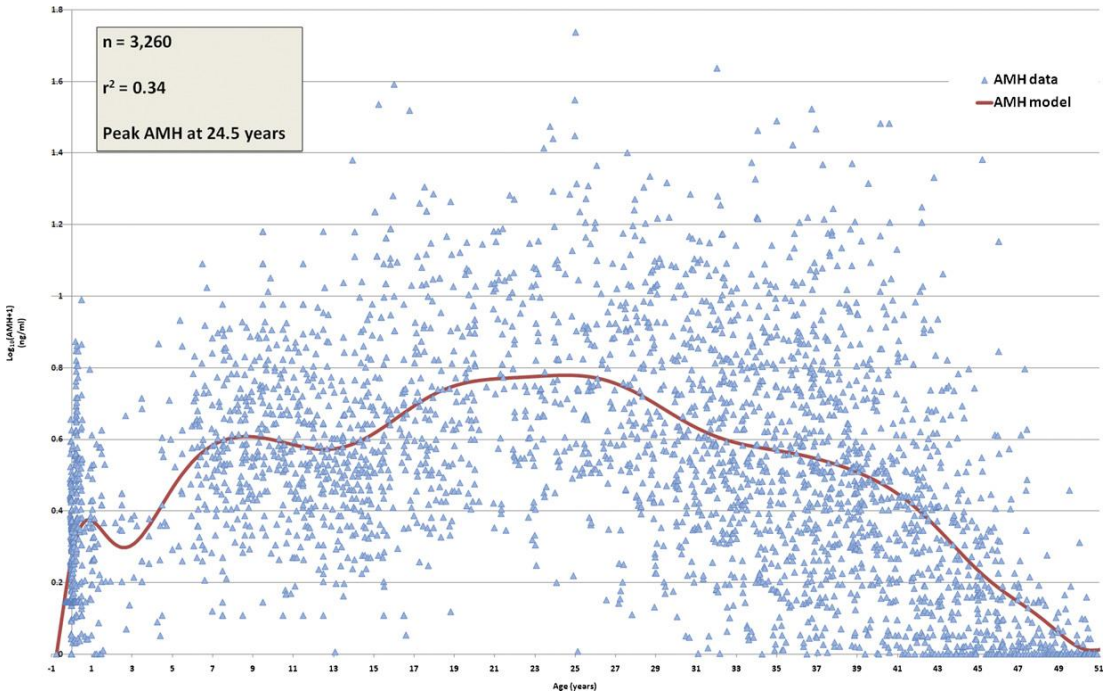


Figure XV. Serum AMH detection during female life: Serum AMH from conception to the menopause. The red line is the peak model that best fits the 3260 datapoints shown as triangles. The coefficient of determination, r2, is 0.34, indicating that 34% of variation in serum AMH levels is due to age alone. Peak serum AMH is at 24.5 years (Anderson et al., 2012).

Ovarian reserve can be correctly predicted by evaluation of the pool of resting primordial follicles in the ovary, but its measuring cannot be directly performed. Therefore, the evidence of a correlation between the size of resting primordial follicle and the number of follicles that enter the pool of growing follicles allowed to find an indirect method to measure ovarian reserve. Indeed, only the growing follicles produce AMH, thus serum AMH levels reflect the size of resting primordial follicle pool and, in turn, ovarian reserve. Comparing AMH levels and non-growing follicle (NGF) recruitment, a strongly correlation between AMH decrease and declining numbers of recruited NGFs after age 25 years was observed. This means that serum AMH levels can be a valid indirect indicator of human ovarian reserve for ages after the mid-twenties (Anderson et al., 2012; Feyereisen et al., 2006; Grynnerup et al., 2012; Visser and Themmen, 2005). Other hormonal markers (FSH, estradiol and inhibin B) and ultrasonographic markers (antral follicle count and measurement of ovarian volume) are also used to predict ovarian reserve by evaluating, directly or indirectly, the size of the antral follicle pool. However, it is reported that AMH is a more accurate measure of ovarian reserve than the other hormones because it shows little or no intracycle fluctuation (Figure XVI).

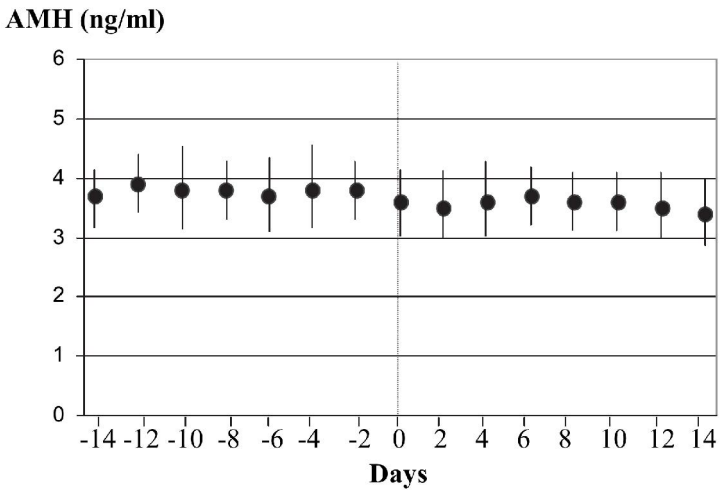


Figure XVI. AMH intracycle fluctuation: AMH levels measurement during menstrual cycle (La Marca et al., 2006).

In addition, its levels are unaffected by pregnancy or the use of contraceptives or other clinical and behavioral variables, such as body mass index and smoking status. On the contrary, inhibin B requires measurement in the early follicular phase of the menstrual cycle due to its variation and FSH is limited by condition that suppress the later FSH-dependent stages of follicle development (i.e. pregnancy, hormonal contraceptives, treatment with gonadotropin-releasing hormone agonist) (Anderson et al., 2012; Grynnerup et al., 2012; Hehenkamp et al., 2006; La Marca et al., 2006). Because of the increase of reliable data that correlates AMH levels with the number of oocytes obtained following superovulation, serum AMH level measurements have been routine practice in clinics for assisted reproduction for several years. Indeed, females with higher concentrations of AMH are demonstrated to respond better to ovarian stimulation and to produce more retrievable oocytes than female with low or undetectable AMH levels. However, women with high AMH levels are likely to incur in ovarian hyperstimulation syndrome (OHSS) due to the excessively response to exogenous gonadotropins. In contrast, women with low AMH levels are likely to poorly respond to stimulation. Serum AMH detection can allow clinicians to correctly assess the patients and to tailor the therapy, on one hand, minimizing the risk of develop OHSS and, on the other, evaluating the real chance of pregnancy.

In addition to ovarian status assessment, serum AMH concentration resulted an indicative marker to evaluate also ovarian function, in particular in polycystic ovarian syndrome (PCOS). In fact, AMH levels are reported to correlate well with the severity resistance of the syndrome to treatment. PCOS elevates serum AMH concentrations because it is associated with the presence of large numbers of small follicles. Furthermore, AMH is applied as early

and sensitive plasma marker of gonadal damage after treatment with chemotherapeutic drugs that usually implicate gonadotoxic side effects. In addition, since AMH concentration increments in granulose cell tumors, serum AMH levels resulted an extremely sensitive and specific marker in the follow-up of ovariectomized patients to determinate the efficacy of surgery and the disease recurrence. Several studies on a possible inhibitory role of AMH on epithelial ovarian cancer cells and on its use as adjuvant therapy for ovarian cancer are undertaken. Finally, in the same way of intersex conditions evaluation in male, in female with minimally virilized phenotype, AMH levels can be measured to evaluate pseudohermaphroditism and gonadal dysgenesis helping to distinguish gonadal and non-gonadal causes of virilization. Indeed, an AMH result above the normal female range is predictive of the presence of testicular tissue, while an undetectable value suggests its absence (Anderson et al., 2012; Dewailly et al., 2014; Dumont et al., 2015; Feyereisen et al., 2006; Grynnerup et al., 2012).

It goes without saying that AMH caught the attention to be the best biomarker for the assessment of ovarian reserve and aging and an attractive one for a wide range of physiological and pathological gonadal conditions in males and females, extending its applicability into different clinical fields. For these reasons, the development of AMH assay was strategically important. In 1990, Dr. Hudson and collaborators developed the first AMH immunoassay composed of two monoclonal antibodies, which recognized human recombinant AMH protein at the pro-region. After that, several AMH assays were developed aiming to augment the performances. In particular, the focus was on the high variability in AMH evaluation due to the storage and the freeze and thaw instability. In 2014, the first fully automated AMH immunoassay was developed by Dr. Gassner and Dr. Young highlighting the necessity to overpass the manual assays for a rapid access to AMH results. Indeed, the pressures of developing an automated immunoassay platform for the measurement into daily clinical practice was growing (Gassner and Jung, 2014; Hudson et al., 1990; Kumar et al., 2010; van Helden and Weiskirchen, 2015). By now, some automated assays are available on the market.

Aim of the work

Recombinant protein production is one of the most significant challenge for research and industrial applications. More and more resources are invested to develop the most efficient system for high protein expression. Mammalian cells are the best host for quality and reliability of expressed recombinant proteins. A diffuse protein production technology is based on dihydrofolate reductase (DHFR) deficient CHO cell line expression system, which has some limitations. Therefore, thanks to the continuous progression in recombinant DNA technology, DHFR knockout cell lines are recently developed using different molecular tools.

In order to achieve high-level of recombinant protein expression in stable cell line, we explored a new generated DHFR knockout cell line in association with *piggyBac* transposon system. Moreover, advances in mammalian cell knowledge on molecular mechanisms of protein expression pathway and epigenetics allowed to understand which elements could push towards high level of protein production also guaranteeing long-term expression stability. In this dissertation, several strategies based on transposon vector engineering were also implemented targeting various steps of protein expression pathway or acting on epigenetic regulations in order to further enhance protein yield and stability over time. To assess the potentiality of our expression platform we used anti-Müllerian hormone (AMH) as a model protein. AMH is a dimeric glycoprotein, member of the TGFβ superfamily that causes regression of the Müllerian duct in male embryos. Its role as diagnostic marker is important in assisted reproduction, where it is widely used to predict the ovarian reserve.

Materials and Methods

Constructs

New Englands Biolabs (NEB, Beverly, Massachussets) was the supplier for all enzymes used in this work.

pBlueScript (LifeTechnologies, Carlsbad, CA) was an empty vector used to balance DNA amount among transfections.

AMH

- *Standard vector*

Synthesis of AMH mutated forms sequences (Q450R and R451T) was commissioned to GeneArt (LifeTechnologies, Carlsbad, CA). NheI/XbaI enzymes were used to extracted AMH mutated sequences from pMK plasmids (GeneArt). Fragments were cloned in pcDNA3.1 plasmid (LifeTechnologies, Carlsbad, CA) digested with the same restriction enzymes. pcDNA3.1_AMH_Q450R and pcDNA3.1_AMH_R451T constructs were generated.

- *Transposon system*

The PB transposon vector and the transposase PB200 plasmid were purchased from System Biosciences (Mountain View, CA). The sequences of mutated AMH obtained from GeneArt were extracted by NheI/XbaI enzymes from the pMA plasmid (GeneArt) and cloned into transposon plasmid (System Biosciences, Mountain View, CA) digested with the same restriction enzymes to generate PB_AMH_Q450R and PB_AMH_R451T. Then, DHFR destabilized sequences were inserted downstream AMH sequence. They were synthetically obtained from GeneArt and extracted from pMA plasmid (GeneArt) by digestion with XhoI and SalI restriction enzymes. Each sequences was cloned into transposon vector containing AMH_R451T sequence.

We obtained:

PB_CAG_HSA_AMH_IRES_DHFR

PB_CAG_HSA_AMH_IRES_DHFR_PESTPEST

PB_CAG_HSA_AMH_IRES_deoptimizedDHFR

PB_CAG_HSA_AMH_IRES_DHFR_deoptimizedDHFR_PESTPEST

PB_CAG_HSA_AMH_FMDV IRES_DHFR

PB_CAG_HSA_AMH_FMDV IRES _DHFR_PESTPEST

PB_CAG_HSA_AMH_FMDV IRES _deoptimizedDHFR

PB_CAG_HSA_AMH_FMDV IRES _DHFR_ deoptimizedDHFR_PESTPEST

The pSF transposon vectors were purchased from Oxford Genetics Ltd, the transposase PB200 plasmid from System Biosciences (Mountain View, CA). The sequences of mutated AMH obtained from GeneArt were extracted by BamHI and XhoI enzymes from PB_AMH_R451T plasmid and cloned into transposon plasmids (Oxford Genetics Ltd) digested with the same restriction enzymes to generate:

pSF_cHS4_CMV_AMH_PGK_DHFR_PEST_ARE_cHS4

pSF_UCOE_CMV_AMH_PGK_DHFR_PEST_ARE

pSF_UCOE_CMV_AMH_PGK_DHFR_PEST_ARE_UCOE

pSF_MAR_CMV_AMH_PGK_DHFR_PEST_ARE

pSF_CMV_AMH_MAR_PGK_DHFR_PEST_ARE

Mammalian Cell Culture

- **CHO DHFR knockout cell line**

Disclosure agreement on cell line strictly avoids to declare its derivation. Cells were cultured in their basal medium supplemented with 1x Hybri-Max™ HT (Sigma-Aldrich, St. Loius, MO), 4mM L-glutamine and 100 UI/ml pen/strep (Lonza, Verviers, Belgium). Cells were grown at 37°C, 5% CO_2 in 125 ml flasks in orbital shaking incubator at 125 rpm. For cell line maintenance, every 3-4 days cells were splitted seeding 0.3 cell/ml in fresh medium. Viability and total cell count were assessed with trypan blue exclusion method and a haemocytometer.

Cryopreservation

Established cell lines or transfected pools were cryopreserved to create working cell banks or backup stocks. 1 x 10^7 cells were harvested and centrifuged at 220 x g for 5 minute. Cell pellets were resuspended in basal medium with the addition of 7% Hybri-Max™ DMSO (Sigma-Aldrich, St. Loius, MO) in 1.5 ml polypropylene vials (Nalgene). Vials were slowly frozen at -80°C and then transferred for long-term storage in a liquid nitrogen tank (at -180°C).

Thawing procedure

A vial of cells removed from the liquid nitrogen storage vessel was quickly thawed at 37°C. Cells were then transferred into 5 ml of basal medium in a 15 ml centrifuge tube and

centrifuged at 220 x g for 5 minute. The supernatant was removed and the cell pellet resuspended in 10 ml of fresh basal medium and dispensed in a 125 ml flask.

Transfection

Cell line was stably transfected applying Amaxa Nucleofector apparatus (Lonza Group, Basel, Switzerland) and the specific nucleofection protocol for CHO cells in suspension. The day before nucleofection, cells were subcultured 1:2 in fresh medium. The day of transfection cells were collected and counted in order to prepare a cell pellet containing 2 x 10^6 cells by centrifugation at 220 x g for 5 minutes. The pellet was resuspended in 100 µl of previously prepared Nucleofector solution (solution V). Then, the cell suspension was mixed with 4.2 µg of total DNA (DNA was prepared with a transposon:transposase ratio of 2.5:1, thus 3 µg of transposon and 1.2 µg of transposase were added) and transferred into a proper cuvette. Once inserted in the Nucleofector instrument, the U-032 program was applied. Subsequently, cell suspension was diluted with 500 µl of basal medium and gently transferred into a 6-well plate containing 2 ml of supplemented basal medium. Transfected cells were incubated at 37°C in orbital shaking incubator at 125 rpm. 48 hours post transfection cells were diluted with basal medium without HT, but supplemented with 4mM L-glutamine, 100 UI/ml pen/strep (Lonza, Verviers, Belgium), 1500 µg/ml Pluronic F-68 and 200 µg/ml Dextran sulfate (Sigma-Aldrich, St. Loius, MO), and transferred into a 125 ml flask. Selective medium was replaced every 3-4 days during routine maintenance. Cell density for inoculation was 0.3 x 10^6 cells/ml.

Gene amplification by MTX

- *Methotrexate preparation*

100 mg of methotrexate powder (Sigma Aldrich) were resuspended in 204 mL H_2O adding 1ml NaOH 1M. The solution was titrated to pH 7, filtered through 0.2 µm filter and stored in 2 ml aliquots at -20°C.

Molecule stability was assessed by mass spectrometry with the following specifications:

- Column Acquity BEH C18 (Waters, MA) 130Å 1.7 µm 2.1x100 mm, 45°C
- Eluent: A) H_2O + 0.02% TFA
  B) Acetonitrile + 0.02% TFA
- Flow rate: 0.4 mL/min
- Gradient: 1 min isocratic from 10%B, 4 min from 10% to 40%B
- Volume injection 2µl
- Sample diluted 1:5 in H_2O + 0.1% TFA

- *Single step protocol*

Cells from stable pool were harvested and counted by trypan blue exclusion and haemocytometer. Cells with a density of 0.3 x 10^6 cells/ml were transferred in a 15 ml falcon and then were centrifuged at 220 x g for 5 minutes. Cell pellets were resuspended in complete selective basal medium containing the appropriate MTX concentration and transferred into a 125 ml flask. Cell culture was incubated at 37°C in orbital shaking incubator at 125 rpm. Every 3-4 days, for in total 2 weeks, a complete media change was performed on the culture by centrifuging the cells, eliminating the exhausted medium and resuspending in fresh complete selective medium with the appropriate MTX concentration. Cells were seeded with a density of 0.3 x 10^6 cells/ml at each passage.

- *Multiple step protocol*

Cells from stable pool were harvested and counted by trypan blue exclusion and haemocitometer. Cells with a density of 0.3 x 10^6 cells/ml were transferred in a 15 ml falcon and then were centrifuged at 220 x g for 5 minutes. Cell pellets were resuspended in complete selective basal medium containing the lowest MTX concentration (i.e. 50 nM) and transferred into a 125 ml flask. Cell culture was incubated at 37°C in orbital shaking incubator at 125 rpm. Every 7 days (2 passages), a complete media change on the culture was performed progressively increasing MTX concentration (i.e. 100 nM, 250 nM, 500 nM, 1 µM, 2 µM MTX concentration). Cells were seeded with a density of 0.3 x 10^6 cells/ml at each passage. The treatment lasted about 6 weeks.

Cell pools evaluation

At each passage during MTX treatment on cell pools, cell number and viability were evaluated by trypan blue exclusion and haemocitometer. At the end of MTX treatment, supernatant from each pool of cells was harvested after 4 days of cell culture for AMH quantification. AMH protein was quantified by AMH Gen II ELISA immunoassay (Beckman Coulter Inc, Brea, CA). Specific Productivity Rate (SPR) was calculated using the following formula:

$$SPR = \frac{\text{ng of protein}}{10^6 \text{ cells x incubation days}}$$

Cloning step

In order to obtain stable single clone, once stopped MTX treatment (6 weeks post-transfection), cells were seeded at 0.5 cell/well in 96-well plates with Excell CHO Cloning Media (Sigma Aldrich) and 1:40 Clonacell-CHO ACF supplement (StemCell technologies, Vancouver, CN). Single clones were identified and maintained in culture for screening to isolate high-producing ones.

Clones screening

Supernatants obtained from clones after 4 days of culture were analyzed with an automated prototype immunoassay for AMH detection in serum and plasma samples developed on Liaison platform (Diasorin, Saluggia, Italy). Liaison Platform is an automated instrument that performs chemiluminescent immunometric analysis of biological matrices (serum, plasma and feces). The chemiluminescent molecule used in Liaison immunoassay is ABEI (N-4(-Amino-Butyil)-N-Ethyl-Isoluminol), a derivative of luminal, that is conjugated with the antibody or antigen used as tracer. The light signal is measured by a photomultipler as relative light units (RLU). Data obtained were normalized on days of culture and cell number. Cell number was evaluated with Dojindo Cell Counting kit-8 assay (Dojindo Molecular Technologies, Rockville, MD). Final value represents the specific productivity rate (SPR) calculated with the formula described above.

Clone productivity and stability over time

- *AMH quantification*

AMH protein quantification was performed on the supernatant obtained from 4 days cell culture with the automated prototype immunoassay on Liaison platform.

- *DNA extraction*

Candidate clones were seeded at a density of $0.3x10^6$ cells/ml in basal medium. After 4 days of culture, cell pellets of 3×10^6 cells were collected. DNA was extracted by using DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) in accordance with manufacturer's instruction.

- *qPCR analysis*

Quantitative PCR (qPCR) analysis was performed using iTaq SYBR Green Supermix vit Rox on a Dyad Disciple Peltier Thermal Cycler (Bio-Rad). All reactions were performed in a total volume of 20 µl containing 12 ng DNA. Amplification cycles consisted of an initial

denaturing cycle at 95°C for 10 minute, followed by 45 cycles of 15 seconds at 95°C and 30 seconds at 62°C. Fluorescence was quantified during the 62°C annealing step, and specificity of the amplification product was confirmed by examination of dissociation plots (55-94°C). A distinct single peak indicated that a single DNA sequence was amplified during PCR. Each sample was tested in duplicate and samples obtained from the three independent experiments were used for the absolute quantification analysis as described below. A ten-fold serial dilution series of the plasmid carrying the target gene, ranging from 3.8×10^5 to 3.8 copies/µl, was used to construct the standard curves. The corresponding copy number was calculated using the following equation (Fu et al., 2009):

$$\text{plasmid copy number} = \frac{6.02 \times 10^{23}\,(\text{copy/mol}) \times \text{DNA amount (g)}}{\text{DNA length (dp)} \times 660\,(\text{g/mol/dp})}$$

Ct values in each dilution were measured in duplicate to generate a standard curve. The Ct values were plotted against the logarithm of their initial template copy numbers. Each standard curve was generated by a linear regression of the plotted points. From the slope of each standard curve, PCR amplification efficiency (E) was calculated according to the equation (Fu et al., 2009):

$$E(\%) = (10^{-1/\text{slope}} - 1) \times 100\%$$

Amounts of target gene were normalized to β2 microglobulin gene (β2m), a reference gene. Two copies of β2m gene are present for genome. Primer sequences and their concentrations in the reaction are listed below (Scheme I). DHFR primer pair was used for transgene copy number evaluation in clones derived from standard vector and PB transposon. AMH primer pair was used for those deriving from psF and pSF_U transposon.

Gene	Forward (10 µM)	Reverse (10 µM)	Amplicon
β2m	AGCTGTCAGATCTGTCCTTCAAC	CTCCTTCAGAGTGGTGTGTGTAA	119 bp
DHFR	ACACGATGATAATATGGCCACA	GTACTTGAACTCGTTCCTGAG	136 bp
AMH	CGAAGTGACCGTGACCAG	TGGAGTCCTCGCCTCTAG	155 bp

Scheme I. Primer pairs: forward and reverse primers used for transgene copy number evaluation. Length of primers and concentration of use are also reported.

Adaptation to other media

Selected clones were adapted to one of the following production media: Cellvento™ CHO-100 Liquid (Merck Millipore, Burlington, MA) and ActiPro (GE Healthcare, Chicago, IL). Media were supplemented with 4 mM L-glutamine (Lonza, Verviers, Belgium), 100 UI/ml pen/strep (PS; Lonza, Verviers, Belgium), non essential amino acids MEM 1X (MEM; ThermoScientific, Waltham, MA) and 200 µg/ml dextran sulfate (DS; Sigma-Aldrich, St. Loius, MO) as described in Scheme II.

CV	CV+DS	CV+MEM	CV+DS+MEM	AP
Glut	Glut	Glut	Glut	Glut
PS	PS	PS	PS	PS
	DS	MEM	DS	
			MEM	

Scheme II. Media compositions: AP: ActiPro; CV: CellVento; DS: dextran sulfate; MEM: non essential amino acids; Glut: L-glutamine; PS: pen/strep.

Clones adaptation was progressive starting from 30% of production medium to reach in 3 steps (30%, 60%, 100%) the complete medium change. Each adaptation step was performed every two passages (Scheme III).

Passages	Basal medium	Production medium
-	100%	0%
2	70%	30%
2	40%	60%
2	0%	100%

Scheme III. New media adaptation protocol

Cells were subcultured in the new medium every 3-4 days, until their doubling time was shorter than 35 hours and they were considered adapted (Nielsen et al., 2010).

Duplication time (DT) was calculated using the following formula:

$$DT = \frac{\ln(2)}{\mu}$$, where µ is specific growth rate, calculated with $$\mu = \frac{\ln x_2 - \ln x_1}{t_2 - t_1}$$

where x_1 and x_2 are viable cell concentrations at time points t_1 and t_2.

Scale-up in mild hypothermia condition

Scale-up procedure was performed in 1-liter flask with a cell density of $0.4x10^6$ cells/ml at 125 rpm orbital shaking in a total volume of 100 ml. Temperature was shifted to 32°C the day after the inoculums. Cell number, viability and glucose measurement were evaluated at each time point. Specific growth rate (μ), specific protein productivity (q) and duplication time were calculated with the formula described below. Protein expression was quantified on supernatants collected with the prototype immunoassay on Liaison platform.

- *Specific growth rate (μ)* was calculated from a semi-log plot of viable cell concentration versus culture time in the exponential growth phase:

$$\mu = \frac{\ln x_2 - \ln x_1}{t_2 - t_1}$$

where x_1 and x_2 are viable cell concentrations at time points t_1 and t_2.

- *Specific protein productivity (q)* represents protein productivity measured in pg/cell/day with the following formula:

$$q = \frac{m \cdot \ln(N/N_0)}{(N - N_0) \cdot t}$$

where m is the total protein mass in the culture supernatant, N_0 and N are initial and final viable cell numbers, respectively and t represents days in culture.

- *Duplication time (DT):* duplication time (DT) was calculated using the following formula:

$$DT = \frac{\ln(2)}{\mu}$$

- *Glucose measurement*: glucose concentration in cell supernatant was measured using Accu-Chek Aviva Plus glucose meter (Roche Diagnostic, Indianapolis, IN). A drop of collected cell supernatant was used for the measurement with a disposable Accu-Chek strip (Roche Diagnostic, Indianapolis, IN), in accordance with manufacturer's instructions.

Protein Gel Electrophoresis

Polyacrylamide gel electrophoresis (PAGE) with denaturant sodium dodecyl sulfate (SDS) was used to separate proteins by size from cell culture supernatants. The running buffer Tris/Glycin (Bio-Rad, Hercules, CA) was diluted 1:10 and used to run the gel electrophoresis. 12.5 µl of clarified supernatant samples were combined with 12.5 µl of sample buffer (reducing or non reducing) followed by 5 minutes at 99°C. The samples were centrifuged briefly and were then loaded into Criterion™ TGX Stain-Free™ Precast 10% polyacrylamide

gels (Bio-Rad, Hercules, CA). Precision Plus molecular weight markers (Bio-Rad, Hercules, CA) were also included. The peculiarities of Criterion™ gels are fast run times (30-40 minutes at 240 V) and the possibility to visualize protein with a ChemiDoc imagers (Bio-Rad, Hercules, CA) without the comassie staining. This ability is due to the trihalo compounds incorporated into the gel matrix that react with tryptophan residues in a UV-induced reaction to produce fluorescence, which can be easily detected by the ChemiDoc imagers. The reaction occurs within gels or on a nitrocellulose membrane after transfer. In our case, UV light was applied for 5 minutes to obtain best sensitivity.

Western Blotting

Proteins were transferred from the gel to a nitrocellulose membrane employing the Trans-Blot Turbo Transfer System (Bio-Rad, Hercules, CA). This system provides a rapid blot transfer thanks to Transfer Pack that includes an optimized buffer, a nitrocellulose membrane and filter paper combination. The process is a semi-dry transfer and occur in 7 minutes on the Trans Blot Turbo blotting apparatus (Bio-Rad, Hercules, CA).

- *AMH immunostaining*: the membrane was blocked in 0.5% BSA in PBS buffer at room temperature for 2 hours. After blocking, blot was immunostained with a commercial anti-C terminal AMH antibody (Cloud-Clone Corp, Huston, TX) (1:200). Post incubation, blot was washed three times with PBS-Tween20 0.1% and probed with goat anti-mouse IgG (Fc) conjugated with horseradish peroxidase (1:5000) (Thermo Fischer Scientific) for 2 hours at room temperature on a rocking platform. Then, the membrane was washed three times with PBS-Tween20 0.1% before the blots were subjected to detection using Clarity Western ECL Substrate (Bio-Rad). Clarity Western ECL is a substrate for HRP-conjugate secondary antibody composed by two components. The two components were mixed in a 1:1 ratio preparing a 0.1 ml of solution/cm^2 of membrane. Then, the mix was added to the membrane and incubated for 2 minutes. Finally, the membrane was imaged by Chemidoc instruments (Bio-Rad) using ImageLab Software (Bio-Rad).

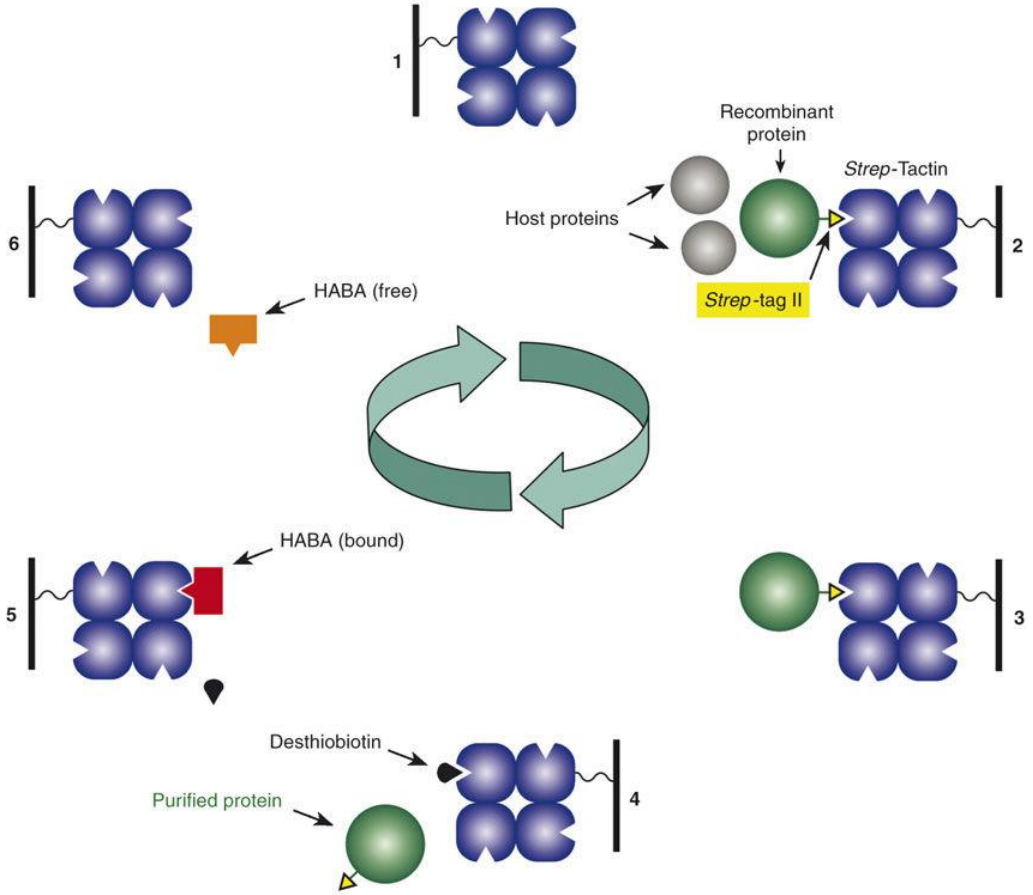
Purification

Purification process based on Strep-tag™ II

This purification method is based on the high-affinity binding between biotin and streptavidin. Strep-tag™ II is a small tag of only eight amino acid residues and a molecular weight of just 1000 Da. The small size of the tag is beneficial, since in most cases it does not interfere with structural and functional studies, and therefore does not have to be removed.

Strep-tag™ II binds specifically to StrepTactin™ ligand, an engineered analogous of streptavidin, immobilized on the base matrix to yield pure target protein. The binding affinity of Strep-tag™ to the immobilized ligand is nearly 100-fold greater than to streptavidin, making this method a very efficient way of purifying proteins. Purification process consists of the following step (Scheme IV):

1. StrepTactin™ ligand is immobilized on the base matrix of the purification column.
2. Supernatant is loaded on StrepTactin™ column. Target protein containing Strep-tag™ II binds the resin with high selectivity.
3. Contaminant proteins in supernatant, which do not contain Strep-tag™ II, do not bind the resin and are removed during washing step (wash buffer A).
4. Target protein is then eluted during elution step (buffer B), which corresponds to wash buffer with the addition of 2.5-3 mM desthiobiotin. This molecule competes with target protein for the specific binding sites on StrepTactin™ and displaces Strep-tag™ proteins at the biotin-binding site of StrepTactin™.
5. In order to regenerate the resin, a wash buffer with the addition of 1 mM di HABA (2- [4'-hydroxy-benzeneazo] benzoic acid) is used. HABA displaces desthiobiotin at the biotin-binding site of StrepTactin™. Furthermore, HABA is yellow colored. This allows to control resin regeneration. Indeed, when HABA binds biotin-binding site of StrepTactin™, the resin changes color to red.
6. HABA is removed by flowing in column the wash buffer. Once red color disappears, the column can be used again.



Scheme IV. Purification process StrepTactin/Strep-tag

Purification protocol using Strep-tag was based on two steps:

Step one. *Tangential Flow Filtration (TFF): protein concentration and buffer exchange.*

Supernatant was thawed and filtrated on 0.22 µm Millipore Express PLUS Membrane (Merck-Millipore) overnight at 4°C. Then, 4 cycles on TFF with Pellicon XL Biomax PES cassette (molecular weight cut off di 30 kDa; Merck-Millipore) were performed.

Ultrafiltration process protocol:

- sample was concentrated from 150 ml to 50 ml, then where added 50 ml of *Buffer A StrepTactin™*: Hepes 100 mM, NaCl 150 mM, EDTA 1 mM, pH 8.0;
- 1° diafiltration cycle till 50 ml and addition of 50 ml of *Buffer A StrepTactin™*.
- 2° diafiltration cycle till 50 ml and addition of 50 ml of *Buffer A StrepTactin™*.
- 3° diafiltration cycle till 50 ml with complete emptying of cassettes. Retentate obtained was moved to step 2.

Step one was entirely checked with SDS-PAGE and Western Blot analysis. Moreover, filtration step was monitored with our automated immunoassay prototype.

Step two. *Protein purification with affinity chromatography: StrepTactin/Strep-tag*

Samples obtained from step one was loaded on pre-charged StrepTactin resin (StrepTactin Sepharose 50% suspension - IBA GmbH) for a total range of resin volume from 5 ml to 10 ml. Charged column diameter was 26mm and height was 10-20 mm in average (column volume = 5.3 ml – h=10 mm - or 10.6 ml – h=20 mm). Column was equilibrated with 5 volume of *Buffer A StrepTactin™*. Sample loading (final volume 50 ml) was performed with a low speed (0.055 ml/min) in order to increase contact time between sample and resin (about 15.5 hours – overnight) and, thus, the efficiency of bond between AMH protein and resin. Then, column was washed (10 volume in column). Elution was performed in one step with 100% of *Buffer B StrepTactin™* (Hepes 100 mM, NaCl 150 mM, EDTA 1 mM, Desthiobiotin 3 mM, pH 8.0) in 5 column volume. Wash step was characterized by a flow of 1-2 ml/min, elution of 1 ml/min; absorbance was monitored at 260 and 280 nm.

StrepTactin resin interacted with high affinity to the Strep-tag present on N-terminal portion of AMH protein. The protein was eluted with a step of desthiobiotin because this molecule competes with target protein to bind StrepTactin resin.

Fractions containing target protein were identified combining SDS-PAGE analysis and chromatography profile. Fractions were mixed to obtain final preparation. Protein concentration was evaluated with the automated immunoassay prototype on Liaison platform. Purified protein was then stored at -30°C.

<u>Purification process based on His-tag</u>

Purification protocol using His-tag was based on two steps:

Step one. *Tangential Flow Filtration (TFF): protein concentration and buffer exchange.*

Supernatant was thawed and filtrated on 0.22 µm Millipore Express PLUS Membrane (Merck-Millipore) overnight at 4°C. Then, 4 cycles on TFF with Pellicon XL Biomax PES cassette (molecular weight cut off di 30 kDa; Merck-Millipore) were performed.

Ultrafiltration process protocol:

- sample was concentrated from 150 ml to 50 ml, then where added 50 ml of *Buffer A IMAC*: Na2HPO4 50 mM, NaCl 100 mM, pH 8.0.
- 1° diafiltration cycle till 50 ml and addition of 50 ml of *Buffer A IMAC*.
- 2° diafiltration cycle till 50 ml and addition of 50 ml of *Buffer A IMAC*.
- 3° diafiltration cycle till 50 ml with complete emptying of cassettes. Retentate obtained was moved to step 2.

Step one was entirely checked with SDS-PAGE and Western Blot analysis. Moreover, filtration steps were monitored with our automated immunoassay prototype.

Step two. *Protein purification with Immobilized Metal Affinity Chromatography (IMAC)*:
Samples obtained from step one was loaded on two HisTrap FF crude (GE Healthcare) IMAC column connected in series for a total resin volume of 10 ml. Columns were activated with 3CVs (30 ml) of 0.1 M nickel chloride solution and, then, washed with 3CVs (30 ml) of MilliQ water. Finally, columns were equilibrated with *Buffer A IMAC.* The loading was performed with a speed of 1 ml/min. After that, a wash step was performed (10 column volume). A gradient elution was performed starting from 0 to 100% of *Buffer B IMAC* (Na2HPO4 50 mM, NaCl 100 mM, 500 mM Imidazole, pH 8.0) in 7 column volume (70 ml). Wash steps were characterized by a flow of 2 ml/min; absorbance was monitored at 260 and 280 nm. Fractions containing target protein were identified combining SDS-PAGE analysis and chromatography profile. Fractions were mixed to obtain final preparation. Protein concentration was evaluated with the automated immunoassay prototype on Liaison platform. Purified protein was then stored at -30°C.

Results

Part 1: Transgene optimization

Large-scale expression of heterologous proteins with high yield is a process as much fascinating as challenging. One of the chances to achieve a protein of good quality and at high expression level is performing an appropriate vector design. This, in turn, implies, for example, the choice of a strong promoter, a proper signal peptide and a correct gene codon optimization. Therefore, it is necessary to investigate which elements allow an optimal protein expression. A transient transfection protocol is ideal for this initial explorative purpose in order to evaluate transgene expression and to characterize the protein quality.

1.1 Secreted AMH protein

In the first step of our project, we applied transient transfection in order to optimize the transgene expression of a model protein, the anti-Müllerian hormone (AMH). Anti-Müllerian hormone (AMH) is a 140-kDa disulfide-linked homodimer glycoprotein member of the TGFβ superfamily. Physiologically, a cleavage occurs at a kex-like ^{448}RXXR/S^{452} site and generates 110 kDa N-terminal and 25 kDa C-terminal homodimers, which remain associated in a non-covalent complex. We decided to set up sequences for two AMH protein mutated forms. Mutations were created in the kex-like site. Q450R mutation generates a more cleavable AMH form, while R451T a non-cleavable one, as reported in literature (di Clemente et al., 2010; Nachtigal and Ingraham, 1996; Papakostas et al., 2010; Pépin et al., 2013). Our aim was to evaluate the effect of those mutations on AMH protein and to characterize mutated proteins for immunochemical purposes. The polypeptide structure was designed with two tag sequences (His-tag and Strep-tag) and a Tobacco Etch Virus (TEV) protease site at the N-terminal (Figure 1A). His-tag was inserted to be sure to have an element suitable for the purification process by Immobilized Metal Affinity Chromatography (IMAC). Strep-tag was added in order to test a new purification method and to verify if it can be more efficient than IMAC purification. Moreover, we decided to insert a TEV site in order to perform a cleavage and remove the tag sequence from AMH protein in case of its application as immunogen. FreeStyle CHO-S cells (CHO-FS), a cell line adapted to grow at high density in suspension, under serum-free condition, was used for this purpose. Transient transfection protocol was performed with plasmids encoding for the AMH protein mutated forms. Culture supernatants were collected at different time points.

Western Blot (WB) analysis and commercial Enzyme-Linked Immunosorbent Assay (ELISA) were performed to analyze recombinant proteins. WB analysis of the culture supernatants confirmed AMH protein production and R451T mutation cleavage prevention (Figure 1B).

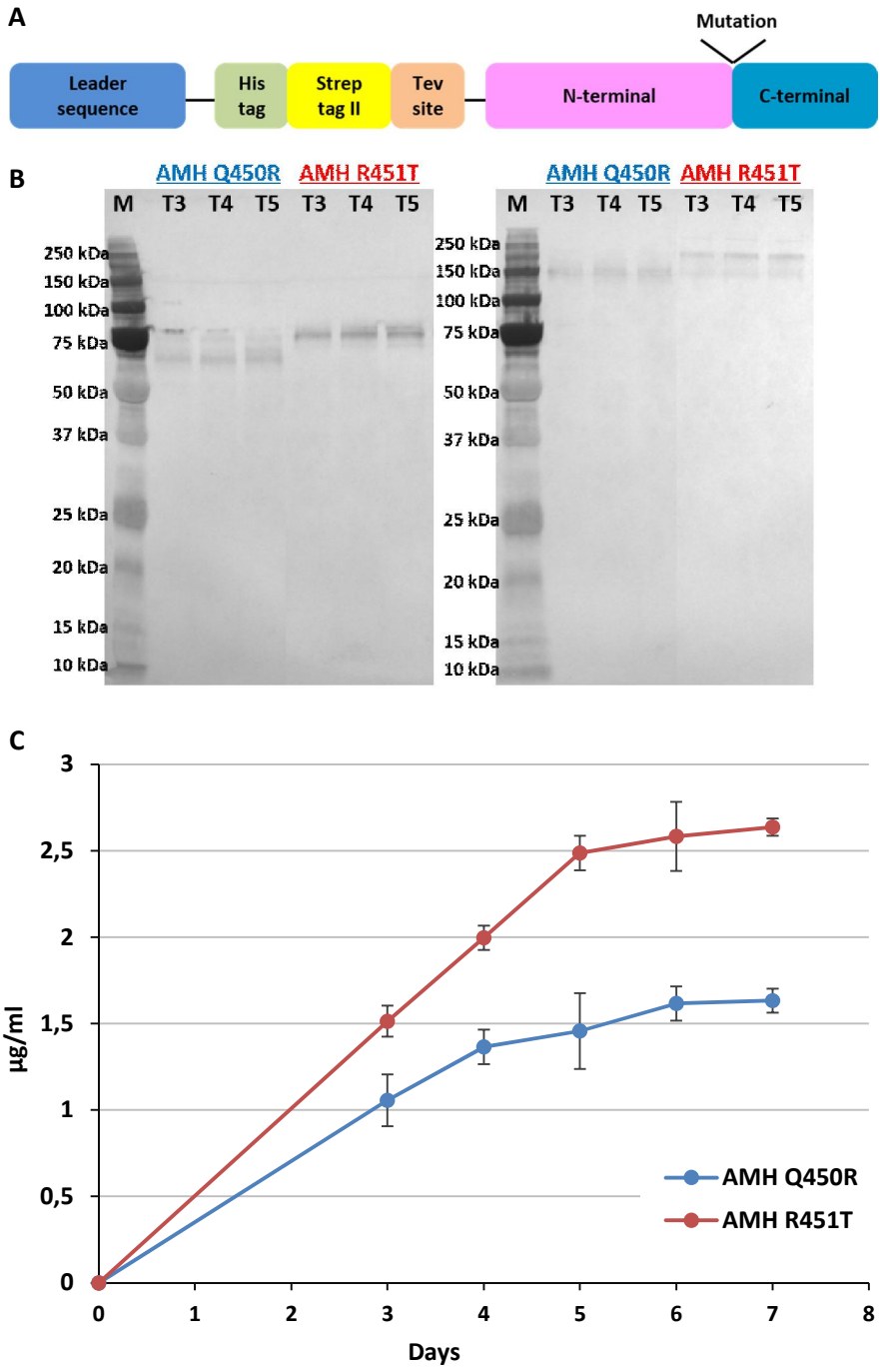


Figure 1. AMH structure: A. Polypeptide structure. B. WB analysis with His probe immunostaining. On the left, WB analysis in reducing conditions. On the right, WB analysis in non reducing conditions. AMH Q450R is the cleavable form; AMH R451T is the non-cleavable form. C. ELISA commercial assay analysis on supernatants collected at different time points.

As written before, the wild type AMH protein is a dimer. The monomer has a molecular weight of about 70 kDa. It is cleaved at amino acid 451 in N- and C-terminal fragments. They have respectively a molecular weight of 57 and 12.5 kDa. Therefore, it is expected that the homodimer of full-length AMH has a molecular weight of 140 kDa, while the N-terminal dimer of 110 kDa and the C-terminal dimer of 25 kDa. Thus, in WB analysis, we observed a band at about 60 kDa coincident with the N-terminal fragment of Q450R mutated form in reducing condition; the C-terminal fragment was not detectable because WB was developed using His-probe and the His-tag is located at the N-terminal. For the R451T mutated form, we were able to detect a band corresponding to the protein monomer at 75 kDa. In non-reducing condition, we observed a band at about 150 kDa corresponding to the AMH N-terminal dimers of cleavable form and a band at about 170 KDa coincident with the homodimers of the non-cleavable one. The molecular weights observed were higher than those we expected and this was probably due to several protein glycosylations. Furthermore, both AMH protein forms were recognized and quantified by a commercial ELISA assay. In particular, supernatant samples were collected each day of a 7 days cell culture for both AMH protein forms (starting from day 3) and were quantified (Figure 1C). Results showed that the non-cleavable mutated form of AMH protein was highly expressed compared to the cleavable one. Reaching a high expression of protein means simplifying the downstream process. This consideration, in addition to the fact that the biological activity of the protein was unnecessary for our immunochemical purposes, allowed us to select the non-cleavable mutated form of AMH protein for further analysis.

1.2 Signal peptide

The recombinant protein production is a complicated and very expensive process. Several efforts are employed to simplify the use and minimize the costs of the entire procedure. In this respect, a great goal would be producing as much of the protein of interest as possible. One way to achieve this is to focus on an efficient protein secretion into the medium for convenient collection and purification. Thus, in order to explore this aspect, the substitution of the endogenous AMH leader sequence with Interleukin-2 (IL2) and human serum albumin (HSA) leader sequences was performed (Pépin et al., 2013; Zhang et al., 2010). We transiently transfected CHO-FS cells with a plasmid containing IL2 or HSA leader sequences fused to the AMH coding sequence. Transfected cells were maintained in culture

for 5 days. At day 5, supernatants were collected and AMH protein was quantified by a commercial ELISA assay. The comparison of the signal peptides in both transfections resulted in the higher performance of HSA leader sequence since it enhanced the AMH secretion (Figure 2).

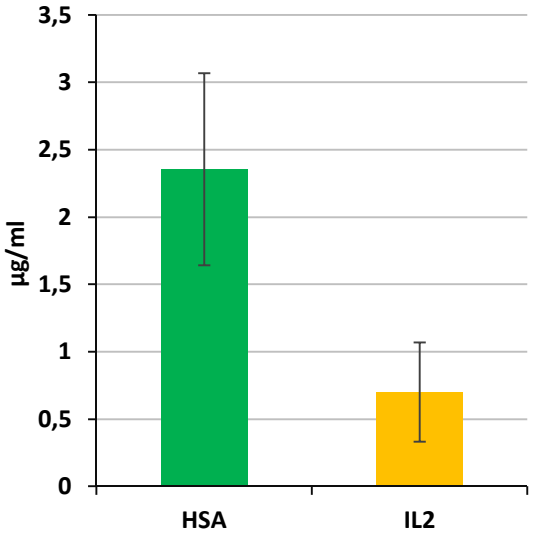


Figure 2. IL2 and HSA leader sequences comparison: CHO-FS cells were transiently transfected with PB_AMH constructs with Interleukin-2 (IL2) or with human serum albumin (HSA) leader sequences and maintained in 125 ml cultures flasks. Supernatants were harvested at day 5 from transfection and AMH protein was quantified by a commercial ELISA assay. Two transient transfections were performed.

Part 2: PB transposon system and CHO DHFR knockout cell line

2.1 Standard vector versus PB transposon system: preliminary data

The productivity of clones isolated from standard vector and *piggyBac* transposon (PB transposon) transfections was evaluated in preliminary experiments.

Briefly, PB transposon, expressing a model protein, was transfected into CHO-K1 adherent cells using a transposon/transposase molar ratio of 2.5:1 (Figure 3A). As a comparative experiment, cells were also transfected with standard vector (Figure 3B).

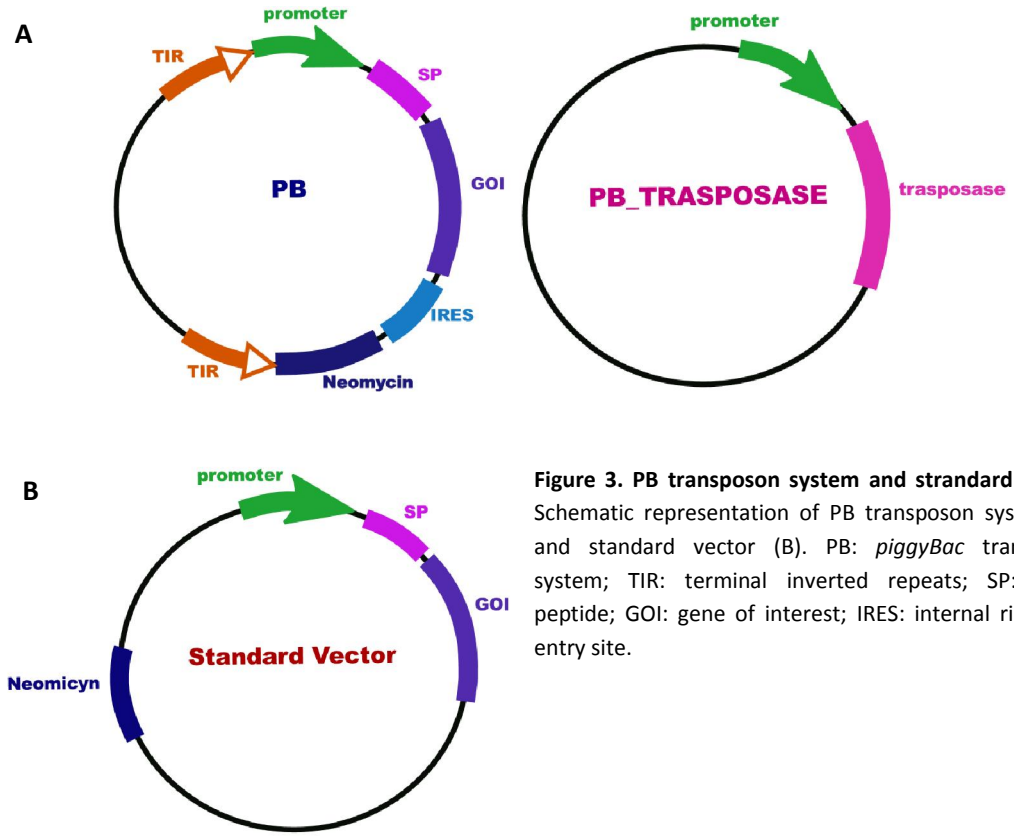


Figure 3. PB transposon system and strandard vector: Schematic representation of PB transposon system (A) and standard vector (B). PB: *piggyBac* transposon system; TIR: terminal inverted repeats; SP: signal peptide; GOI: gene of interest; IRES: internal ribosome entry site.

One hundred of single clones were isolated from each condition and maintained in culture for 2 weeks under selection in order to ensure a stable expression. A quantitative screening was then performed on supernatants obtained from 2 days cell culture. Collected samples were analyzed by an automated prototype assay for the detection of the model protein. Quantitative values were normalized to both clone cell number and days of culture. The resulted value represented the production rate, or Specific Productivity Rate (SPR), expressed as nanograms (ng) of protein produced by 10^6 cells per day (ng/10^6 cells/day).

Production categories were arbitrarily defined in which data obtained were included (0-80, 80-500, 500-1000, and >1000 ng/10^6 cells/day). Final data were expressed as percentage of clones for each category in which clones were divided (Figure 4).

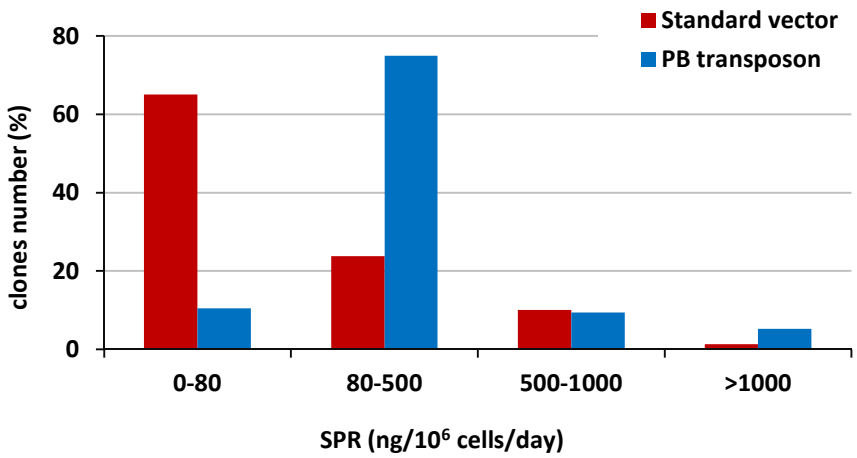


Figure 4. Productivity of clones: Quantitative screening of clones generated with standard plasmid transfection or PB transposon system. Clones were classified based on a model protein expression level. Cell supernatants were analyzed with an automated prototype assay for the detection of the model protein. Values were normalized based on clone cell number (Dojindo cell counting assay) and days of culture. Data were reported as percentage of clones classified into arbitrary categories of production. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

Clones derived from standard vector resulted in a distribution shifted towards low-producing clone categories with the peak in the lowest producing one. On the contrary, clones derived from PB transposon were distributed in the medium-high producing categories, mainly populating the medium one. These preliminary data showed that by using the PB transposon system it is possible to enhance the frequency of high producing clones compared to conventional transfections. In this way, the probability to isolate the highest-producing clone by using PB transposon can be greatly increased.

2.2 Gene amplification process

A widespread protein production technology is the dihydrofolate reductase (DHFR) deficient CHO cell line expression system. Different mutant strains of DHFR deficient CHO cell line are employed in industry, which are characterized by mutation or deletion in one or both alleles of the DHFR gene. DHFR deficient cell lines currently used are DuxB11 and DG44. Thanks to the continuous progression in engineered mammalian cells, DHFR knockout cell lines were recently developed by using different molecular tools. There is no

literature until now about the new engineered cell line used in this work, therefore its performances are unknown. Based on the encouraging data on the PB transposon system, described in the previous paragraph, we decided to evaluate this powerful tool of transgene integration in combination with the new platform DHFR knockout cell line. The hypothesis is that the genetic characteristic of this cell line could guarantee an efficient selective pressure and an extensive amplification process. If this assumption is true, exploiting the PB transposon potentiality in this cell line could mean improving the results even more. In order to explore this aspect and to analyze the effects of amplification process on the integrated transgene, our research plan involved the comparison of the conventional gene transfer method and the PB transposon system. For the PB transposon system, we co-transfected DHFR knockout cells with the plasmid carrying the expression cassette and the helper plasmid encoding for the transposase using a transoson/transposase molar ratio of 2.5:1, as established in preliminary data. In order to operate a correct comparison, we performed a co-transfection also for standard vector. We transfected cells with the plasmid containing the expression cassette and a "mock" plasmid to mime the same transfection condition of the PB transposon system. The expression cassette for both systems consisted of AMH gene followed by DHFR gene linked through an encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) sequence (Figure 5).

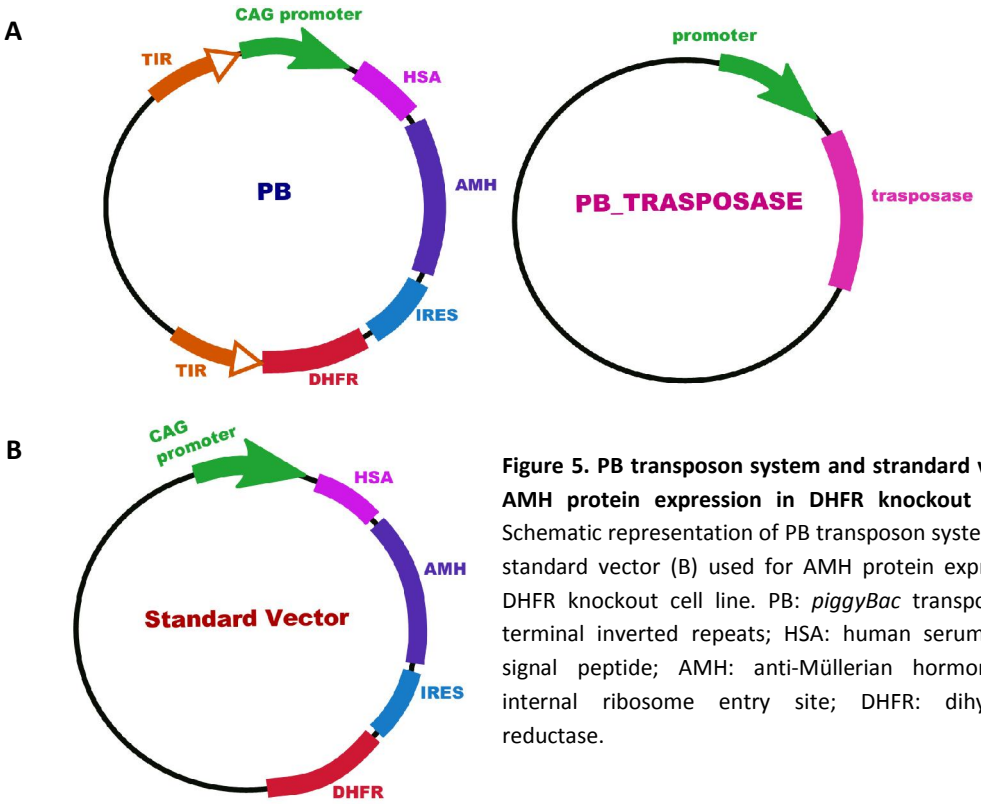


Figure 5. PB transposon system and strandard vector for AMH protein expression in DHFR knockout cell line: Schematic representation of PB transposon system (A) and standard vector (B) used for AMH protein expression in DHFR knockout cell line. PB: *piggyBac* transposon; TIR: terminal inverted repeats; HSA: human serum albumin signal peptide; AMH: anti-Müllerian hormone; IRES: internal ribosome entry site; DHFR: dihydrofolate reductase.

In order to set up a protocol for transfection of suspension cells we compared different methods: Nucleofection™ (Lonza), FectoPro™ (Polyplus), JetPEI® (Polyplus) and FreeStyleMax™ (ThermoFisher Scientific). They differ from each other in the method used to insert DNA into cells. Nucleofection™ is an electroporation-based transfection method, JetPEI® works exploiting a water-soluble polymer, FectoPro™ and FreeStyleMax™ use a lipid-based transfection technique. The experiments were performed on CHO-FS cells because of their easiness of handling. Indeed, these cells are prone to be efficiently transfected, easily grow in suspension at high density with a fast doubling time and do not need any specific requirements in the medium. A vector containing eGFP gene was exploited to evaluate the transfection efficiency. After 24 hours, cells from each condition were counted establishing the viability. Cells that were correctly transfected were evaluated assessing eGFP expression by fluorescent microscopy. The transfection efficiency was calculated as percentage of fluorescent transfected cells on viable cells (Figure 6).

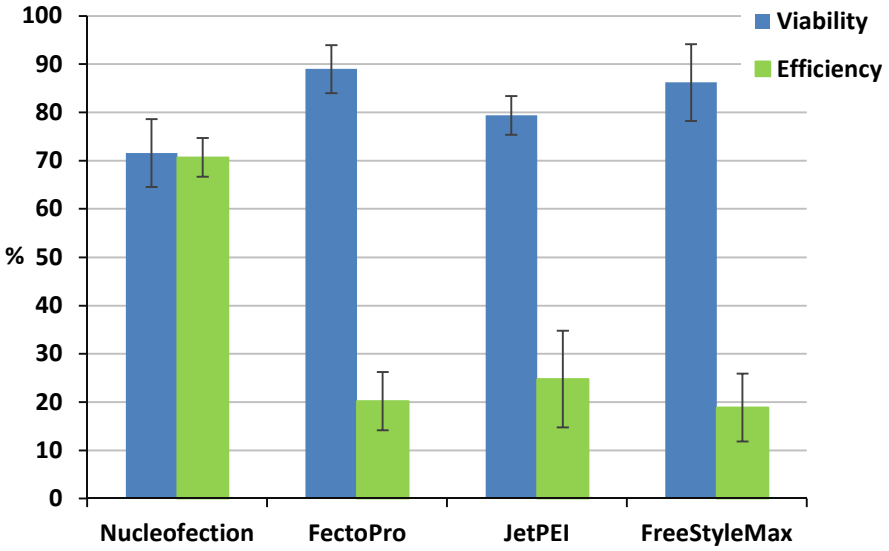


Figure 6. Comparison of transfection methods: Viability and transfection efficiency evaluation of Nucleofection™ (Lonza), FectoProTM (Polyplus), JetPEI® (Polyplus) and FreeStyleMax™ (ThermoFisher Scientific) transfection methods. CHO-FS cells were transfected with a vector containing eGFP gene exploiting the different methods. After 24 hours from transfection, viability and transfection efficiency were established for each condition. Fluorescence of transfected cells was evaluated using a fluorescent microscope. Transfection efficiency was calculated as percentage of fluorescent transfected cells on viable cells.

Despite all methods allowed to preserve high cell viability, they showed a very low transfection efficiency, with the exception of Nucleofection™. Indeed, it consented to obtain a very high transfection efficiency maintaining an acceptable cell viability. A possible explanation on Nucleofection™ success can be due to its different transfection strategy.

Most of transfection techniques allow the DNA insertion mainly into the cytoplasm, thus the nucleus can be accessible only during cell division when the nuclear membrane is eliminated. In contrast, Nucleofection™ combine cell-specific solution with electrical pulses consenting the DNA to enter directly into the nucleus.

Accordingly with previous evaluations, DHFR knockout cells were transfected exploiting Nucleofection™ protocol with either standard vector or PB transposon system, both carrying the DHFR gene and the AMH gene. Cells were cultured in a medium without HT for two weeks in order to select the transfected ones. Afterwards stable pools of cells were obtained, they were treated with methotrexate (MTX) for gene amplification. MTX was added directly to the selective medium following a single step protocol. In order to evaluate the effects of MTX on the amplification process, cells were treated with different concentrations of it: 50 nM, 100 nM, 250 nM, 500 nM, 1 µM and 2 µM (Figure 7).

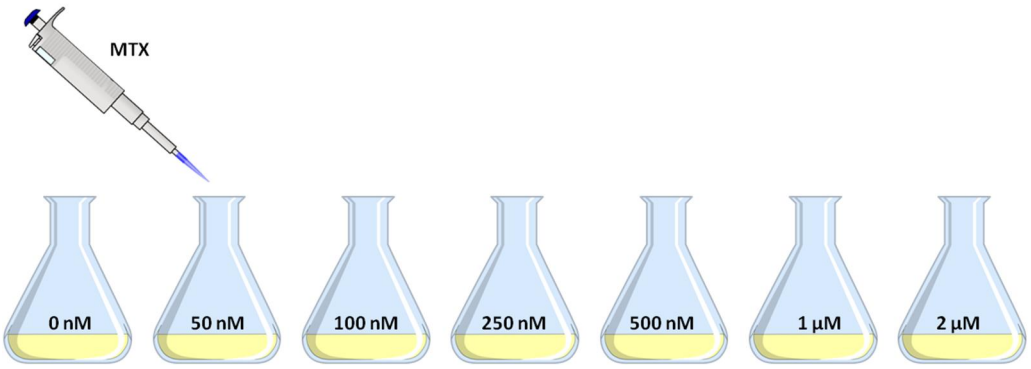


Figure 7. Single step protocol of MTX treatment: 2 weeks culture of 7 cell pools each treated with a different MTX concentration (i.e. 0 nM; 50 nM; 100 nM; 250 nM; 500 nM; 1 µM; 2 µM).

Cell pools were grown for 2 weeks in medium containing the proper MTX concentration. At each passage, the cell viability was evaluated (Figure 8A and 8B). After 5 passages in MTX treatment, each pool of cells had a viability above 90%. However, data showed that the stable pools derived from standard vector (from now on referred to as SV pools) treated with 1 µM and 2 µM MTX concentrations resulted in high mortality at the beginning; in particular the 2 µM MTX treatment had decreased the viability up to 30% at the second passage. On the contrary, the stable pools derived from the PB transposon system (from now on referred to as PB pools) maintained immediately a very high viability, superior to 80%, at each MTX concentration.

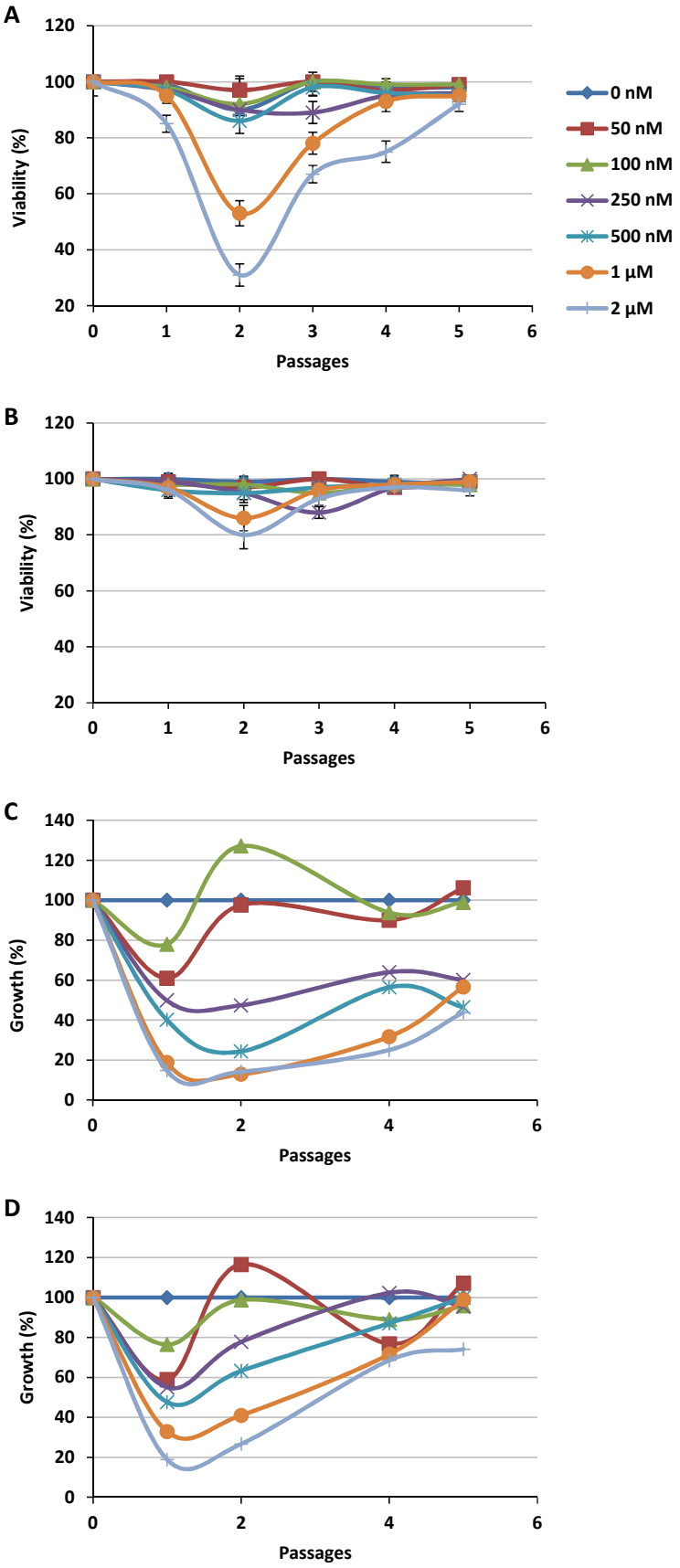


Figure 8. Viability and growth percentage of cell pools during MTX treatment at different concentrations: Viability and percentage of growth of SV stable pools (A, C) and PB stable pools (B, D) treated with different MTX concentrations (50 nM, 100 nM, 250 nM, 500 nM, 1 µM and 2 µM) compared to not treated one (0 nM). Stable cell pools were grown for 2 weeks (5 passages) in medium containing the proper MTX concentration. At each passage cell viability and cell growth were evaluated.
Growth percentages were normalized on values obtained with not treated pools that functioned as controls. SV: standard vector; PB: *piggyBac* transposon.

The same analysis was performed evaluating the cell growth rate, i.e. the number of cells at each passage expressed as a percentage of growth calculated on the values obtained in not treated cells (Figure 8C and 8D). Data showed that the growth percentage of SV pools treated with higher MTX concentrations resulted very lower than not treated pool. In contrast, over passages, the growth rate of PB pools was similar to that of not treated pool except for the cell pool treated with the highest MTX concentration (i.e. 2 μM). Even if the PB pool treated with 2 μM MTX concentration did not completely recover, it reached a good percentage of growth. We hypothyzed that the increased resistance of pools transfected with PB transposon to the selective agent could predict a more significant gene amplification.

At the end of the MTX treatment, the supernatant was harvested from each pool of cells (50 nM, 100 nM, 250 nM, 500 nM, 1 μM and 2 μM of MTX concentration) after 4 days of cells culture for the AMH quantification. AMH protein was quantified by a commercial ELISA assay and the results are shown in Figure 9. The productivity was evaluated using SPR (Specific Productivity Rate) that indicates nanograms of protein produced by a million of cells in a day. Each cell pool was compared to the control (pool not treated with MTX). Results demonstrated that the AMH expression augmented expecially after MTX treatment at the highest concentrations even if differences were observed between pools obtained from standard vector and the PB transposon system. In SV pools the maximum increment was only 3-3.5 fold. On the contrary, in PB pools, the MTX treatment with 500 nM, 1 μM and 2 μM allowed to obtain 5-, 10- and 19-fold increase, respectively. However, the increment was lower than what was expected referring to literature, except for that obtained with PB pool after 2 μM MTX treatment which fall within the reported values (Lucas et al., 1996; Wurm, 2004).

The MTX addition was also performed in a different way to evaluate its efficiency on the transgene amplification. A multiple step protocol was performed on the original stable cell pool treating it progressively with increasing MTX concentrations. Thus, the stable cell pool was treated starting from 50 nM MTX concentration. Every 7 days in selection, the viability was evaluated and, if it resulted above 90%, the MTX concentration was increased proceeding with the following steps: 100 nM, 250 nM, 500 nM, 1 μM, 2 μM MTX concentrations. In total, the treatment lasted about 6 weeks (Figure 10).

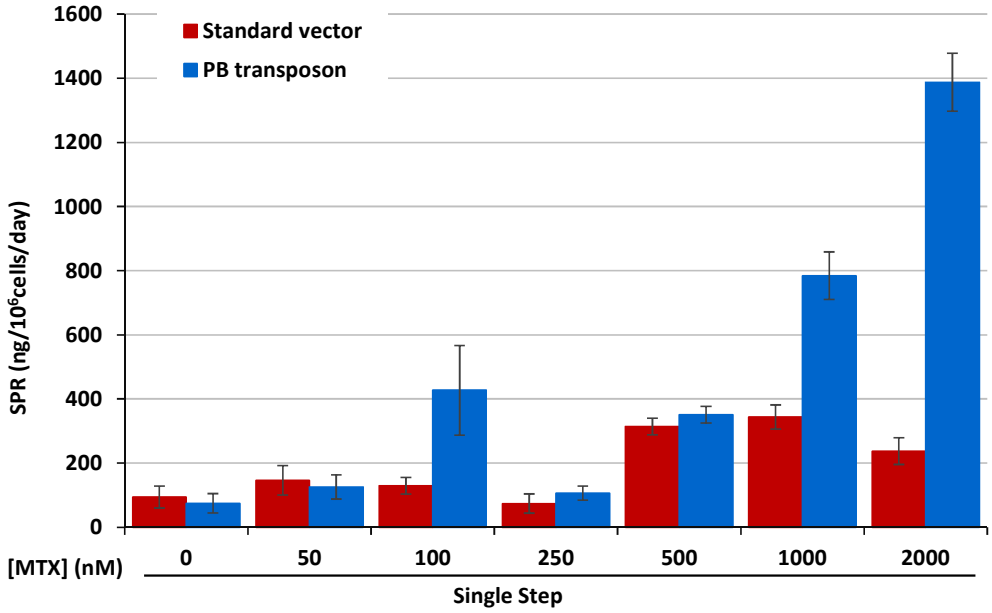


Figure 9. Productivity of SV and PB stable pools after single step MTX treatment protocol: AMH protein quantification on supernatant samples collected from SV and PB pools after MTX treatment at different concentrations (50 nM, 100 nM, 250 nM, 500 nM, 1 µM and 2 µM) compared to not treated one (0 nM). The supernatant from each pool of cells was harvested after 4 days of cell culture at the end of MTX treatment. The quantification of AMH protein was performed with a commercial ELISA assay. The productivity was evaluated using SPR (Specific Productivity Rate) that indicates nanograms of protein produced by a million of cells in a day. SV: standard vector; PB: *piggyBac* transposon.

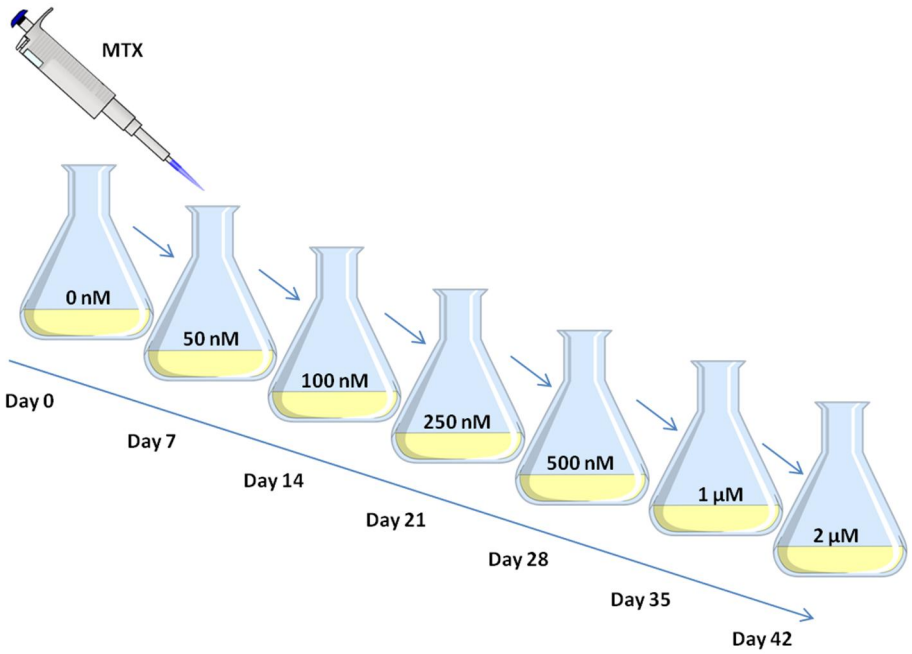


Figure 10. Multiple step protocol of MTX treatment: Culture of a cell pool treated with progressively increasing MTX concentrations (i.e. 0 nM; 50 nM; 100 nM; 250 nM; 500 nM; 1 µM; 2 µM). The increment was performed each 7 days if the cell viability was > 90%. The treatment lasted about 6 weeks.

The supernatant was harvested after 4 days of cell culture at each step of MTX concentration and AMH protein was quantified by the commercial ELISA assay. The AMH protein yields in supernatants collected from pools obtained at each round of MTX concentration were compared to the control (pool not treated with MTX) (Figure 11). Data showed a progressive increment of the AMH production at each increasing MTX concentration step. At the last step, the AMH expression in pools derived from transfection with standard vector or PB transposon compared to not treated stable pool was enhanced of about 5-fold and 7-fold, respectively. No significant difference in productivity increase was observed between SV and PB stable pools.

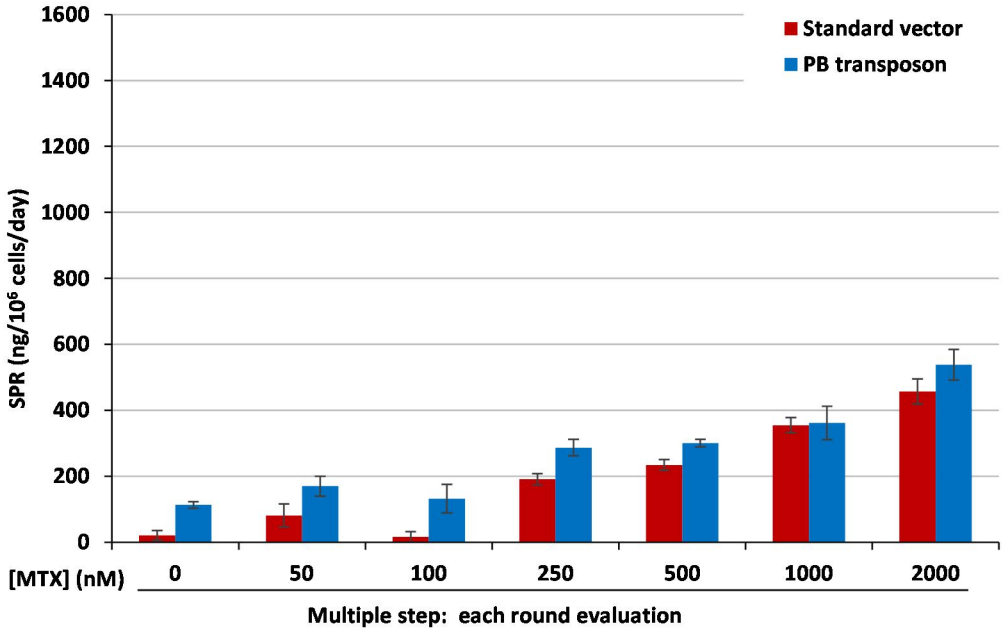


Figure 11. Productivity of SV and PB stable pools after multiple step MTX treatment protocol: AMH protein quantification on supernatant samples collected from SV and PB pools after multiple step treatment increasing progressively MTX concentrations (starting from 50 nM to 2 µM) compared to not treated one (0 nM). The supernatant from each pool of cells was harvested after 4 days of cell culture at each step of MTX concentration. The quantification of AMH protein was performed with a commercial ELISA assay. The productivity was evaluated using SPR (Specific Productivity Rate) that indicates nanograms of protein produced by a million of cells in a day. SV: standard vector; PB: *piggyBac* transposon.

Finally, the AMH protein expression of each pool treated at different MTX concentrations following the single step protocol was compared to the last round of the multiple step one and to the control (pool not treated with MTX) (Figure 9 and 11). The first observation to consider is that the AMH expressions of SV and PB not treated pools (0 nM) in the single step differ from the multiple step protocol. That is because in the single step protocol PB

and SV pools were evaluated after 2 weeks, which is the time necessary for other pools to be treated. Consequently, the difference in the AMH expression was due to the pools heterogeneity and the progressively dilution of high-producing cells over passages. However, PB pool resulted more productive than SV pool as shown in the first stage of the multiple step protocol. Focusing on SV pools, the AMH expression in pool derived from the multiple step protocol was higher than that obtained in all pools derived from the single step one. On the contrary, this did not occur in PB pools. The AMH expression in PB pools derived from the single step protocol treated with 1 µM and 2 µM MTX concentrations was higher than that obtained in the pool derived from the multiple step one. Moreover, the highest productivity was reached with the single step protocol at 2 µM concentration on PB pools. This was more than double of AMH yield obtained with the multiple step protocol.

Another consideration is that the amplification effects were more evident in PB pools compared to SV pools after the single step protocol, while no significant differences between PB and SV pools were observed in the multiple step one. In accordance with these evaluations, the multiple step amplification protocol resulted less convenient than the single step one because it is time-consuming and no great advantage was gained using it. Therefore, we decided to select SV and PB pools derived from the single step protocol with 0 nM (no treatment), 500 nM and 2 µM MTX addition as candidates for further analysis. Stable pools with no treatment functioned as control. The highest productivity reached in PB pool treated with 2 µM MTX treatment convinced us to examine in depth SV and PB pools derived from this MTX concentration. However, expression instability problems and MTX resistance mechanisms other than DHFR-mediated gene amplification could occur with high MTX concentrations (Kim et al., 2001; Kim and Lee, 1999). Thus, in order to have an alternative option in case of instability taking place at 2 µM MTX concentration we decided to also analyze another condition. We selected stable pools treated with 500 nM MTX because at 1 µM we supposed to encounter the same problems described for the 2 µM MTX treatment.

2.3 High producing clones frequency

Single cell cloning by limiting dilution was performed on stable pools treated with 0 nM (not treated), 500 nM and 2 µM MTX concentration. the single clones derived from each stable pool were isolated and their productivity was evaluated (Figure 12). Cell number in each

well was calculated by CCK-8 Dojindo assay. A DiaSorin automated prototype immunoassay for the detection of AMH in serum and plasma samples developed on Liaison platform was set up for quantification analysis. The quantitative values were normalized to both clone cell number and days of culture, obtaining a final production rate value (Specific Productivity Rate or SPR) expressed as nanograms of protein produced by 10^6 cells per day (ng/10^6 cells/day). Results obtained from the quantitative screening were plotted arbitrarily defining production categories and estimating the percentage of clones for each experimental condition that fell in each category.

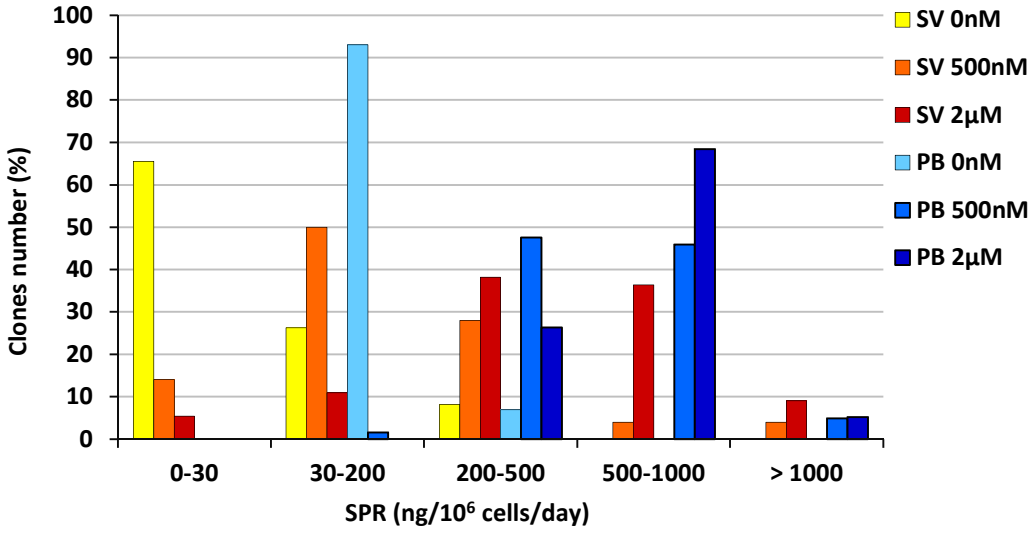


Figure 12. Productivity of clones: Quantitative screening of clones generated from SV or PB pools treated with 500 nM and 2 µM MTX concentration compared to not treated ones (0 nM). Clones were classified based on AMH protein expression level. Cell supernatants were analyzed with an automated prototype assay for the detection of the AMH protein. Values were normalized based on clone cell number (Dojindo cell counting assay) and days of culture. Data were reported as percentage of clones classified into arbitrary categories of production. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day. SV: standard vector; PB: *piggyBac* transposon.

Standard vector (not treated, yellow bars) generated more than 50% of clones with a very low productivity. This is probably due to random integration of standard plasmid, since PB transposon (not treated, light blue bars) allowed to obtain only productive clones, confirming previous data obtained by our lab. However, by using the transposon system and the DHFR knockout CHO cell line without MTX gene amplification, no clones were obtained in the highest productive category, that is in contrast to what observed previously by using the CHO-K1 adherent cell line (see paragraph 2.1).

After gene amplification by MTX, the PB transposon system generated a high percentage of clones with a medium-high productivity by using both 500 nM and 2 µM MTX treatment. This result was partially reached by clones derived from standard vector only if it was applied the maximum amplification strength, i.e. 2 µM MTX.

At this point, we selected the two highest-producing clones from each stable pools (not treated, 500 nM and 2 µM MTX treatment) for further characterization. Our aim was to select the best clone for industrial AMH protein production, that is a clone with the highest AMH protein productivity along with genetic and expression stability over time.

2.4 Clones productivity and stability over time

Industrial recombinant protein production requires a severe characterization of selected clones in order to avoid the productivity reduction over time. This event could occur due to genetic or expression instability of clones. Therefore, once the two highest-producing clones from stable pools treated with 0 nM, 500 nM and 2 µM MTX concentrations were selected, they were maintained in culture for up to 3 months to evaluate fundamental parameters. The selected clones isolated from SV and PB stable pools from now on referred to as SV and PB clones, respectively, are listed in Table 1.

origin stable pool	SV clones	PB clones
0 nM	49A.13	49D.213
	49A.16	49D.221
500 nM	49B.82	49E.259
	49B.84	49E.279
2 µM	49C.158	49F.330
	49C.170	49F.372

Table 1. Selected clones: schematic representation of selected clones. SV and PB selected clones were divided based on their pool of origin. SV: standard vector; PB: *piggyBac* transposon.

A cell line is deemed unstable if the titer declines by more than 30% when subjected to 45–55 population doubling (PDs) (Dorai et al., 2012). In our case, cells made 2 PDs among each passage, thus 55 PDs correspond to about 27 passages. This duplication time is sufficient to cover the time for culture scale-up and harvest needed for protein production at industrial levels. The selected clones were maintained in suspension culture for 30 passages. Samples of cell pellet were collected after 11, 22 and 30 passages in order to study the genetic stability over time. A qPCR analysis on genomic DNA extracted from the cell pellets was realized to evaluate the copy number of the transgene integration (Figure 13A). Moreover,

the AMH expression was checked at the same time points (11, 22 and 30 passages) for each clone. The quantification was performed on supernatants obtained from 4 days cell culture with the automated prototype immunoassay on Liaison platform (Figure 13B).

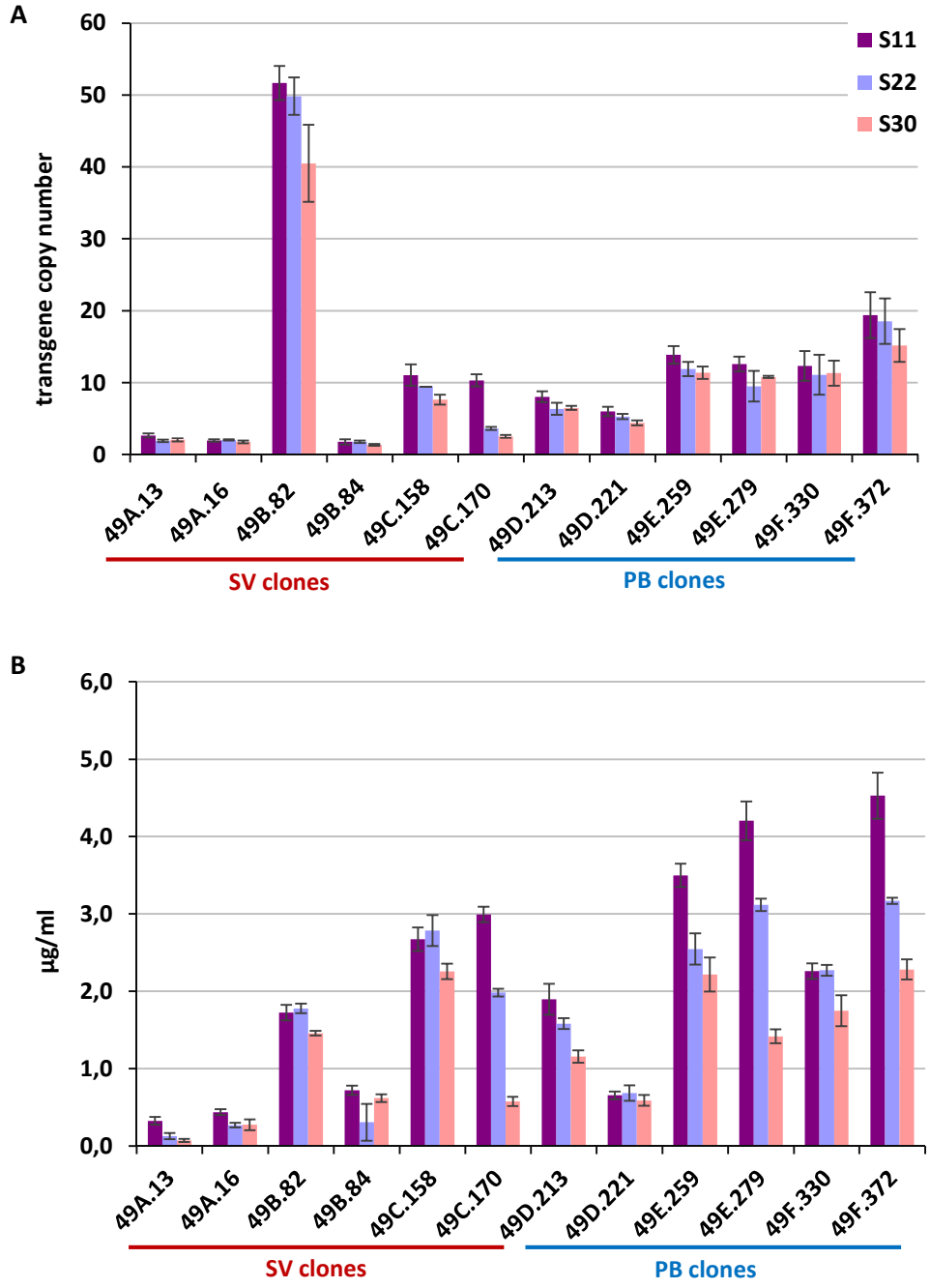


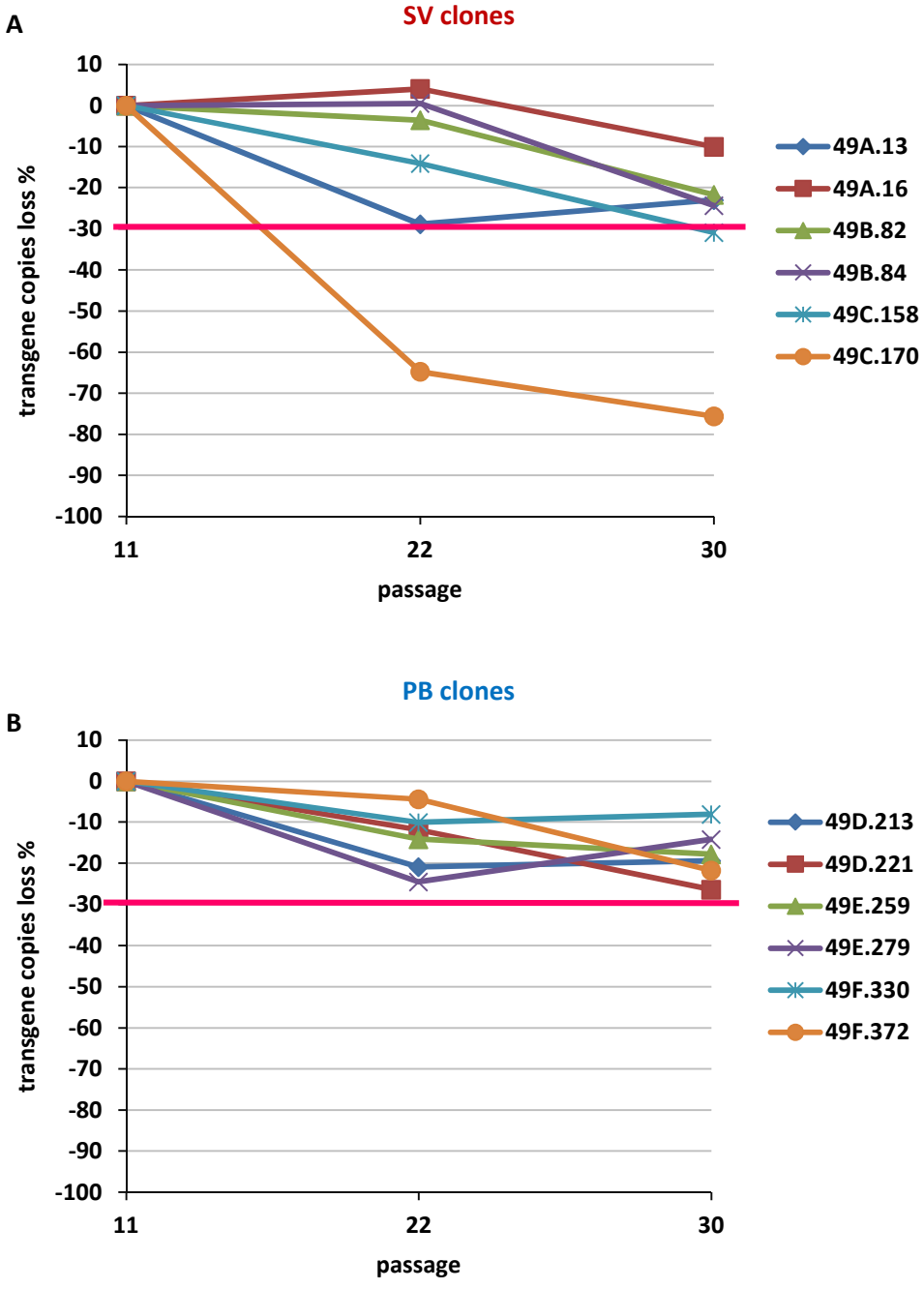
Figure 13. Genetic and expression stability of selected clones: A. Genetic stability of selected clones over time. qPCR analysis on genomic DNA extracted from selected clones at passages 11, 22 and 30 (S11, S22, S30). B. Productivity of selected clones over time. AMH protein expression quantification of selected clones with the automated prototype immunoassay on Liaison platform from 4 days cell culture supernatants collected at passages 11, 22 and 30 (S11, S22, S30).

Although only a few clones were analyzed data revealed some considerations to take into account. Regarding the genetic analysis (Figure 13A), focusing on not treated SV and PB clones, we firstly observed that PB clones showed 2-3 fold more transgene integrated copies than SV ones, as expected by the active integration mediated by transposase. The treatment with 2 µM MTX helped to reduce differences in integration of transgene copies between SV and PB clones because it consented a 5-fold enrichment of integrated copies in SV clones but only a 2-fold enrichment in PB ones. However, such a high selective pressure could generate genetic instability as probably happened in SV clone 49C.170. The treatment with 500 nM MTX resulted in a 2-fold increase of integrated transgene copies in PB clones compare to not treated one. In contrast, the same treatment did not lead to gene amplification in SV clones, as shown in clone 49B.84. An exception occurred in the SV clone 49B.82 that accounted for the highest transgene copy number, about 50 integrated copies. We hypothesized that such high copy number insertions derived from an amplification process of an integrated concatemer. Despite the higher amplification level occurred in treated SV clones compared to treated PB clones, the latter showed higher copy number integration than SV clones (with the exception of clone 49B.82) confirming the greatest potential of the transposon system than the standard delivery method not only in absence of MTX treatment, but also after it.

The same consideration can be done for the protein productivity (Figure 13B), focusing on not treated SV and PB clones we observed a 3.5-fold increase in the AMH expression in PB clones compared to SV ones that was congruent with their difference in copy number of integrated transgenes. Consistently with the genetic data, the treatment with 2 µM MTX allowed to reduce the gap in protein productivity between SV and PB clones because it consented about a 7.5-fold increment in the AMH expression in SV clones but only 3-fold in PB ones. In addition, a significant expression instability was observed in SV clone 49C.170 which reflected its genetic instability. In PB clones, the treatment with 500 nM MTX resulted in an average of 3-fold increment in the AMH expression compared to not treated ones showing, also in this case, congruency with the genetic data. On the contrary, an inconsistency with the gene amplification results was observed in 500 nM treated SV clones. Despite the invalid effect of the MTX treatment on the clone 49B.84 in amplifying transgene copy number, a 2-fold increase in protein productivity was obtained. Mechanisms other than gene amplification occurred in this clones that enhanced the AMH expression. However, the most evident inconsistency was observed in the clone 49B.82

whose AMH protein expression did not absolutely correlate with the integrated transgene copies. Since we supposed that it was an amplified concatemer, it could undergo silencing, a phenomenon that can easily occur in such repeated sequences (McBurney et al., 2002).
In order to focus on the genetic stability over time, data were evaluated as percentage of transgene copy loss over passages based on the values obtained at passage 11 (Figure 14A and 14B). A transgene copy loss ≤ 30% (fuchsia line) was considered acceptable for our needs. We observed that there was a genetic instability over time in clones derived from both standard vector and PB transposon. However, the transgene copy loss was more significant in SV clones than in PB ones. In particular, the increase in MTX treatment led to a greater instability in SV clones that was more evident at passage 30. On the contrary, PB clones preserved a more homogeneous genetic stability profile among different culture conditions. Anyway, the clones derived from both standard vector and PB transposon resulted in a maximum transgene copy loss of 30% after 30 passages, with the exception of the SV clone 49C.170 that showed a loss of 80%. An analogous evaluation was performed on the productivity in order to put emphasis on the expression stability over time. Data were evaluated as percentage of productivity loss over passages based on the values obtained at passage 11 (Figure 14C and 14D). A percentage of productivity loss that went beyond 30% (fuchsia line) indicated clone expression instability. At passage 22, PB clones showed less instability than SV clones. In fact, all PB clones showed a productivity loss beneath 30%, even if a progressive slight drop was observed. On the contrary, at the same passage, only two of SV clones preserved the protein expression, without any decrease. The unexpected outcome was that these clones, that were 49B.82 and 49C.158, derived from pools treated with MTX instead of not treated ones, which were not stressed by high selective pressure. Nevertheless, after 30 passages, seven of twelve total clones were shown to be remarkably unstable with a productivity loss superior than 30%, i.e. clones 49A.13, 49A.16, 49C.170, 49D.213, 49E.259, 49E.279 and 49F.372. Evaluating in details the PB clones loss at passage 30, the unstable PB clones resulted in a loss of 40-50% with a maximum loss of about 66% in the clone 49E.279. Surprisingly, the two PB stable clones were one derived from not treated pool and one from 2 µM MTX treated pool (i.e. clones 49D.221 and 49F.330), which exactly underwent to the opposite treatment. On the other hand, regarding SV clones at passage 30, a loss as high as 80% was observed in two of SV clones (i.e. clones 49A.13 and 49C.170). Note that it was not clear why the clone 49B.84 showed a very high productivity loss at passage 22, while it did not occur at passage 30. A

possible explanation could be a problem in the supernatant sample harvesting at passage 22.



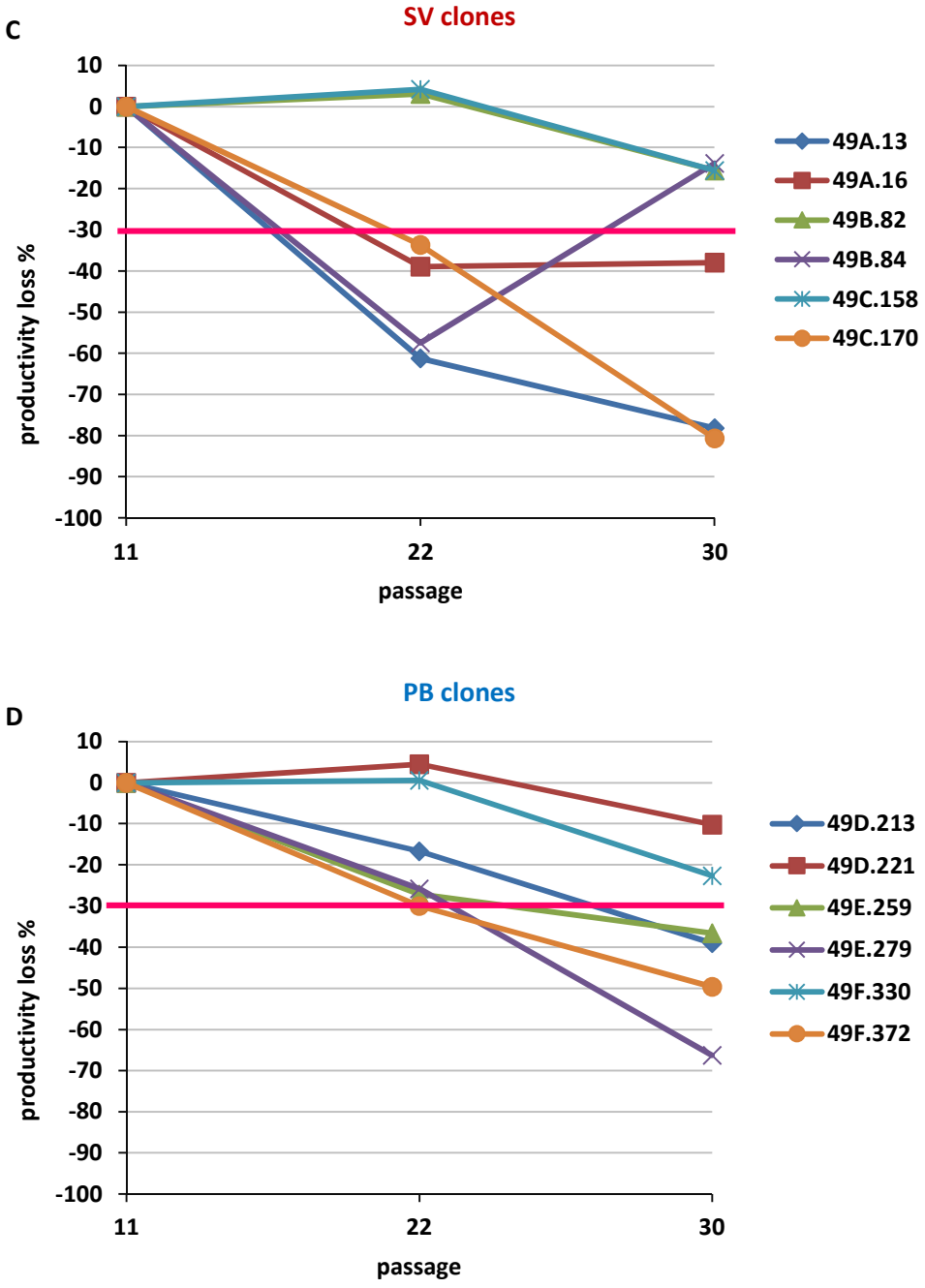


Figure 14. Genetic and expression stability of selected clones expressed as percentage of transgene copy loss and productivity loss: A, B. Percentage of transgene copy number loss in SV and PB selected clones at passages 22 and 30 compared to passage 11. Fuchsia line indicates the 30% cut-off of acceptable transgene loss. C, D. Percentage of productivity loss in SV and PB selected clones at passages 22 and 30 compared to passage 11. Fuchsia line indicates the 30% cut-off of acceptable AMH expression loss.

All these data taken together convinced us to proceed in depth analysis with PB clones derived from 500 nM MTX treatment. We preferred PB clones derived from the middle MTX treatment because they allowed to obtain the same AMH protein expression of clones derived from 2 µM MTX treatment avoiding an excessive pressure strength. Besides, we did not consider SV clones derived from the same MTX treatment a good choice due to their low protein yields, the consistent genetic and expression loss observed in one of the two clones and the risk linked to a possible integration as concatemer. Based on these considerations, the PB transposon system in combination with DHFR knockout cell line and a 500 nM MTX treatment was demonstrated to be the most confident strategy to obtain high-producing clones not prone to silencing and maintaining an acceptable stability over time. Among clones showing these characteristics, we selected the clone 49E.259 because it was the best.

2.5 Culture condition optimization

At this point, our aim was to evaluate the best culture condition for the AMH protein production for industrial purposes. Therefore, an adaptation to different media and several additives was performed on the selected clone 49E.259. The culture medium is an essential component of the in vitro environment and it can dramatically affect the protein production. Different commercial media are now available. We tested two media on the selected clone in comparison with its basal medium: Cellvento™ CHO-100 Liquid (Merck Millipore) and ActiPro (GE Healthcare). Cellvento (CV) medium is used for several types of CHO suspension cells, it is particularly rich in nutrients and suitable for batch and perfusion application. It has been already tested in our laboratory and gave the best performance in term of protein yield compared to PROCHO4 (Lonza) (data not shown). ActiPro (AP) is also a versatile medium. It was selected because it supports DHFR deficient cell lines with appropriate metabolites and it is suitable for batch and fed-batch culture. Indeed, it is commercialized in association with boosts specifically formulated for fed-batch culture, one of our outcomes for industrial AMH protein production. Since Cellvento medium has not a specific formulation for DHFR deficient CHO cell line unlike ActiPro medium, we decided to test the same additives used in the basal medium such as the dextran sulfate (DS, Sigma Aldrich) and the non-essential amino acids (MEM, ThermoFisher). In particular, the dextran sulfate was added to prevent cell aggregation, thereby increasing cell viability. The clone

adaptation was progressive and required one month (see Material and Methods). Once the adaptation was finished, 4 days cell culture supernatants were collected. The quantification with our immunoassay prototype (Table 2) and Western Blot analysis in reducing and non-reducing condition were performed on the supernatants to analyze the AMH protein expression (Figure 15). Results demonstrated that a comparable AMH protein yield was quantified in the clone adapted to ActiPro and Cellvento media in presence of dextran sulfate (CV+DS, CV+DS+MEM) and in the clone maintained in the basal medium. Cellvento medium without any adds or in presence of non-essential amino acids resulted in a lower protein expression compared to that obtained with the clone cultured in other media. Immunochemistry data were corroborated by WB analysis. An immunodecoration with a commercial α-AMH antibody showed the non-cleavable AMH protein as a monomer at 75 kDa in reducing condition and as a homodimer at 150 kDa in non-reducing condition confirming the protein profile.

Medium	µg/mL	SD
BM	5,11	0,5
CV	3,18	0,4
CV + DS	4,49	0,7
CV + MEM	2,51	0,4
CV + DS + MEM	4,21	0,6
AP	4,12	0,3

Table 2. Productive media comparison: Clone productivity after adaptation to different media expressed as mean ± SD. Supernatants were collected after 4 day of cell culture and quantified with DiaSorin immunoassay prototype.
BM: basal medium; CV: Cellvento; AP: ActiPro; DS: dextran sulfate; MEM: non-essential amino acids.

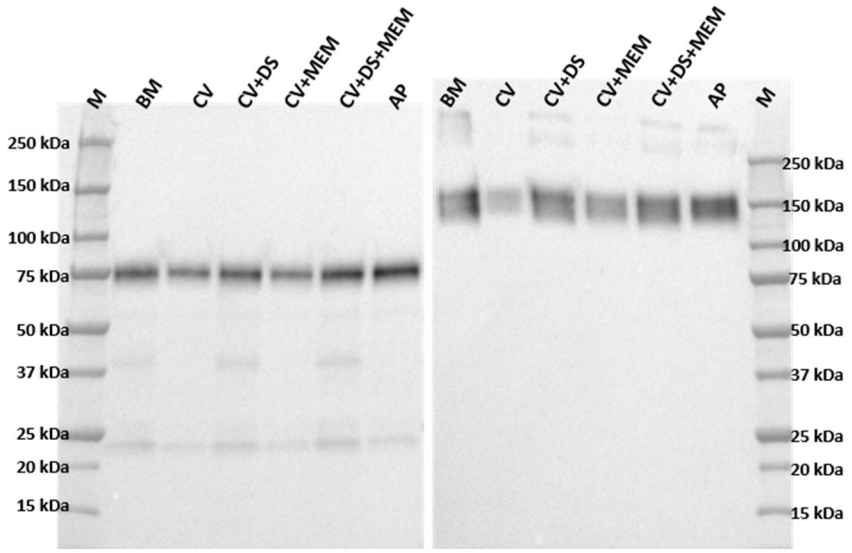


Figure 15. Productive media comparison: WB analysis on AMH protein expression in selected clone adapted to different media.
On the left, reducing condition. On the right, non-reducing condition.
A commercial α-AMH antibody was used for immunodecoration.
M: marker; BM: basal medium; CV: Cellvento; AP: ActiPro; DS: Dextran sulfate; MEM: non-essential amino acids.

This analysis revealed that the adaptation of clone in Cellvento medium did not lead to a better protein production compared to the basal medium in which the clone already grew. It is not the case when either additives were added, even if a slight improvement in protein yield was observed. The dextran sulfate was confirmed to operate an anti-aggregation effect as demonstrated by the cell aggregation in cell cultures in absence of this additive which, in turn, resulted in a lower protein expression (CV and CV+MEM). Curiously, cells treated with dextran sulfate showed a band at around 40 KDa in reducing condition; such band was absent in media without this additive. In non-reducing condition, no significant differences in aggregation pattern were observed. On the contrary, ActiPro medium showed a good protein expression and it can be a valid alternative medium to the basal one. Despite this evidence, we decided to continue experiments focusing on the clone with its basal medium for a couple of reasons. Firstly, the clone grown in the basal and ActiPro media resulted in a similar productivity. Secondly, the disadvantage for a clone adapted to a different medium is its aging because at least 1 month (8 passages) is required to complete the adaptation procedure. This means that the derived clonal cell line has a more limited time for its industrial application compared to the origin clone.

2.6 Production scale-up

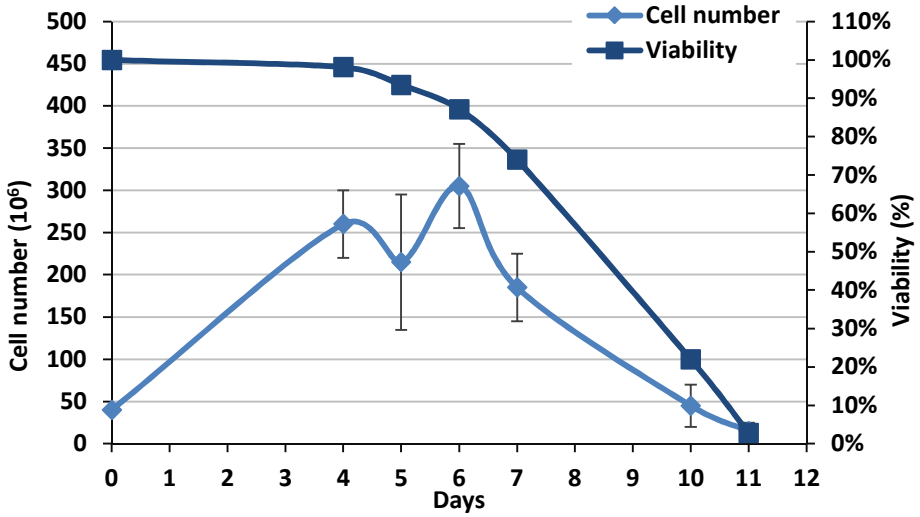
We proceeded exploring the production scale-up. Our experiments aimed also to investigate the mild hypothermia condition on the protein production, since it can enhance the protein expression (Kumar et al, 2008 and Sunley et al., 2008). Indeed, it is reported that a temperature decrease below the optimal condition (30-33°C) during cell culture results in an increase of the cell-specific recombinant protein production rate. The mild hypothermia impacts on the regulation of several cell pathways such as cell cycle arrest, decrease of apoptosis, reduction in nutrient uptake and in waste metabolites accumulation. In this way, the upregulation of the heterologous protein expression is promoted as well as the reduction of the intermolecular aggregation of their products. The scale-up procedure was performed in 1-liter flask with a cell density of 0.4 x 10^6 cells/ml at 125 rpm orbital shaking in a total volume of 100 ml. The temperature was shifted to 32°C the day after the inoculum in order to obtain a high cell density. The cells cultured at 37°C reached a peak of growth at the day 6^{th} with 300 x 10^6 cells and a good viability, which was superior to 80%. After that time point, the viability drastically decreased and the cell death occurred at the

day 11th (Figure 16A). The quantification analysis performed with our prototype immunoassay on Liaison platform showed that the maximum AMH protein yield obtained was about 8 µg/ml (Figure 16B). On the contrary, the cell culture grown at 32°C maintained the viability superior than 70% until the day 11th. The cell number did not exceed 230 x 10^6 confirming the growth arrest due to the mild temperature condition (Figure 16C). This strategy led to increase the clone productivity of about 2-fold reaching a protein yield of 15 µg/ml (Figure 16D). Western Blot analysis confirmed the higher protein production obtained in mild hypothermia condition compared to 37°C culture. The protein profile was not affected, actually the AMH protein degradation decreased at the secondary cleavage site (Figure 16E).

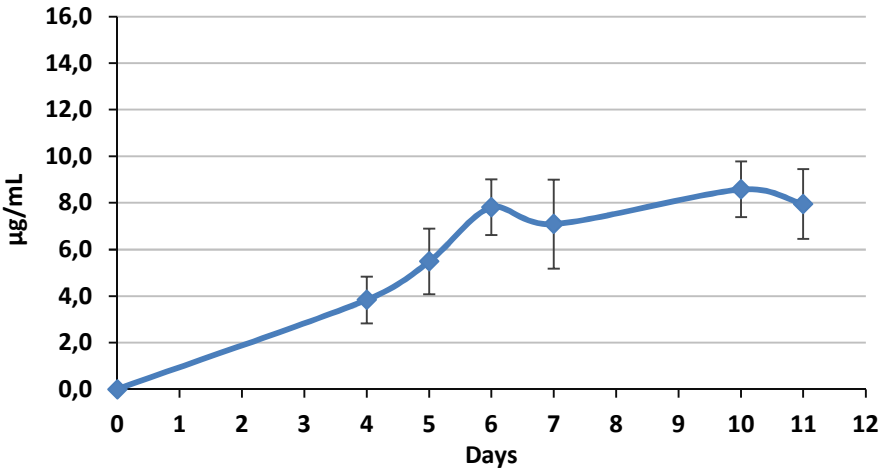
In standard condition, the AMH protein degradation started at the beginning of the culture, while it was slightly visible from the time point 10 in mild hypothermia condition. This was probably due to the initial phase of cell death, suggesting that in this medium the best time point to harvest supernatant was the day 7th and not the day 10th despite the higher protein expression. Following the mild hypothermia protocol, it was possible to produce AMH protein at a concentration of 13 µg/ml.

After the culture scale-up, we expected a higher clone productivity than what we obtained. We also tested the mild hypothermia scale-up protocol on the selected clone adapted to the ActiPro medium (data not shown). The aim was to improve the protein production and to explore the fed-batch culture with boost additions. Despite the addition of feeds, we obtained only a slight improvement in protein expression. Therefore, we decided to restart the whole process starting from an accurate vector engineering.

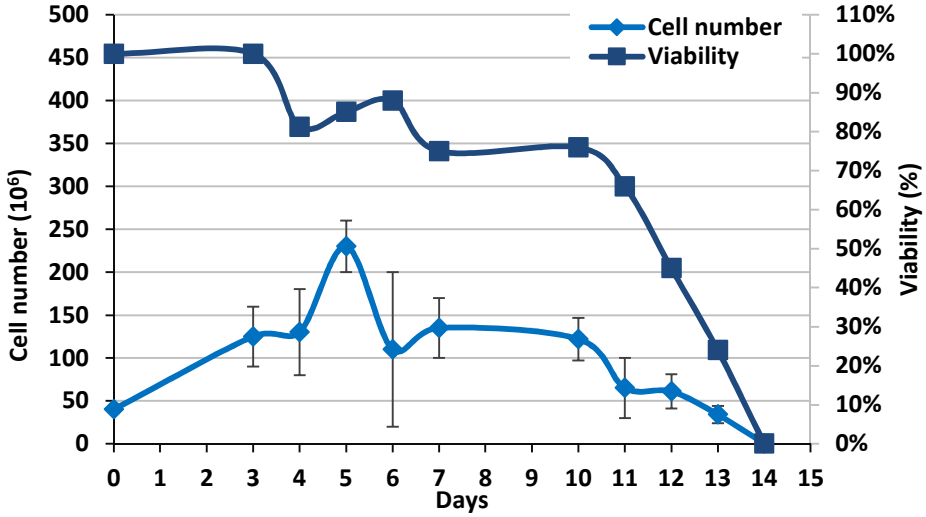
A


Cell number
Viability
Cell number (10⁶)
Viability (%)
Days

B


µg/mL
Days

C


Cell number
Viability
Cell number (10⁶)
Viability (%)
Days

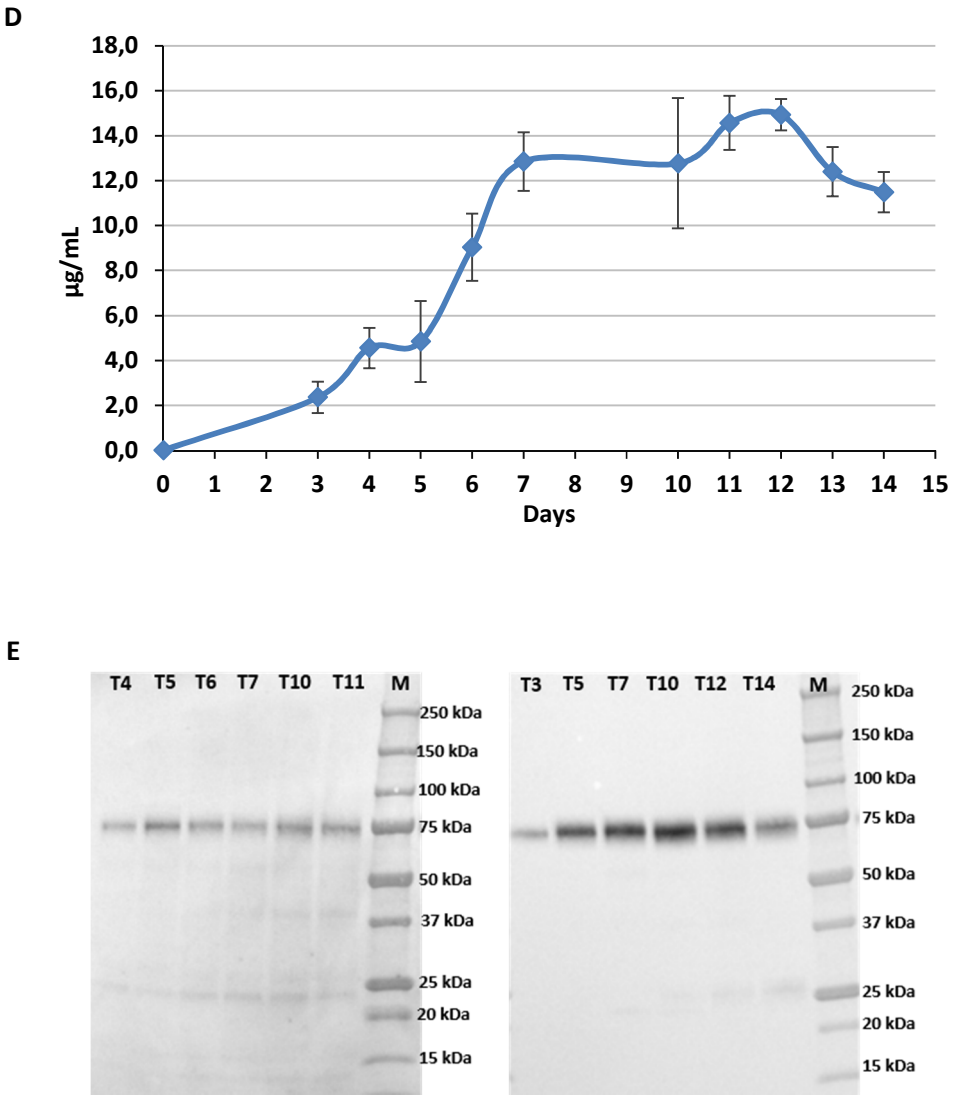


Figure 16. Production scale-up in standard and mild hypothermia condition: Growth curve of cell culture in scale-up (1 liter flask with a cell density of 0.4 x 10^6 cells/ml at 125 rpm orbital shaking in a total volume of 100 ml). At each time point, cell number and percentage of viability were evaluated. A. Growth curve maintained at a temperature of 37°C. Supernatant samples were collected at day 4, 5, 6, 7, 10 and 11. B. Curve of AMH protein quantification on supernatant samples at each collected time point in production scale-up maintained at 37°C. C. Growth curve of cell culture in scale-up in mild hypothermia condition. Supernatant samples were collected at day 3, 4, 5, 6, 7, 10, 11, 12, 13 and 14. D. Curve of AMH protein quantification on supernatant samples at each collected time point in mild hypothermia production scale-up. E. WB analysis of production scale-up in standard (on the left) and mild hypothermia condition (on the right) immunodecorated with a commercial α-AMH antibody.

Part 3: Strategies for optimization of gene expression

3.1 DHFR destabilization and amplification

In spite of the medium change and the mild hypothermia condition, the AMH protein yield in scale-up was unsatisfactory. Since the MTX gene amplification was expected to be higher than obtained, we focused on construct engineering to improve the clone productivity. Thus, strategies to reduce the DHFR activity by acting on its expression pathway were performed to enhance the AMH protein production. Several approaches are reported in literature that aim the reduction of the DHFR expression and result in the increase of the DHFR gene amplification as a cell response to the higher selection stringency leading to enhanced recombinant protein expression. Strategies can be performed editing the plasmid vector in order to target various steps of the protein expression pathway, acting on:

- DNA, performing a DHFR codon deoptimization in order to reduce the translation efficiency (Westwood et al., 2010). Here, the codon deoptimization was performed manually by using a codon usage table.
- mRNA, inserting a less efficient internal ribosome entry site (IRES) than the standard encephalomyocarditis virus (EMCV) one. We selected the foot and mouth disease virus (FMDV) IRES, in order to reduce the ribosome-mRNA affinity (Ho et al., 2013; Ng et al., 2012; “Other Plasmid Components,” n.d.);
- protein half-life, introducing PEST sequences (proline (P), glutamic acid (E), serine (S) and threonine (T)) from the ornithine decarboxylase. Two PEST sequences in tandem were inserted, the second one carrying the mutation S445A in order to strengthen the protein degradation (Li et al., 1998; Ng et al., 2007).

The transfections were performed with eight different PB transposon bicistronic constructs. They were realized based on the classic PB transposon system used in previous experiments (Figure 5A) with the insertion of the elements listed before.

The first four vectors consisted of the unmodified EMCV IRES with or without the insertion of PEST sequences and deoptimized DHFR (Figure 17A). The other four vectors consist of the FMDV IRES with the insertion of the same modifications designed for the previous ones (Figure 17B).

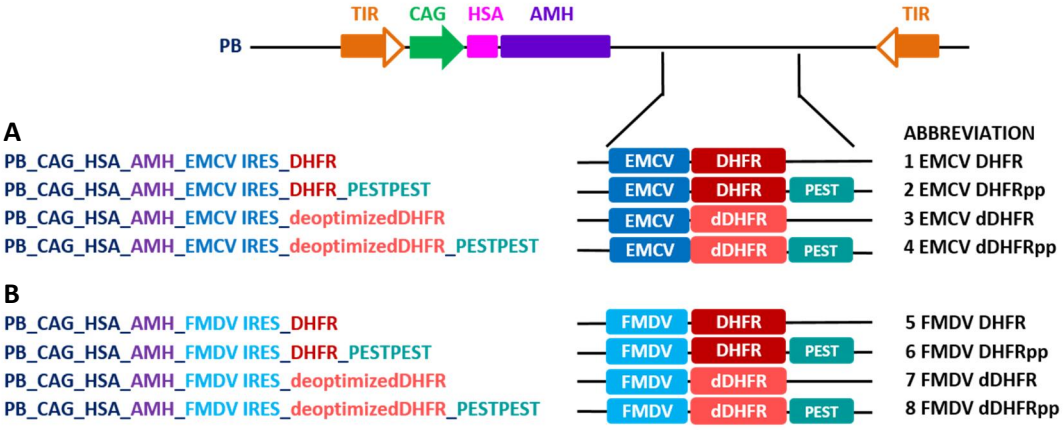


Figure 17. PB transposon vectors engineering: schematic representation of engineered PB vectors containing EMCV IRES (A) or FMDV IRES (B) and different modifications affecting DHFR expression pathway. Abbreviations used in the paragraph are also reported. PB: *piggyBac* transposon; TIR: terminal inverted repeats; HSA: human serum albumin signal peptide; AMH: anti-Müllerian hormone; IRES: internal ribosome entry site; EMCV: encephalomyocarditis virus; FMDV: foot and mouth disease virus; DHFR: dihydrofolate reductase.

Cells were transfected with designed vectors and transferred in the selective medium. Initially, the viability was very low. After 12 days from transfection, the cells reached both good viability and cell density to be transferred in flasks (125 ml) in an orbital shaker incubator. Then, a 500 nM MTX treatment was started. We chose to test only the 500 nM MTX treatment in single step protocol in order to maintain the reference conditions previously tested. After two weeks, the MTX treatment was stopped and the supernatants from cell pools were collected. From now on we refer to the pools derived from transfection with constructs containing EMCV IRES as EMCV pools and those from transfection with constructs containing FMDV IRES as FMDV pools. Not all of the cell pools survived. Death occurred in pools carrying the most intense constrains, i.e. pools 6 FMDV DHFRpp, 7 FMDV dDHFR and 8 FMDV dDHFRpp. The destabilization in the DHFR pathway for these pools was probably too detrimental to survive with a 500 nM MTX treatment.

At each passage, the cell viability was evaluated (Figure 18A). Only three of all pools maintained the viability above 80%, i.e. 1 EMCV DHFR, 2 EMCV DHFRpp and 4 EMCV dDHFRpp. The others pools showed a significant drop from the passage 2 and only the pools 3 EMCV dDHFR and 5 FMDV DHFR recovered from the passage 4. As disclosed before, the pools 6 FMDV DHFRpp, 7 FMDV dDHFR and 8 FMDV dDHFRpp did not survive.

The same analysis was performed evaluating the growth rate normalized to the values from the classical PB transposon (1 EMCV DHFR) that acted as control (Figure 18B). Results demonstrated that EMCV pools had an initial drop at passage 3, but they maintained a

growth percentage over 20% followed by a gradual recover. In contrast, the growth percentage of all FMDV pools reached about 10% at passage 3. The only FMDV pool that survived showed a slow recover, but the growth percentage was still very low (18%) at passage 5.

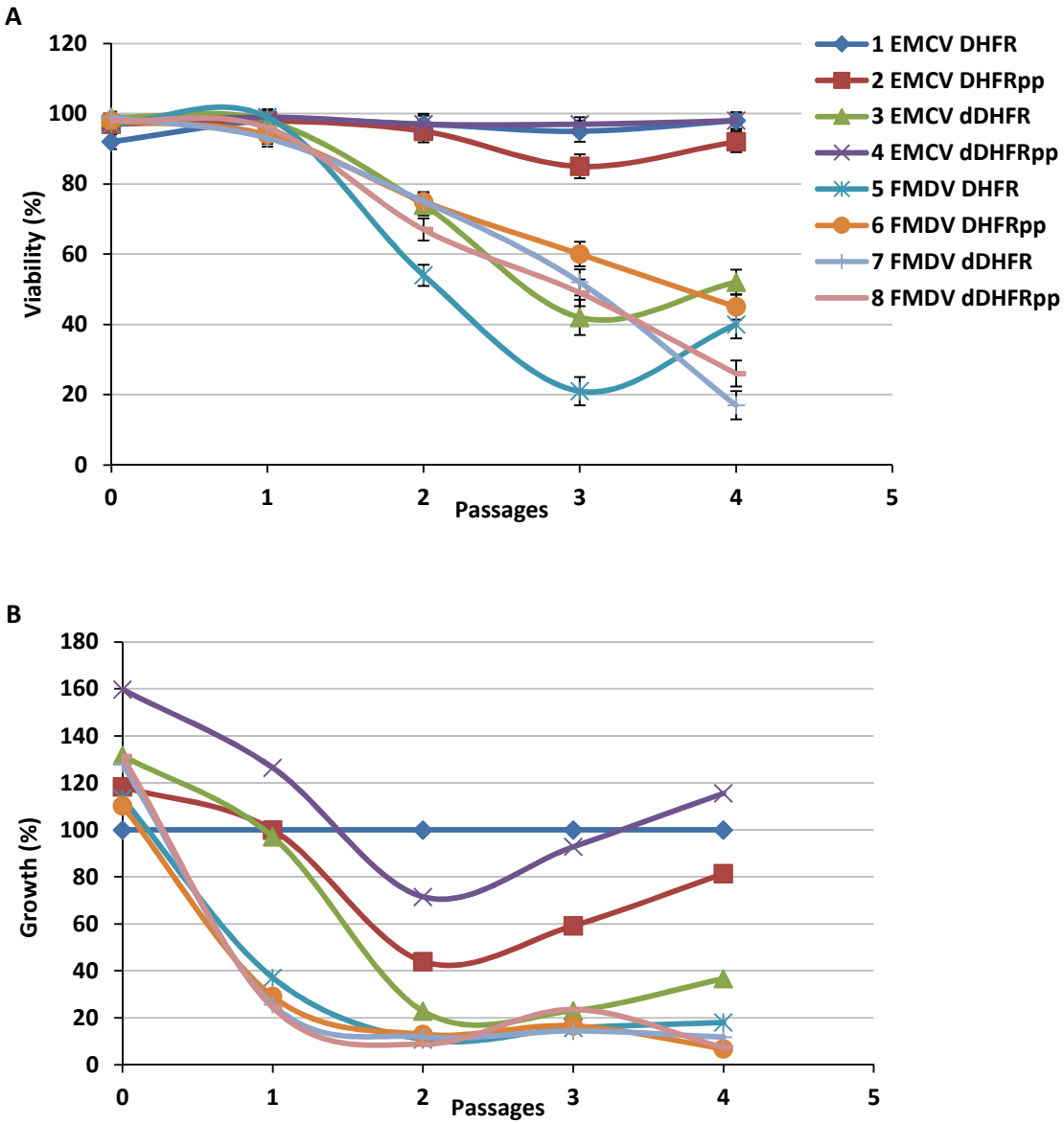


Figure 18. Growth curves of cell pools during 500nM MTX treatment: Growth curves of cell pools transfected with engineered PB transposons affecting DHFR expression pathway. Cells were cultured for 5 passages. At passage 1, 500 nM MTX treatment was started and maintained until passage 5. At each passage, cell number and percentage of viability were evaluated. A. Percentage of viability representation for each cell pools at each passage B. Percentage of growth for each cell pool at each passage. Values were normalized on that obtained from cell pool 1 EMCV DHFR that functioned as control.

The collected supernatants were quantified for AMH protein using our prototype immunoassay on Liaison platform (Figure 19). EMCV pools did not show advantages in protein expression compared to cells transfected with the classic construct used in previous experiments. It is not the case either after the MTX treatment at 500 nM concentration. On the contrary, FMDV pools showed a 1.5-2 fold increase in protein production even before the MTX treatment compared to the control. However, it was impossible to evaluate the protein production after the MTX treatment because these cell pools did not survive to the MTX amplification, except the one derived from the construct with only IRES modification. Data obtained after the MTX treatment in the cell pool that survived was encouraging, therefore, we decided to try again the MTX amplification process using lower MTX concentrations to evaluate the amplification effect in FMDV pools.

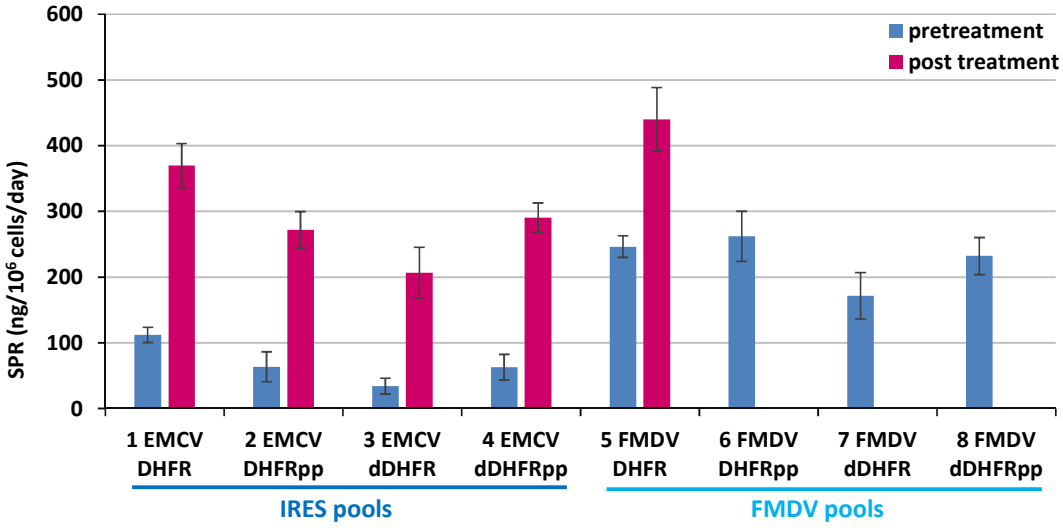


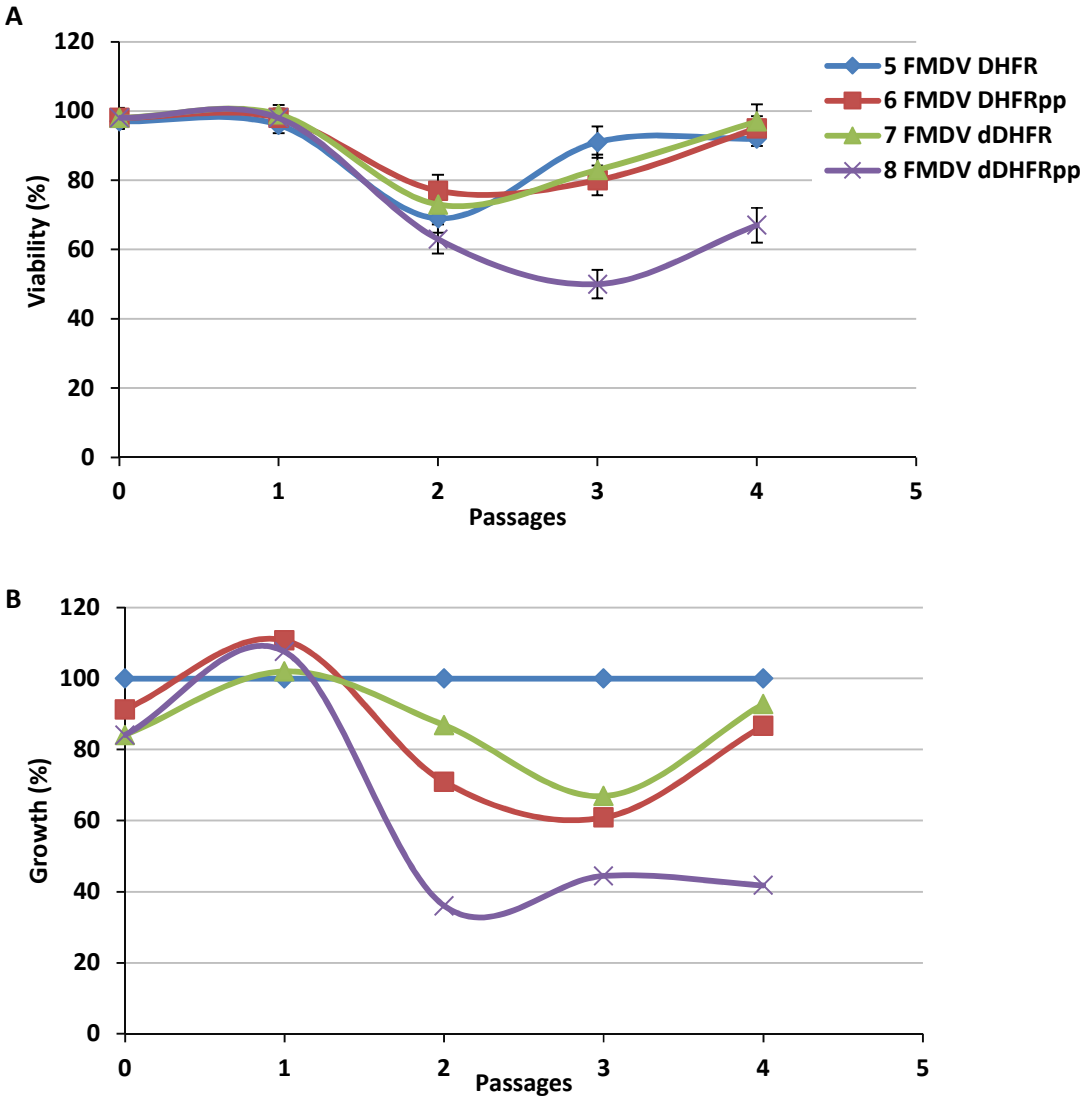
Figure 19. Cell pools productivity pre- and post- 500nM MTX treatment: AMH protein expression of cell pools transfected with engineered PB transposons affecting DHFR expression pathway pre- and post- 500 nM MTX treatment. Supernatants were collected after 4 days of cell culture and quantified with DiaSorin immunoassay prototype on Liaison platform. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

3.1.1 Gene amplification process: second attempt

In accordance with previous evaluations, FMDV pools were thawed at an early passage and treated again with 250 nM and 50 nM MTX concentrations. Cell growth and viability were evaluated at each passage as analyzed after the 500 nM MTX treatment (Figure 20). The growth rate was normalized on the values obtained from the pool 5 FMDV DHFR that functioned as control.

Regarding the 50 nM MTX treatment, results on viability percentage showed that the pools tolerated well this treatment, including the pool 8 FMDV dDHFRpp, even if a very slow recover was observed starting from the passage 3 (Figure 20A). The growth percentage analysis showed that pool with the highest constrains (8 FMDV dDHFRpp) was more impaired in growth than the other pools, which, on the contrary, reached a good growth rate at last passage (Figure 20B).

Regarding the 250 nM MTX treatment, results demonstrated that the pools survived also at this MTX concentration, but the viability percentages showed that the pools had high difficulties to survive (Figure 20C). The growth percentage analysis has an analogous profile to the 50 nM MTX treatment (Figure 20D).



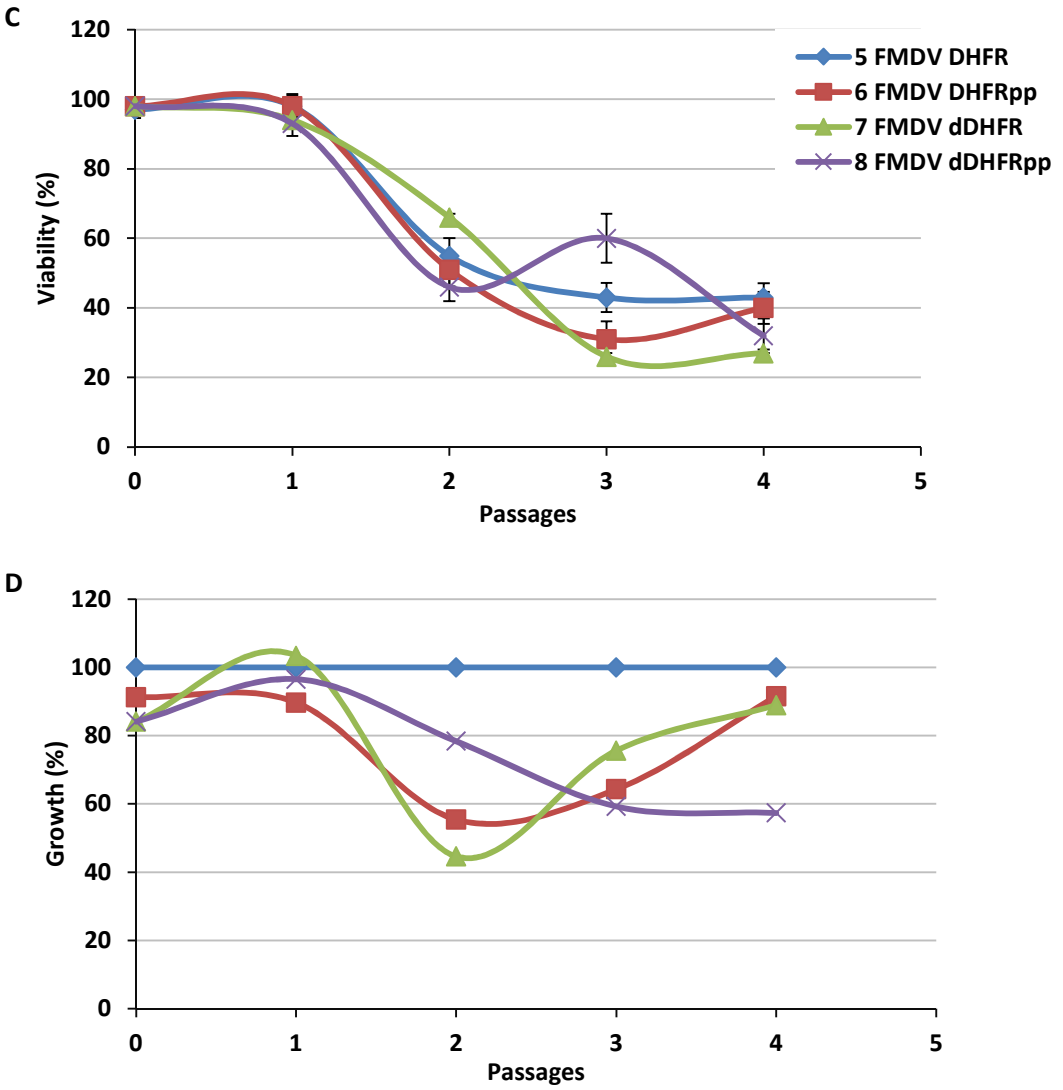


Figure 20. Growth curves of cell pools during 50nM and 250nM MTX treatment: Growth curves of cell pools transfected with engineered PB vectors affecting DHFR expression pathway. Cells were cultured for 5 passages. At passage 1, 50 nM or 250 nM MTX treatment was started and maintained until passage 5. At each passage, cell number and percentage of viability were evaluated. Percentage of viability and percentage of growth for each cell pool at each passage during 50 nM (A, B) and 250 nM (C, D) MTX treatment. Values were normalized on that obtained from cell pool 5 FMDV DHFR that functioned as control.

The protein expression was quantified on the collected supernatants with immunoassay prototype on Liaison platform (Figure 21). Positive results were obtained after 50 nM MTX treatment with an increase in protein production of about 2-3 fold. Anyway, this protein production increase was less than our expectation based on data reported in literature.

The AMH protein quantification obtained from the 250 nM MTX treatment were not reliable due to the viability and growth rate impairment of pools (data not shown).

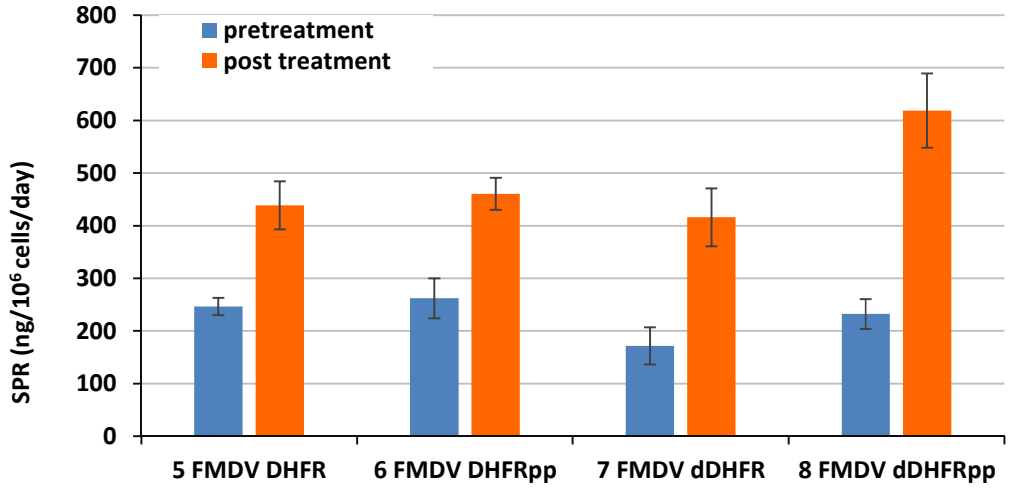


Figure 21. Cell pools productivity pre- and post- 50 nM MTX treatment: AMH protein expression of cell pools transfected with engineered PB vectors affecting DHFR expression pathway pre- and post- 50 nM MTX treatment. Supernatants were collected after 4 days of cell culture and quantified with DiaSorin immunoassay prototype on Liaison platform. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

3.2 Epigenetic regulatory elements

Despite the improvement in production scale-up and the implementation of strategies targeting the DHFR expression pathway, our aim to hugely enhance the AMH protein yield was not achieved. For this reason, we decided to change target. The protein expression pathway in mammalian cells is a finely controlled system based on transcription, post-transcriptional events and translational efficiency. Hence, we shifted the focus on epigenetic regulatory elements in order to force the transcriptional efficiency. We turned to a service company, Oxford Genetics Ltd, to produce customized vectors with the insertion of different epigenetic regulatory elements. Our aim was to set up a proper system to exploit the DHFR knockout CHO cell line potentiality. The strategy was based on the insertion of cis-acting DNA elements in the transposon in order to reduce or eliminate epigenetic modifications that negatively affect the transgene expression also in the case of transgene integration into an area of closed heterochromatin. The basal cloning vector used was pSF_OG2 with the insertion of the *piggyBac* inverted terminal repeat sequences (ITRs) in order to exploit the *piggyBac* transposon system.

The DNA elements chosen were:

- Insulator cHS4: the DNase I hypersensitive site 4 of the chicken β-globin locus control region (cHS4) shows classical insulator properties. It confers stability of the transgene expression during time. This element was present also in the PB transposon vector used for our previous experiments.
- UCOE: Ubiquitous chromatin opening element (UCOE) is a promoter region of ubiquitously expressed housekeeping genes. This element is a methylation free CpG island. It is demonstrated that the UCOE maintains a favorable chromatin conformation and protects the transgene from silencing. We selected a 1.5 kb sized A2UCOE from the human HNRPA2B1-CBX3 housekeeping gene.
- MAR: Matrix associated region (MAR) is an element that binds the nuclear matrix. This element realizes a structural order in the chromatin protecting against position effects and enhancing the gene expression. We selected the human MAR 1-68.

The backbone of the expression cassette for all pSF transposons contained the CMV (Cytomegalovirus) and the weak PGK (murine phosphoglycerate kinase 1) promoter, the former controlled the gene of interest and the latter the DHFR gene (Figure 22). CMV promoter was chosen because it is reported that epigenetic regulatory elements such as UCOEs work acting on this promoter (Kwaks and Otte, 2006). Moreover, in order to strengthen the weakening effect on the DHFR expression, we cloned the DHFR gene under a distinct and weak promoter (PGK promoter) instead of inserting it after an IRES sequence, as in PB transposon. In addition, a PEST and an ARE (Adenylate-Uridylate Rich Element) sequence were inserted downstream of the expression cassette. We added only a PEST sequence on the evidence of the results obtained in the previous experiments where two PEST sequences in tandem did not give an advantage. ARE sequences target mRNAs for a rapid degradation. We did not explore this sequence in our previous PB transposon because its insertion in combination with the IRES would have negatively affected not only the DHFR mRNA but also the AMH one. In pSF transposons, the DHFR gene was under a different promoter and hence it was consistent to insert the ARE sequence. Indeed, it acted only on the DHFR expression. Each pSF transposon contained an epigenetic regulatory element, except for pSF and psF_U2 where two cHS4 and two UCOE sequences, respectively, were incorporated: one at 5' and the other at 3' end of the expression cassette. In order to make a comparison with pSF_AMH, the transposon used for previous experiments (Figure 5A), from now on referred as PB_AMH, was also analyzed. It contained two cHS4 sequences.

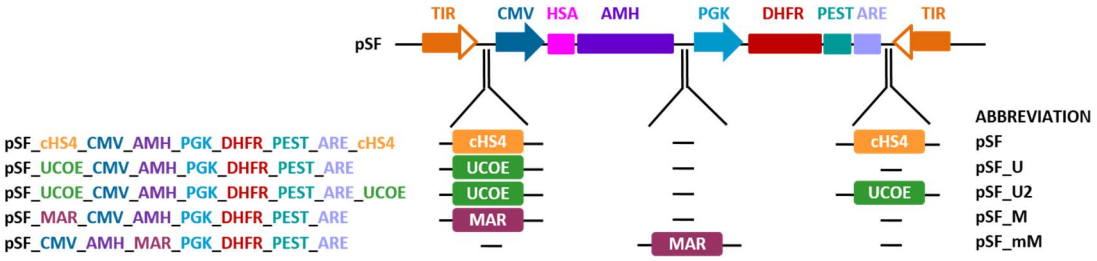


Figure 22. pSF transposon engineering: schematic representation of engineered pSF vectors containing different epigenetic regulatory elements. Description of epigenetic regulatory elements inserted in each pSF transposon and their position in the expression cassette. Abbreviations used in the paragraph are also reported. pSF: transposon; TIR: terminal inverted repeats; CMV: Cytomegalovirus promoter; PGK: murine phosphoglycerate kinase 1 promoter; HSA: human serum albumin signal peptide; AMH: anti-Müllerian hormone; DHFR: dihydrofolate reductase; ARE: Adenylate-Uridylate Rich Element; UCOE: Ubiquitous chromatin opening element; cHS4: DNase I hypersensitive site 4 of the chicken β-globin locus control region; MAR: Matrix associated region.

Therefore, eight different stable transfections were performed on DHFR knockout cells. Six of them were carried out using the constructs described before (5 pSF transposons and the PB one) in co-transfections with the plasmid carrying the gene encoding for the transposase in a molar ratio transposon/transposase of 2.5:1. As controls, other two conditions were evaluated co-transfecting pSF_AMH and PB_AMH with a "mock" plasmid, pBluescript vector (pBlue), to verify if the insertions in the cell genome occurred by transposition or classical integration. From now on, we referred to the pool derived from the transfection with pSF transposon as pSF pool, that from pSF_U transposon as pSF_U pool and so on. Therefore, pSF_U2 pool, pSF_M pool, pSF_mM pool and PB pool correspond to the pools obtained from the transfections with pSF_U2, pSF_M, pSF_mM and PB transposons, respectively. After the transfections, the cells were maintained in MW6 plates in a static incubator for 8 days; a selective medium without HT was used. In this phase, the cells have never reached a satisfying viability, not exceeding 60% (Figure 23A). Despite the low viability, we decided to transfer all pools in flasks in a shaking incubator with 6.5 ml of medium. The cells were cultured in flasks for other 4 passages progressively increasing volume until 10 ml. After the transfer, the cells started to significantly grow, improving their viability up to 99% in 3 passages. Only pSF_U2 pool showed a greater impairment to stabilize its viability compared to the other pools. Regarding cell growth (Figure 23B) after a slight drop when the cells were transferred, all cell pools reached a good growth rate in 2 weeks from the transfection. An exception occurred in pSF_U2 pool whose cell density increased but it never reached 1 x 10^6 cells, neither after 20 days from the transfection.

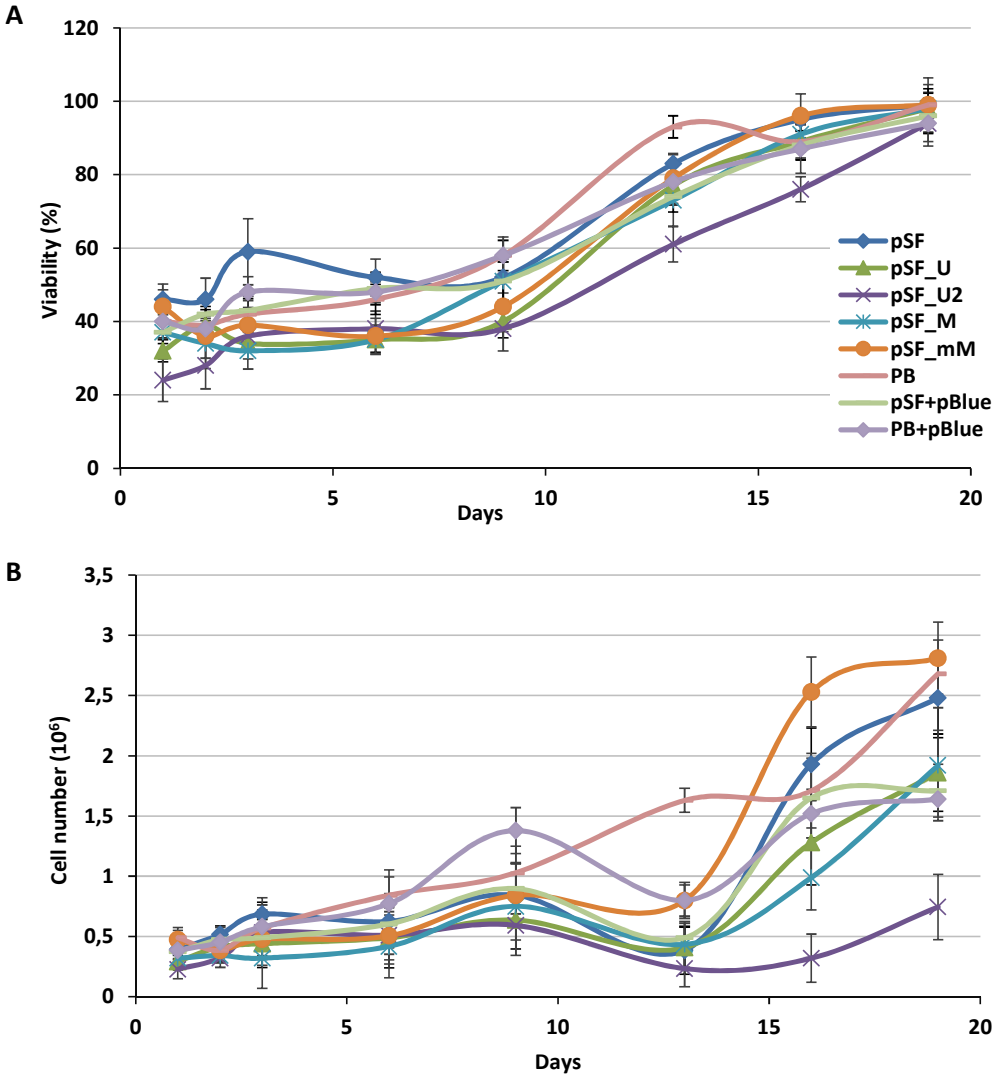


Figure 23. Growth curves of cell pools derived from engineered pSF transposons: Cell pools were monitored for 19 days after transfection. At each passage percentage of viability (A) and cell number (B) were evaluated.

After 4 days of culture, the cell supernatants were collected for the quantification on Liaison by using the AMH prototype immunoassay (Figure 24). Results demonstrated that the pSF transposon integration occurred by transposition because the AMH expression of the pools co-transfected with transposases was higher than that obtained by the pSF "mock" transfection. The same happened for PB transposon, as also demonstrated in previous experiments. Moreover, the pool transfected with PB transposon showed a higher AMH expression than that transfected with the basal pSF vector. Good results were obtained in pSF_U and pSF_U2 pools in which high protein expression levels were reached. A 6-fold and a 3-fold increase in AMH yield compared to the basal pSF pool and PB pool, respectively, demonstrated the potent impact of UCOE sequences. Unexpectedly, no effect was obtained in the pools transfected with constructs containing MAR sequence.

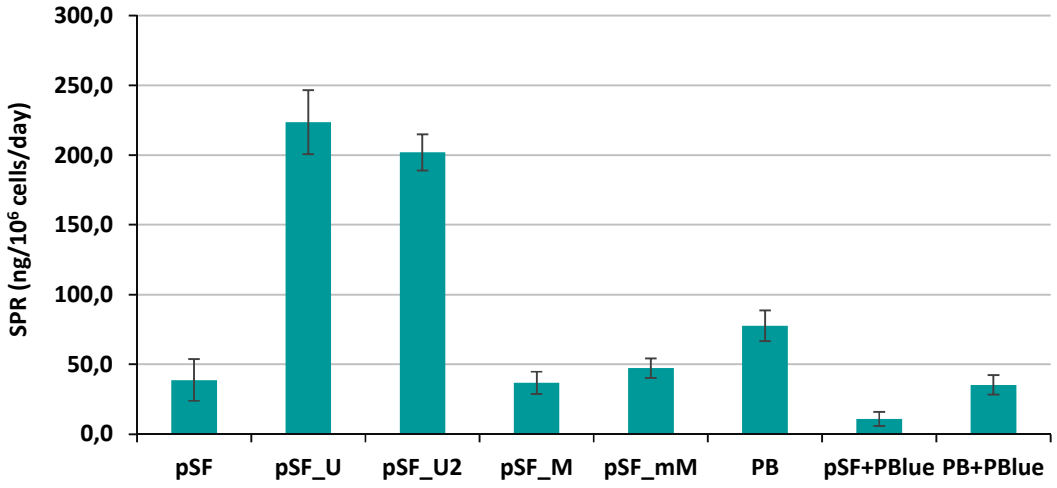


Figure 24. Productivity of cell pools derived from engineered pSF transposons: AMH protein expression of cell pools transfected with engineered pSF transposons. Supernatants were collected after 4 days of cell culture and quantified with DiaSorin immunoassay prototype on Liaison platform. pSF_AMH and PB_AMH were also co-transfected with the "mock" plasmid pBluescript vector (pBlue) as controls. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

3.2.1 Gene amplification process

Given these results, we decided to explore the gene amplification with the MTX treatment in pSF and PB pools. Therefore, the cell pools were treated with 50 nM MTX concentration. We chose to test only the 50 nM MTX treatment to avoid an excessive impact on the pools, as happened for FMDV pools. Indeed, as explained before, the sequences that impair the DHFR expression pathway were also present in pSF transposons used for the transfections. After two weeks, the MTX treatment was stopped and the supernatants from the cell pools were collected. All cell pools survived after the 50 nM MTX treatment without any difficulty. At each passage, the cell viability was evaluated, all cell pools maintained a viability ranging from 96% to 99% over time with or without MTX treatment (data not shown). The growth rate was also analyzed, the values were normalized to those obtained from PB pool that acted as control. Results demonstrated that the growth percentage of not treated cell pools was similar to that of the control over time (Figure 25A). The growth percentage of cell pools during the 50 nM MTX treatment showed a slight drop after the first passage in MTX treatment (passage 1) compared to the control, but no significant impairment in the growth rate was observed, except for pSF_M pool in which the growth

percentage lowered until 54% (Figure 25B). However, pSF_M pool showed a very unstable growth profile, in contrast to that obtained in pSF_mM pool. Note that the pSF pool growth percentage was similar to that of the control, moreover, pSF_U and pSF_U2 pools maintained a comparable growth profile over time. The AMH protein expression was quantified on the collected supernatants with immunoassay prototype on Liaison platform (Figure 25C). Unexpectedly, no effect was observed after MTX treatment. It is important to highlight that the results obtained in the pool transfected with PB transposon correlated with the experiments previously performed (see paragraph 2.2). The cell pools derived from pSF transposons probably resulted more resistant to the MTX effects than FMDV pools. Therefore, we decided to strengthen the gene amplification performing a 250 nM MTX treatment on the cell pool derived from the transfection with pSF_U transposon (Figure 26). This cell pool was selected because it was the most promising compared to the others. Indeed, it allowed to obtain the highest AMH protein expression and it maintained a very good growth curve profile from the beginning. These are promising properties of an easy handling cell pool with great performances. Besides, pSF and PB pools were treated at 250 nM MTX as controls. The 250 nM MTX treatment was performed andthe cell viability was evaluated at each passage. All pools maintained a viability ranging from 96% to 99% over time, except for pSF_U pool at passage 2 and 3, where a slight drop in viability (94%) was observed, probably because it suffered the MTX treatment (data not shown). The growth rate at each passage was also analyzed and values were normalized to those obtained from PB pool that acted as control (Figure 26A). The profile of pSF pool growth percentage oscillated around that of the control, as happened for MTX treatment at 50 nM concentration. In contrast, an impairment in the growth was shown in pSF_U pool in particular from passage 1 to 3 and then a recovery was observed. After the MTX treatment, the AMH protein expression was quantified on the collected supernatants with our immunoassay prototype on Liaison platform. Data demonstrated that the protein expression was not enhanced in pSF_U pool even after the treatment with higher MTX concentration (Figure 26B). In contrast, an increase in the protein production was observed in pSF and PB pools even if it was slight. Also in this case, the results obtained from PB pool were consistent with previous data (see paragraph 2.2).

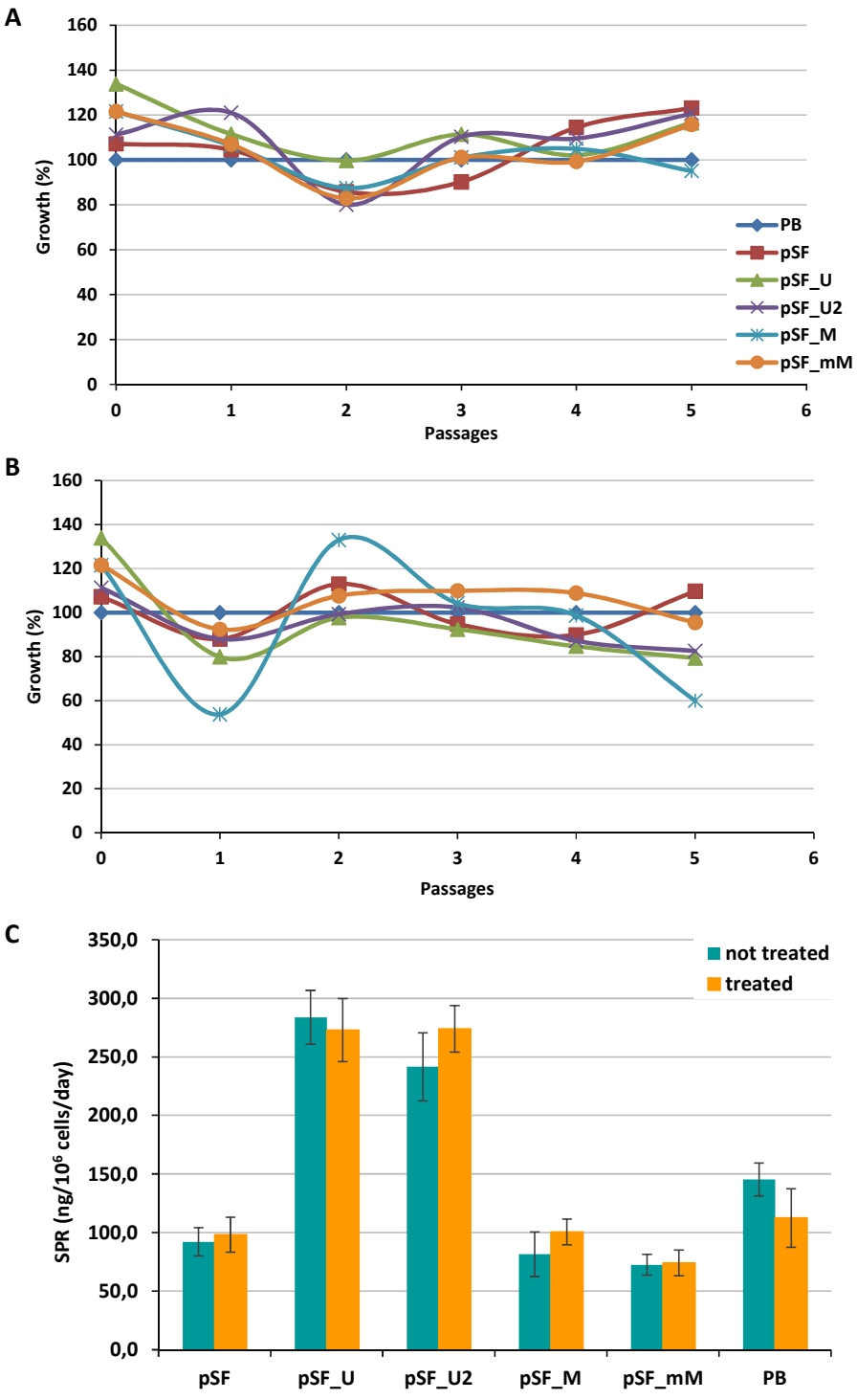


Figure 25. Growth percentage and productivity of cell pools derived from engineered pSF transposons with or without 50 nM MTX treatment: Cells were cultured for 5 passages. At passage 1, 50 nM MTX treatment was started and maintained until passage 5. At each passage, cell number and percentage of viability were evaluated. A, B. Percentage of growth for each cell pool over passages without (A) and with 50 nM MTX treatment (B). Values were normalized on that obtained from PB pool that acted as control. C. AMH protein expression of cell pools treated with 50 nM MTX concentration compared to not treated ones. Supernatants were collected after 4 days of cell culture and quantified with DiaSorin immunoassay prototype on Liaison platform. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

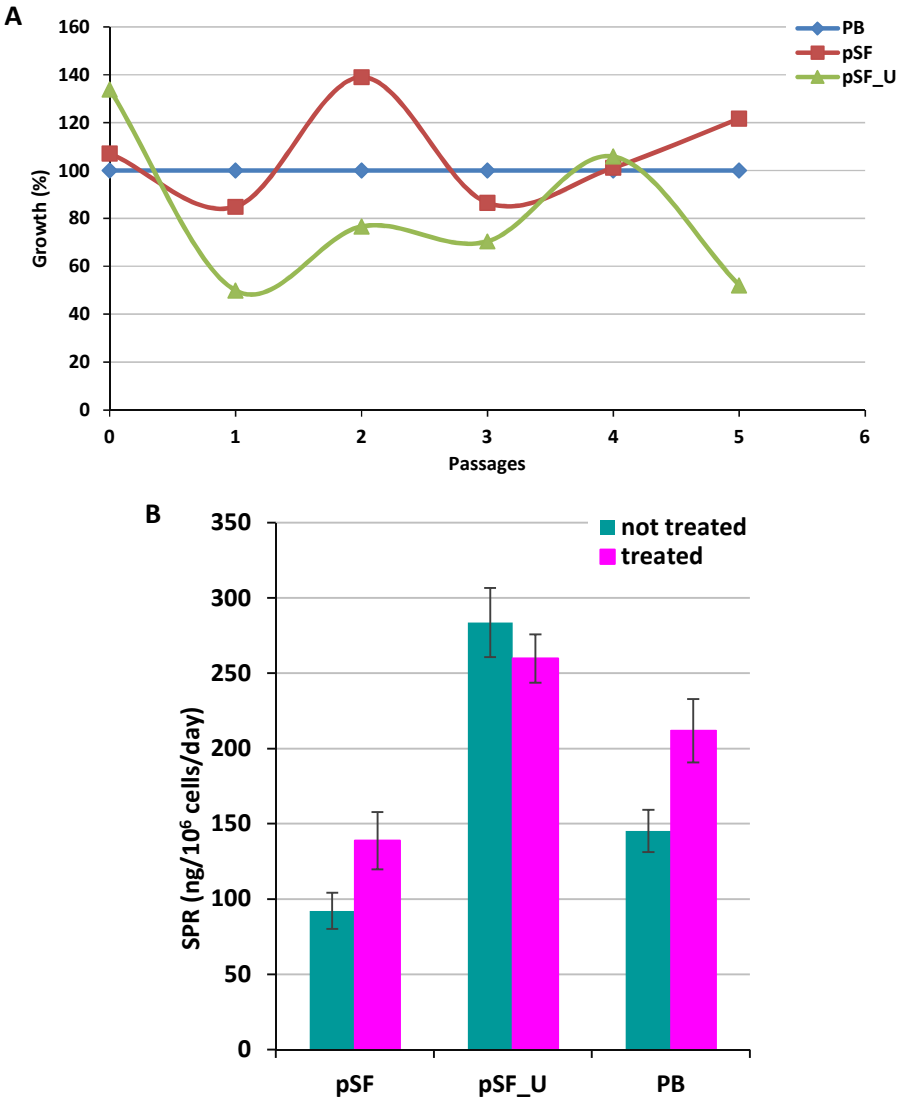


Figure 26. Growth percentage and productivity of cell pools derived from engineered pSF transposons with or without 250 nM MTX treatment: Growth curves of selected cell pools transfected with engineered pSF transposons. Cells were cultured for 5 passages. At passage 1, 250 nM MTX treatment was started and maintained until passage 5. At each passage, cell number and percentage of viability were evaluated. A. Percentage of growth for each cell pool over passages with 250 nM MTX treatment. Values were normalized on that obtained from PB pool that acted as control. B. AMH protein expression of selected cell pools treated with 250 nM MTX concentration compared to not treated ones. Supernatants were collected after 4 days of cell culture and quantified with DiaSorin immunoassay prototype on Liaison platform. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

3.2.2 High producing clones frequency

Despite the ineffectiveness of the MTX treatment on the AMH protein expression in pSF_U pool, we decided to proceed in depth analysis due to its promising performance. Therefore, we performed a cloning step on not treated pSF_U pool in order to isolate high-producing clones. Our aim was to compare their productivity to that of the clones obtained with the PB transposon system in our previous experiments and more importantly to study if the UCOE can assure the expression and genetic stability over time avoiding problems encountered with PB transposon. We decided to clone also pSF pool in order to highlight the advantages of the UCOE and the difference between pSF and PB transposons. A single cell cloning by limiting dilution was performed on the selected pools. Single clones derived from each stable pool were isolated and their productivity was evaluated (Figure 27). The cell number in each well was calculated by CCK-8 Dojindo assay. The AMH protein expression was quantified with our automated prototype immunoassay on Liaison platform. The quantitative values were normalized to both clone cell number and days of culture, obtaining a final production rate value (Specific Productivity Rate or SPR) expressed as nanograms of protein produced by 10^6 cells per day (ng/10^6 cells/day).

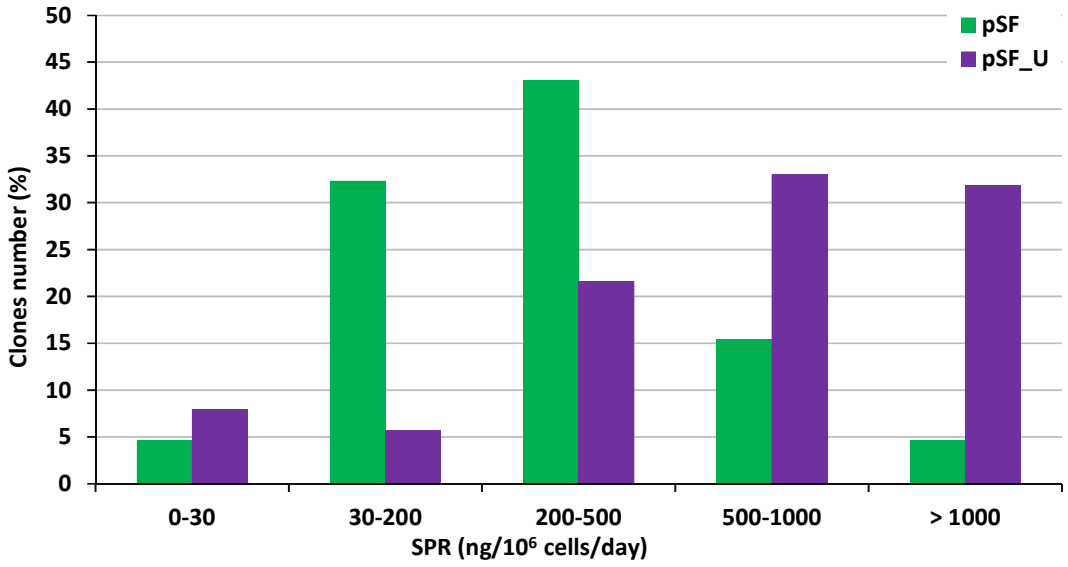


Figure 27. Productivity of clones: Quantitative screening of clones generated from pSF or pSF_U stable pools. Clones were classified based on AMH protein expression level. Cell supernatants were analyzed with an automated prototype assay for detection of the AMH protein. Values were normalized based on clone cell number (Dojindo cell counting assay) and days of culture. Data were reported as percentage of clones classified into arbitrary categories of production. SPR (Specific Productivity Rate) indicates nanograms of protein produced by a million of cells in a day.

Results obtained from the quantitative screening were plotted arbitrarily defining production categories and estimating the percentage of clones for each experimental condition that fell in each category. Data showed that pSF transposon generated more than 75% of clones with a low-medium productivity, thus only 20% of clones fell into the high and very high category. On the contrary, the productivity of the clones derived from pSF_U transposon was completely shifted towards the highest producing category. Indeed, pSF_U transposon showed a productivity profile that was exactly the opposite than that obtained with pSF transposon because it allowed to obtain 65% of clones with the high and very high productivity, while only 27% of them had a low-medium one. These results demonstrated the efficacy of the UCOE sequence to enhance protein expression.

3.2.3 Clones productivity and stability over time

Once selected the high-producing clones, it is necessary to evaluate clones productivity and stability over time in order to guarantee a consistent reproducibility in the protein production day by day, as described in paragraph 2.4. Therefore, four high-producing clones from both pSF and pSF_U stable pools were selected and maintained in culture for 3 months to evaluate the expression and genetic stability. The selected clones isolated from pSF and pSF_U stable pools from now on referred to as pSF and pSF_U clones, respectively, are listed in Table 3.

pSF clones	pSF_U clones
83A.8	83B.130
83A.9	83B.135
83A.20	83B.136
83A.54	83B.161

Table 3. Selected clones: schematic representation of the selected clones. pSF: transposon; pSF_U: transposon carrying UCOE sequence.

The selected clones were cultured for 24 passages and a qPCR analysis on genomic DNA extracted from them was realized to evaluate the copy number of transgene integration. Samples were collected after 1, 13 and 24 passages in order to study the genetic stability over time (Figure 28A). The AMH expression was quantified at passage S1, S13 and S24 for each clone. The quantification was performed on supernatants obtained from 4 days cell culture with the automated prototype immunoassay on Liaison platform (Figure 28B).

Also in this case, only a few clones were analyzed, but data disclosed some substantial evidence. Regarding the genetic analysis, we observed that pSF clones showed an average of 5-fold more transgene integrated copies than pSF_U ones, except for the clone 83A.8 that was similar. The pSF_U clones integrated copies resulted also slightly less (2-fold in average) than those obtained by using PB transposon. The transposition was probably reduced in pSF_U transposon due to its higher size compared to pSF one. The genetic stability over time was maintained in all clones with a remarkably consistency in pSF_U ones. Surprisingly, an increase in copy number was shown in the clone 83A.20 during passages. We speculated that a possible gene amplification could have occurred over time due to the DHFR impairments caused by the weak DHFR promoter, ARE and PEST sequences. If we shifted the focus on the productivity, we observed an exactly opposite situation where pSF_U clones showed their supremacy in productivity with a 4-fold increment, in average, compared to pSF and also to PB ones. This corresponded to our expectation based on the enhanced frequency of high producing clones, previously observed, compared to pSF ones. Unexpectedly, in pSF clones, in particular in the clones 83A.9 and 83A.20, no correlation between transgene integrated copies and productivity was observed. This means that a silencing phenomenon occurred in these clones, which is in contrast with the integration in actively transcribed regions mediated by the transposase. A possible explanation could be due to the different transposon vector used. Indeed, PB transposon carried the IRES sequence between the AMH and DHFR genes, on the contrary pSF had the two genes under different promoters. This could allow a more accessibility and sensibility of the AMH gene to silencing effects. We also hypothesized that the promoter interference could have occurred which displayed perturbation of one transcription unit by another (Curtin et al., 2008; Eszterhas et al., 2002; Uchida et al., 2013). However, the expression stability was maintained over time in all clones.

Our results supported the potentiality of the UCOE sequence to substantially enhance the clone protein productivity sustaining the expression and genetic stability over time. pSF_U clones satisfied our needs. Therefore, the highest producing clone among pSF_U ones, i.e. the clone 83B.135, was selected to proceed with a further characterization.

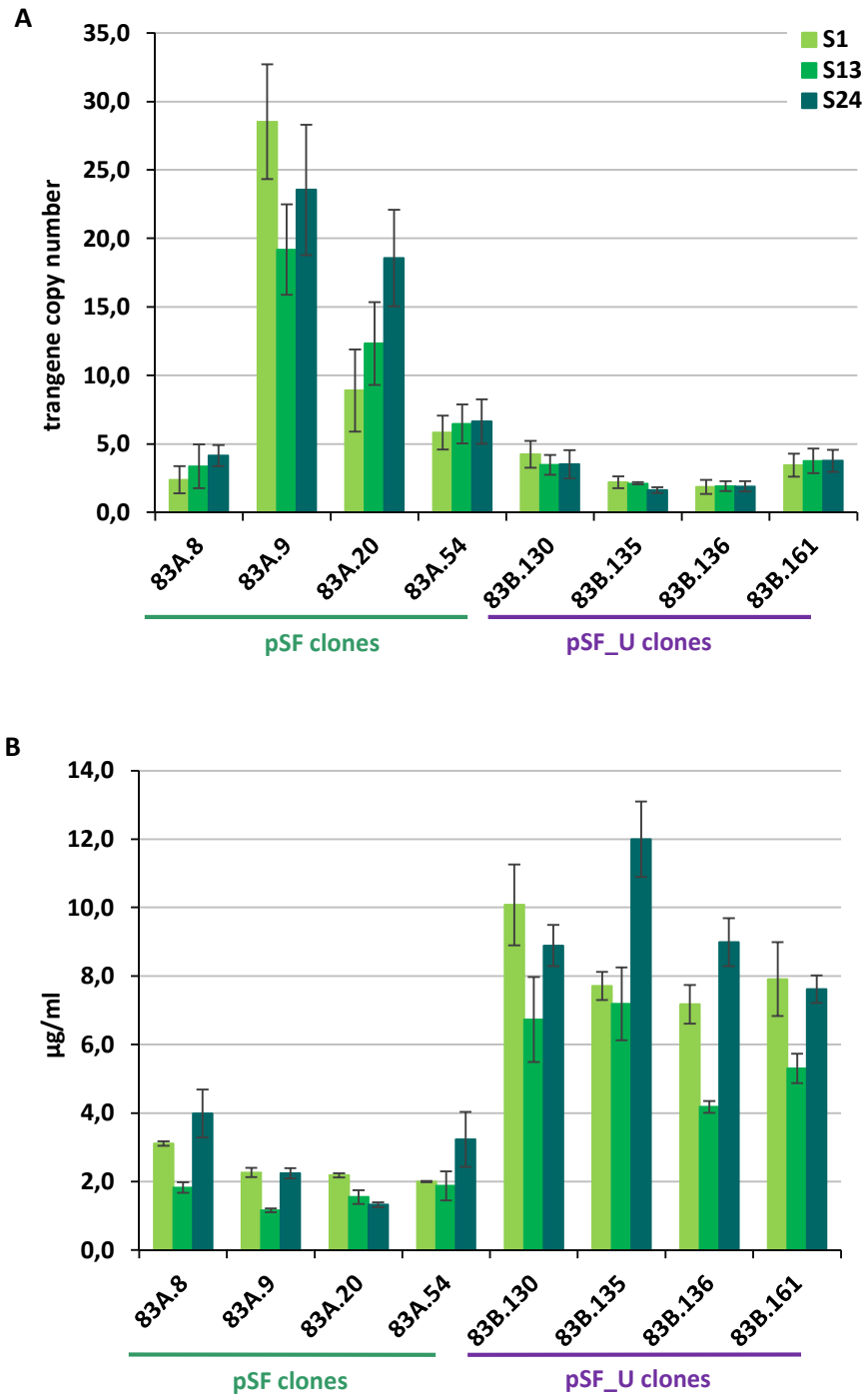


Figure 28. Genetic and expression stability of selected clones: A. Genetic stability of selected clones over time. qPCR analysis on genomic DNA extracted from selected clones at passages 1, 13 and 24 (S1, S13, S24). B. Productivity of selected clones over time. AMH protein expression quantification of selected clones with the automated prototype immunoassay on Liaison platform from 4 days cell culture supernatants collected at passages 1, 13 and 24 (S1, S13, S24).

Part 4: Purification

Since the downstream process is vital to gaining a purified recombinant protein for industrial purposes, we inserted two tag sequences in the vector in order to increase the chance to getting a robust purification process. Therefore, we compared two purification procedures. One exploits a tag, the Strep-tag, that we have not tested before and the other uses the His-tag that is widely applied in our lab. At the beginning, we explored these two methods on several batch cultures of the clone 49E.259, even if it was not our final choice. The batch cultures were performed in 1-liter flask with a cell density of 0.4 x 10^6 cells/ml at 125 rpm orbital shaking in a total volume of 100 ml in mild temperature condition. The AMH protein quantification on the supernatants collected at day 7^{th} of cell culture was about 20 µg/ml.

4.1 Purification process based on Strep-tag

Purification protocol using Strep-tag was based on two steps:

- **Step one.** *Tangential Flow Filtration (TFF): protein concentration and buffer exchange*

The supernatant was filtrated (0.22 µm pores diameter) and then the TFF with a PES (polyethersulfone) cassette was performed. During this step the buffer was extensively exchanged and the initial volume was 3-fold reduced.

- **Step two.** *Protein purification with affinity chromatography: StrepTactin/Strep-tag*

The samples obtained from the step one was loaded on a pre-charged StrepTactin resin (StrepTactin Sepharose 50% suspension - IBA GmbH). The StrepTactin resin interacted with high affinity to the Strep-tag present on the N-terminal portion of the AMH protein. The protein was eluted with a step of desthiobiotin because this molecule competes with the target protein to bind the StrepTactin resin. The fractions containing the target protein were identified combining the SDS-PAGE analysis (Figure 29A) and the chromatography profile (data not shown). The fractions were then mixed to obtain the final preparation. The protein concentration was evaluated with the automated immunoassay prototype on Liaison platform. The purified protein was then stored at - 30°C. The SDS-PAGE analysis showed that all contaminant proteins were thrown away in the flow through. Therefore, the target protein was eluted with a high purity degree. As highlighted with the red rectangle, the fractions from 5 to 9 were mixed, the pool obtained represented the final protein preparation.

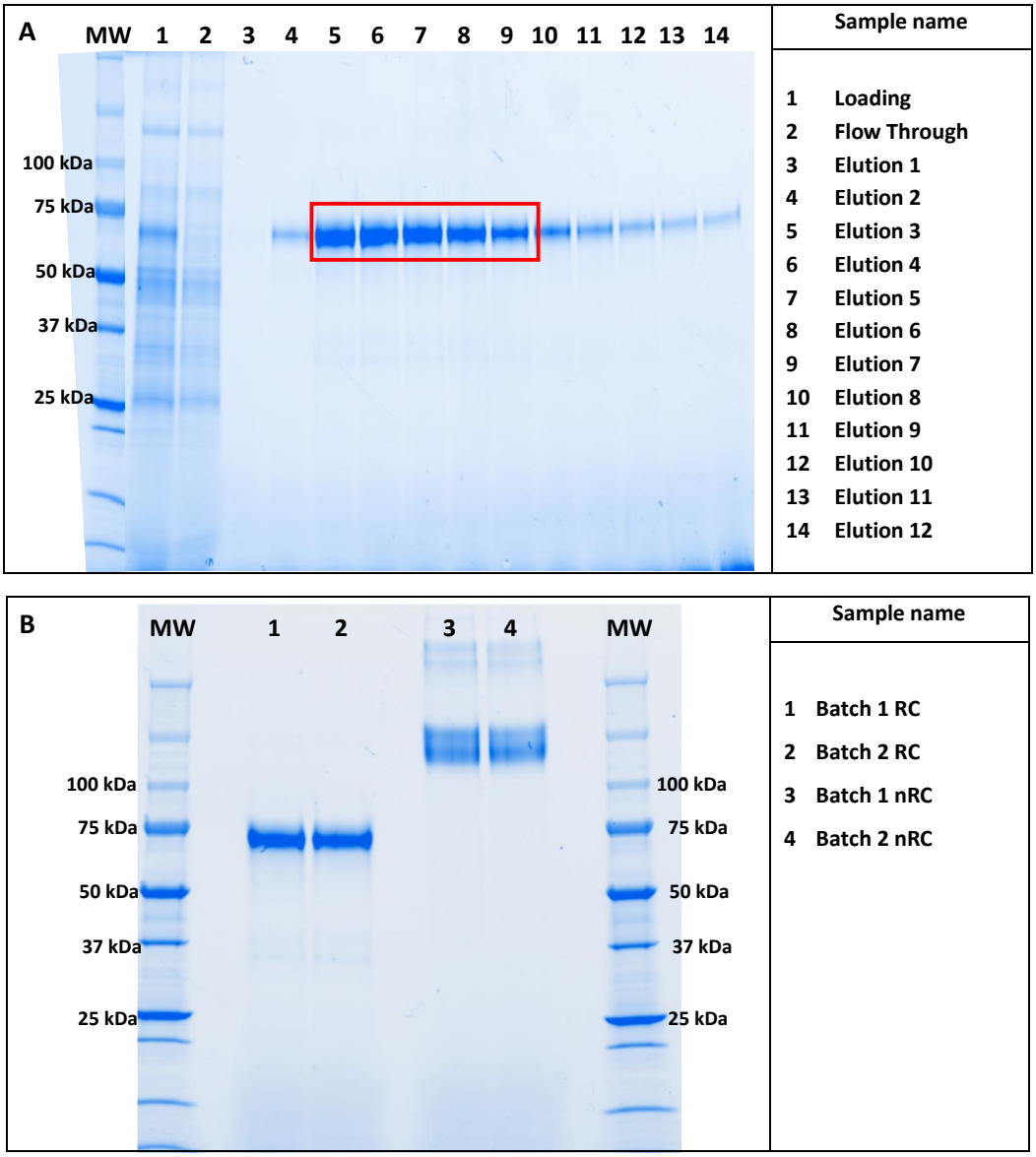


Figure 29. Purification process evaluation exploiting Strep-tag and protein characterization: A. SDS-PAGE analysis on protein affinity chromatography for a batch. Red rectangle represented the eluted fractions that contained the target protein. B. SDS-PAGE analysis for protein characterization on two batches (batch 1 and batch 2). RC: reducing condition; nRC: non reducing condition.

The final concentration was measured with our immunoassay prototype on Liaison platform and resulted in a good yield of about 6 mg/L on average. The last two batches were also characterized by the SDS-PAGE analysis (Figure 29B). In reducing condition, the two batches reached a high purity degree (≥ 95%); in non-reducing condition, the protein dimer at about 140 kDa was detected in both batches with only minimal aggregation forms at higher molecular weights.

4.2 Purification process based on His-tag

In order to compare the efficiency of the Strep-tag and the His-tag in the purification process, we produced other batches to explore the Immobilized-Metal Affinity Chromatography (IMAC) procedure exploiting the His-tag.

Purification protocol using the His-tag was based on two steps:

- **Step one.** *Tangential Flow Filtration (TFF): protein concentration and buffer exchange*

The supernatant was filtrated (0.22 µm) and then the TFF with a PES (polyethersulfone) cassette was performed. During this step the buffer was extensively exchanged and the initial volume was 3-fold reduced, exactly the same step performed for the Strep-tag/StrepTactin purification.

- **Step two.** *Protein purification with IMAC*

The samples obtained from the step one was loaded on two IMAC columns (HisTrap FF crude 5ml linked in series). The target protein was eluted by a linear gradient of imidazole. The fractions containing the target protein were identified combining the SDS-PAGE analysis (Figure 30) and the chromatography profile (data not shown).

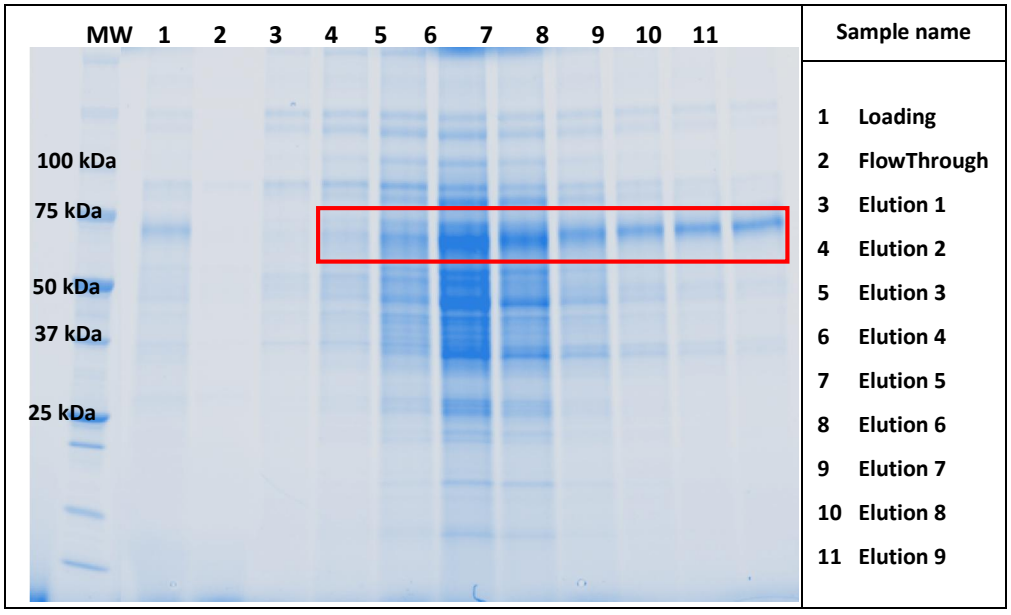


Figure 30. Purification process evaluation exploiting His-tag: A. SDS-PAGE analysis on protein affinity chromatography for a batch. Red rectangle represented the eluted fractions that contained the target protein.

The SDS-PAGE analysis showed that the target protein (red rectangle) was eluted with all contaminant proteins. The IMAC purification was not efficient and it was not a valid alternative to the purification exploiting the StrepTactin resin.

Given these results, we selected the Strep-tag/StrepTactin procedure to purified the AMH protein on the supernatants collected from batch cultures. The process showed a good reproducibility and allowed to obtain a final preparation with a purity degree higher than 95% and a final protein yield of about 30-35%. The protein amount obtained abundantly satisfied our productive demand.

Part 5: Characterization of the best productive clonal cell line

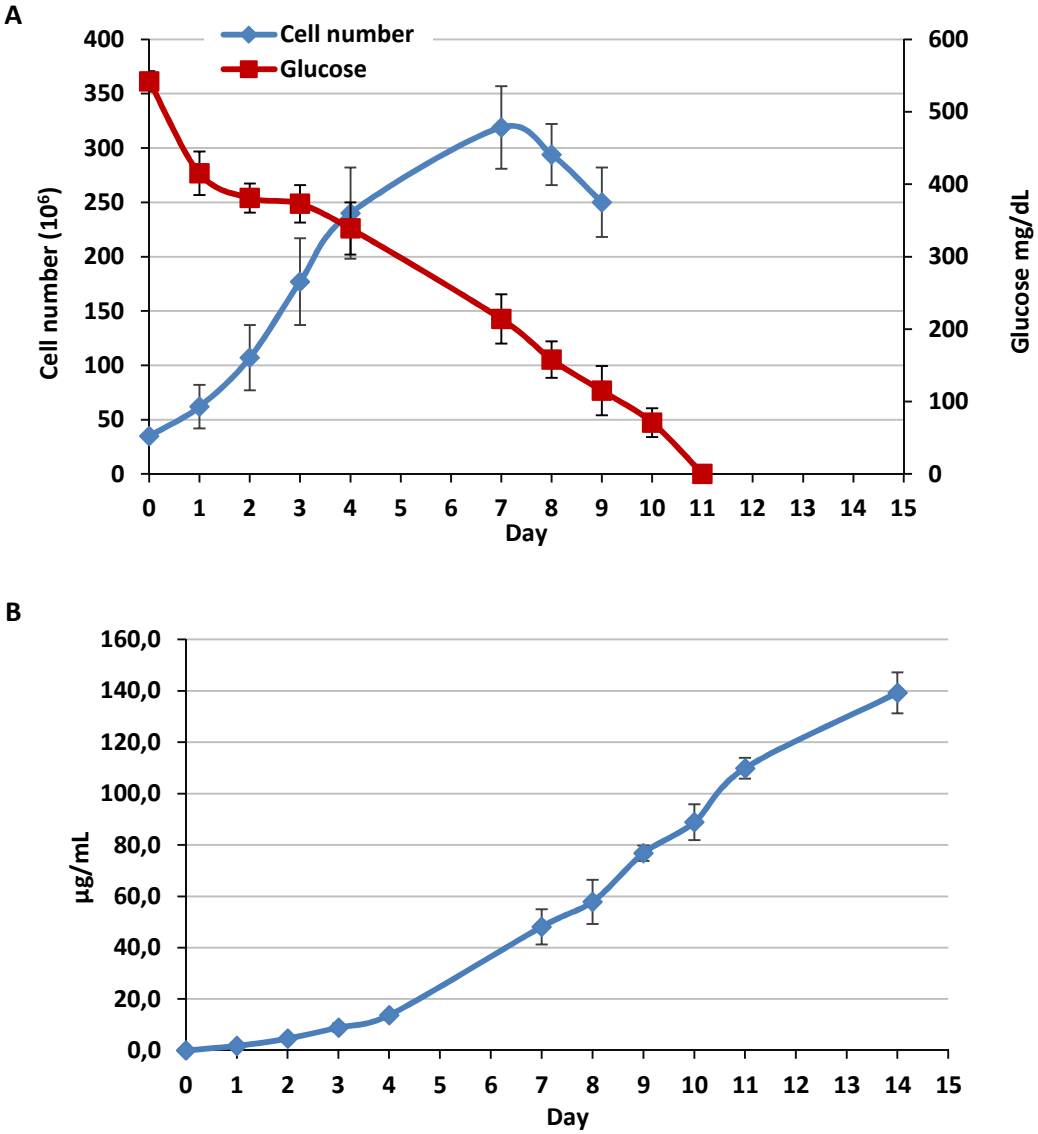
The selected clone parameters were evaluated to precisely define the clone characteristics. Accurate features including specific growth rate (μ), specific protein productivity (q), viability, duplication time (DT) and glucose consumption are required to establish the clone performances in standard condition. It is strictly fundamental that the profile of selected productive clonal cell line is preserved over time. Whenever the clone is handled, it is necessary to check that each parameter falls into the described ranges to guarantee reproducibility and quality of the protein product.

5.1 Parameters of growth and production

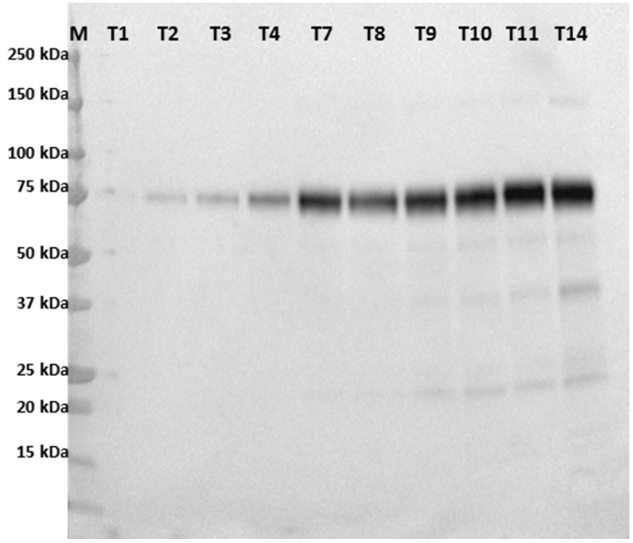
The parameters of the selected clone 83B.135 were established. Specific growth rate (μ) represents the number of doublings that occur per unit of time; duplication time (DT) indicates the period occurred for cell duplication. These parameters were calculated during the clone maintenance. Specific growth rate (μ) and duplication time (DT) were estimated in 0,029 h^{-1} and 23 hours, respectively.

The batch culture production was performed following the mild hypothermia optimized protocol. The cells were seeded with a density of 0.4 x 10^6 cells/ml in 1-liter flask at 125 rpm orbital shaking in a total volume of 100 ml. The temperature was shifted to 32°C the day after the inoculum. The cells reached a peak of growth at day 7^{th} with 320 x 10^6 cells and a good viability (99%), exactly the same values obtained in production performed with the clone 49E.259. The cell number started to decrease from day 7^{th}. The counts were considered unreliable after day 9^{th} due to cell aggregation and cellular debris; therefore, those time points were not shown in the graph (Figure 31A). The cell viability was very high until 10^{th} day, ranging from 94% to 100%, and then rapidly decreased (data not shown). The glucose analysis showed a gradual nutrient consumption. The AMH protein was quantified with our prototype immunoassay on Liaison platform resulting in a great protein yield of 140 μg/ml at day 14^{th} of culture (Figure 31B). WB analysis confirmed the substantial AMH protein expression. The protein continued to be produced over days. However, starting from day 8, protein degradation was clearly visible in WB analysis and increased each passing day (Figure 31C). Based on cell viability, growth rate, glucose consumption, protein quantification and protein degradation pattern, the best time point for supernatant collection was day 7^{th} because harvesting after that day could impair purification process.

Specific growth rate (μ) and duplication time (DT) resulted 0.022 h^{-1} and 31 hours, respectively, confirming the reduction of the growth rate due to the mild temperature culture condition. The specific AMH protein productivity (q) was 5 pg/cell/day (Figure 31D).



C



D

Clone ID	μ (h^{-1})	DT (h)	q_{AMH} (pg/cell/day)
83B.135	0,029	23	5

Figure 31. Parameters of production: Growth curve of cell culture with the best productive clonal cell line (1-liter flask with a cell density of 0.4 x 10^6 cells/ml at 125 rpm orbital shaking in a total volume of 100 ml). At each time point, cell number, viability and glucose level were evaluated. A. Growth curve and glucose evaluation. Supernatant samples were collected at day 1, 2, 3, 4, 7, 8, 9, 10, 11 and 14. B. Curve of AMH protein quantification on supernatant samples at each collected time point. C. WB analysis of production immunodecorated with a commercial α-AMH antibody. D. Clone 83B.135 specific parameters; μ and DT referred to standard clone maintenance; μ: specific growth rate (h^{-1}); DT: duplication time (h); q_{AMH}: specific AMH protein productivity (pg/cell/day).

5.2 Purification

Since the excellent performance obtained in the purification process with the strep-tag, we purified the batch production of the selected clone using the Strep-Tactin resin. The fractions containing the target protein were identified combining the SDS-PAGE analysis (Figure 32A) and the chromatography profile (data not shown). The top fractions of the eluted peak were pooled together and represented the final protein preparation. Congruently with what obtained in the purification on the supernatant of the clone 49E.259, the SDS-PAGE analysis showed that all contaminant proteins were thrown away in the flow through allowing the target protein elution with a high purity degree. Moreover, the protein was eluted in less fractions (red rectangle) compared to the purification on the

clone 49E.259 supernatant, giving the advantage to have a more concentrated purified protein. The final concentration was measured with our immunoassay prototype on Liaison platform and it resulted in a good yield of about 15 mg/L. The final protein yield was 37%, also better than that obtained in the purification on the clone 49E.259 supernatant.

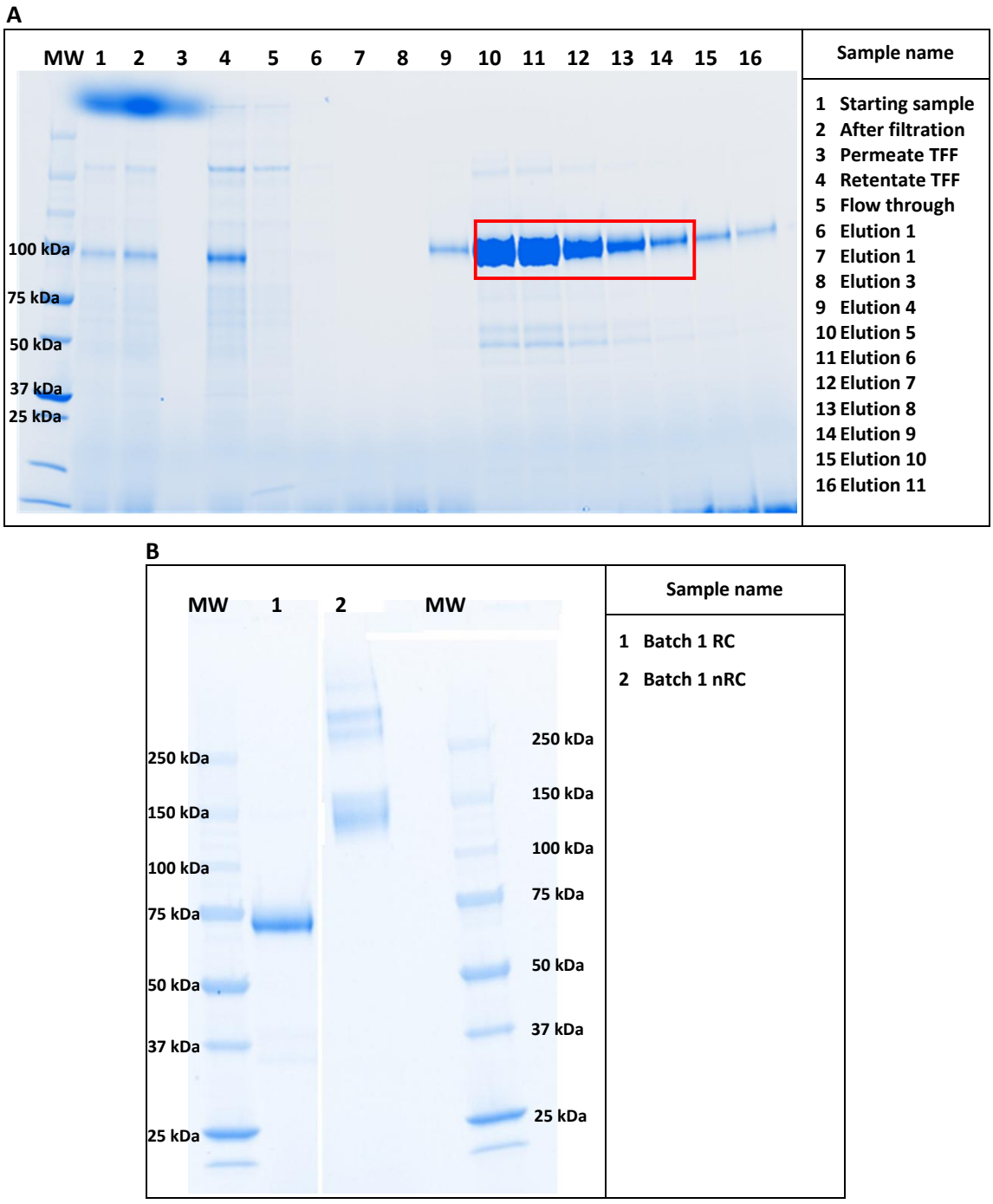


Figure 32. Purification process exploiting Strep-tag and protein characterization: A. SDS-PAGE analysis on protein affinity chromatography for a batch. Red rectangle represented the eluted fractions that contained the target protein. B. SDS-PAGE analysis for protein characterization on one batch. RC: reducing condition; nRC: non reducing condition.

The last batch was also characterized by the SDS-PAGE analysis (Figure 32B). Results confirmed what observed after processing of the clone 49E.49 batches. In reducing condition, the purification reached a high purity degree (≥ 95%); in non-reducing condition, the protein dimer at about 140 kDa was detected with only minimal aggregation forms at higher molecular weights. The consistency of data observed in the purification of this production confirmed the robustness, reliability and reproducibility of the process. The AMH protein yield obtained was satisfying.

Discussion

Progresses in recombinant protein production exploiting mammalian cells have been achieved since the first protein derived from this platform entered the market 30 years ago. In particular, CHO cell lines gained an attractive position in industry as the most used host for the majority of products both on the market and in clinical development, at the expense of other expression platforms. The reasons of CHO cell line success are its culture condition, safety properties and post-translational modifications without forgetting that the more CHO cell line was used the more advances were implemented. In addition, the genomic variability of CHO cells and the fact that they are functionally hemizygous for many genes are other relevant features that allowed the isolation of mutant lines with deficiencies in metabolic enzymes. Indeed, three CHO hosts have been routinely used to express recombinant products: CHO cells deficient for DHFR gene, i.e. DUXB11 and DG44, and those deficient for GS gene, i.e. CHOK1SV. These cell lines represented a readily manipulated phenotype suitable to select for genome integration and stable expression of transgene. Moreover, an inhibiting step with drugs that block DHFR and GS enzymes could be performed to force the selection process allowing the gene amplification and, in turn, improving the linked transgene expression. However, the attractiveness for these expression systems has been impaired by two unrelated factors. Firstly, the intellectual property on the GS technology that has only recently expired. Secondly, the new engineering cell lines obtained with tools such as zinc finger endonucleases, meganucleases, TALENs and CRISPR (Estes and Melville, 2014; Zhu, 2012).

Besides the cell line platform, other elements play a pivotal role in the final productivity. The most relevant are: the transgene integration efficiency; the design of the expression construct to assure high level of transgene transcription; the development of high-throughput screening method to allow the isolation of clones with the highest performance; the optimization of culture condition to improve the viability and, hence, the productivity. The development of recombinant CHO clonal cell lines has been hampered by unstable and variable transgene expression caused by random integration. Therefore, transgene integration strategies were implemented to overcome the limitations of conventional transfection methods. The recombinase mediated cassette exchange (RMCE) allowed to obtain site specific integration in active and stable region (hotspot) of the genome, promising speed and efficiency improvement in the generation of high producer cell lines. However, the yields reported seem to be only moderate (Bandaranayake and Almo, 2014). Another strategy is the viral delivery system since the viruses naturally infect

cells to integrate their genetic material into the host genome. Nevertheless, drawbacks such as the modest packaging size of viral vector and the regulatory concerns about the risk linked to the handling of viruses should be considered (Bandaranayake and Almo, 2014). An attractive alternative are the DNA transposons, which support the integration of recombinant genes into cultivated mammalian cells (Ding et al., 2005). Recently, these transposable elements have been investigated for industrial recombinant protein production. The advantages are the safe and easy handling, which do not differ from the conventional plasmid vector (Matasci et al., 2011). In this work, two of the most promising technologies in the recombinant protein production are combined in order to enhance the final productivity. Thus, a new generated DHFR knockout cell line is exploited in association with the *piggyBac* (PB) transposon system aiming to assure a stable transgene integration and a high level of expression.

Evaluation of PB transposon system in CHO DHFR knockout cell line

In our previous data we observed that by using the PB transposon system it is possible to enhance the frequency of high producing clones compared to the conventional transfection. Therefore, we decided to evaluate this powerful tool of transgene integration in combination with a new generated DHFR knockout cell line. We observed that this new cell line in association with the PB transposon system allowed to obtain a substantial improvement in the AMH protein yield.

Gene amplification process properly succeeded in PB pools

The DHFR deficiency in CHO cells can be exploited not only as a selection marker but also performing a gene amplification process by MTX treatment. The gene amplification process on standard vector pools and *piggyBac* pools (SV and PB pools, respectively) was firstly performed in a single step protocol, with different concentrations of MTX. In SV pools the decrease in cell viability and growth during the amplification was correlated to the level of MTX which cells were exposed to. In contrast, PB pools maintained a percentage of growth similar to that of not treated pool. Unexpectedly, both pools survived to the MTX treatment at the highest concentration. Probably a few copies of the exogenous transgene are sufficient for these cells to survive. After the gene amplification, the AMH expression resulted comparable in both pools from 50 nM to 500 nM MTX treatment but differences in

the growth rate were observed. SV pools showed growth impairments suggesting that in these pools all efforts were invested in survival through the DHFR enzyme expression and, in turn, AMH protein, rather than in cell division. On the other hand, despite PB pools treated at the same concentration did not show great differences in protein expression than SV pools, they maintained an optimal growth rate meaning that the protein expression pathway in these pools could be more efficient. The evidence was different after MTX treatment with 1 µM and 2 µM concentrations; PB pools resulted more productive compared to SV pools at the same treatment condition of 2.5- and 6-fold, respectively. Since PB transposon allows to enhance the frequency of high producing clones compare to standard vector, it is reasonable that the potentiality of PB transposon resulted favorable in condition of forced selective pressure. Probably, the rapid and high stress caused by the treatment did not allow the few high producing clones in SV pool to reach a good growth rate and enrich the population. On the other hand, since the high producing clones are more numerous in PB pool, the selective pressure is not only well tolerated, but resulted also in an efficient gene amplification due to the less impairment in survival.

Anyway, the MTX treatment allowed to augment the protein expression in both pools compared to not treated ones, expecially at the highest concentrations. In SV pools, the maximum increase was 3-3,5 fold. In PB pools, the MTX treatment with 500 nM, 1 µM and 2 µM led to obtain 5-, 10- and 19-fold increase, respectively. The increment was lower than what was expected referring to literature, except for that obtained with PB pool after 2 µM MTX treatment which fall within the reported values (Lucas et al., 1996; Wurm, 2004). We supposed that this phenomenon occurred due to the single step protocol. Indeed, the high-level resistance to MTX could result in cells synthesizing an MTX-resistant DHFR mutant or in cells with altered MTX-transport properties as response to the rapid stress applied with a single round MTX treatment. Therefore, gradually increasing MTX concentrations with a multiple round protocol is usually employed for the gene amplification (Cacciatore et al., 2010; Chusainow et al., 2009; Voronina et al., 2016). A multiple step protocol was also performed on the original stable cell pool treating it progressively with increasing MTX concentrations. Results showed a direct correlation between the increment in MTX concentration and the increase in AMH production with a final augment of about 5-fold and 7-fold for SV and PB stable pools, respectively, compared to not treated stable pool. The multiple step protocol resulted advantageous for SV pools compared to the single step, because we obtained a higher AMH expression at each MTX concentration step. However,

it was not the same for PB pools, where the AMH expression after the single step protocol with 2 µM MTX concentration was more than double of the AMH yield obtained with the multiple step one. Moreover, the amplification effects were more evident in PB pools compared to SV pools after the single step protocol, whereas no significant difference between PB and SV pools was observed in the multiple step one. Although the multiple step protocol is recommended to avoid the establishment of mechanisms other than gene amplification, a few reasons discouraged its application. Firstly, it is labor intensive and extremely time-consuming. Secondly, increasing the MTX concentration consents to select high copy number clones in a limited way. Beyond a certain concentration, the increment in MTX concentrations does not lead to generate clones with higher transgene copy number. Thus, the saturation effect of using MTX to select for amplified cell lines also limits the extent to which high producers can be isolated (Cacciatore et al., 2010). Thirdly, in our experiments no great advantage was gained using it, in particular in PB pools. Our observation is reinforced by other evidence in literature (Gustafsson, 2010). An option to consider for the future is performing a 2 step amplification protocol, which resulted a valid procedure in different experiments, also in association with the transposon system (Kacherovsky et al., 2015).

Gene amplification process shifted clone frequency towards high producing category

Once isolated the single clones, the high producing clone frequency analysis was performed. Our results on the clones derived from not treated pools confirmed previous data obtained by our lab, since standard vector generated more than 50% of clones with a very low productivity, contrary to PB transposon, which allowed to obtain only productive clones. These data highlighted the potency of an active transposition instead of an episomal DNA random integration. However, by using the transposon system and the DHFR knockout CHO cell line without MTX gene amplification, no clones were obtained in the highest productive category, that is in contrast to what observed previously by using the CHO-K1 adherent cell line. A possible explanation could be that, even if the MTX treatment was not applied, the absence of HT in the medium caused a great selective pressure that impacted substantially in the DHFR knockout cell line. This could, in turn, have affected the cloning step in which it is necessary to establish a delicate balance between clone survival (i.e. viability and growth rate) and transgene expression. However, after gene amplification, the PB transposon system allowed to shift the frequency of clones toward a medium-high

productivity, an achievement that was partially reached also by clones derived from standard vector. Since standard vector integrates randomly and often as concatemer, it originates a low number of productive clones. Nevertheless, increasing the concentration of MTX probably consented an efficient gene amplification because large part of the concatemer could be duplicated due to the proximity of multiple copies of sequences linked in series. On the contrary, the process mediated by the transposase assures single copy integrations, thus it resulted in a higher number of productive clones, but in a less efficient process of MTX gene amplification since only a copy could be duplicate at each event of recombination.

To sum up, this analysis showed that by using the PB transposon system it is possible to enhance the frequency of high-producing clones compared to standard vector, confirming our preliminary data. After MTX gene amplification, the frequency distribution was substantially shifted towards the high producer categories in clones derived from standard vector, but it was more evident in clones derived from PB transposon, in fact the majority of them fell into the high and very high categories. The most producing category was equally populated by clones derived from both standard vector and PB transposon. This is probably due to the concatemer integration of standard vector that facilitates the gene amplification process compared to the single copy integration that is a PB transposon signature. However, a consideration to take into account is that concatemer amplification is also prone to silencing (McBurney et al., 2002).

PB clones prevailed in stability and overcame silencing

Even if only a few clones were analyzed, the genetic data on the selected clones revealed substantial considerations. PB clones derived from not treated pool showed 2-3 fold more trangenes integrated copies than SV clones, as expected by the active integration mediated by the transposase. As hypothesized, the MTX treatment at the highest concentration helped to reduce the difference in integration of transgene copies between SV and PB clones probably due to the easier amplification process on concatemers than on single copy integrations. However, the MTX treatment at high concentration generated instability as occurred in SV clone 49C.170. A possible explanation could be that an excessive selective pressure forced cells to a rapid adaptation in order to contrast the negatively effects of MTX, but, upon removal of the MTX selective pressure, rearrangements in cells genome occurred because a high level of transgene is no more required to survive.

The treatment at 500 nM MTX resulted efficient on PB clones with a 2-fold increase of integrated transgene copies but not on SV clones, except for the clone 49B.82 which accounted for about 50 integrated copies. These high copy number insertions are suggested to derive from an amplification process of an integrated concatemer. Despite the higher amplification level occurred in treated SV clones compared to treated PB clones, the latter showed higher copy number integrations than SV clones (with the exception of the clone 49B.82) confirming the greatest potential of the transposon system than standard delivery methods not only in absence of MTX treatment, but also after it. We observed that there was a genetic instability over time in clones derived from both standard vector and PB transposon. However, after 30 passages, they resulted in a maximum transgene copy loss of 30%, which was considered acceptable. An exception was the SV clone 49C.170 that showed a loss of 80%. PB clones preserved a more homogeneous genetic stability profile among different culture conditions resulting in less significant transgene copy loss than in SV clones.

Results on productivity analysis reflected the genetic data on not treated SV and PB clones and on those treated with 2 µM MTX, as well as a correspondence to the genetic data was observed in the protein expression of 500 nM treated PB clones. However, inconsistencies were obtained in 500 nM treated SV clones. Indeed, despite an invalid gene amplification in the clone 49B.84 a 2-fold increase in the protein productivity was observed. We suggested that mechanisms other than gene amplification occurred in this clones that enhanced the AMH expression. The other inconsistency was observed in the clone 49B.82, which showed the highest copies of integrated transgene, but no correlation with the AMH expression. As mentioned before, it is reported that multiple copy integrations undergo transgene silencing and instability compared to single copy integrations (McBurney et al., 2002). After 22 passages, PB clones showed less instability than SV clones maintaining a productivity loss beneath 30%, which is the threshold of acceptable loss (Dorai et al., 2012). However, after 30 passages, seven of twelve total clones were shown to be remarkably unstable with the maximum loss (80%) reached by two SV clones. Supporting our results, Dr. Chusainow and collaborators in their work observed a decrement up to 93% in antibody mRNA levels over 30 passages without MTX selective pressure, despite any significant changes in gene copy number. They supposed that this event was due to a lower transcription efficiency rather than a loss of transgene copies. That is because several mechanisms could lead to the loss of transcriptional activity of the transgene. The gene silencing could occur through

methylation, histone deacetylation, or chromatin condensation. The mRNA stability could also be impaired due to the decrease in transgene mRNA level. Even though a correlation between decrease in transgene copy number and productivity loss during long-term culture without MTX is more frequent, some studies reported a productivity loss in clones that maintained the genetic stability (Chusainow et al., 2009).

In literature, it is reported that exploiting CHO-DG44 or similar cell lines, the gene amplification by MTX usually results in several hundred to a few thousand copies of the integrated plasmid with an increase in productivity up to 10- to 20-fold even if a variation among individual clones is observed (Kingston et al., 2002; Wurm, 2004). Since we did not observe such copy number amplification or increment in productivity, we could consider different strategies for the gene amplification. An example could be to perform the MTX treatment on isolated individual clones or to isolate individual clones from the first level of MTX selection, treating each clone to a higher level of MTX, again isolating single clones, and so on for each level of selection. This procedure is reported to be more efficient than the gene amplification protocol on cell pools. The reason could be that within each pool the high-producing clones can be overtaken by lower producing clones due to a lower growth rate. However, the labor intesive and time consuming process that required extra cloning steps dissuades to pursue this method (Cacciatore et al., 2010; Voronina et al., 2016). This strategy can be reasonable applied if a high-troughput screening is available.

The advantages of PB transposon as transgene delivery method compared to standard vector emerged also in the DHFR knockout cell line, resulting convenient and compatible to the gene amplification process with MTX treatment. In paritcular, our results supported that the PB transposon system in combination with the DHFR knockout cell line and a 500 nM MTX treatment resulted the most confident strategy to obtain high-producing clones not prone to silencing and maintaining an acceptable stability over time.

Cold does it better

The culture medium represents the cell environment. Therefore, the optimization in its composition determines a pivotal impact on cell growth, productivity and product quality including glycosylation. Several media are now available on the market in association with feeds. A media screening results a valid procedure to improve the recombinant protein production. In our work, only a few media could be evaluated due to a manual screening process. No significant increase in the AMH expression was obtained with tested media.

Anyhow, we do not exclude that a different production medium can be worth for the selected clones.

Different strategies have been applied to enhance the viable cell densities and to increase the culture longevity that, in turn, reflect on production titers. The mild hypothermia growth condition allowed us to improve the clone productivity of about 2-fold. Our findings are in accordance with literature. Indeed, the biphasic culture system is a well established procedure and an increase in the recombinant protein or antibody production of about 2-3 fold in CHO batch cultures is observed (Fogolín et al., 2004; Fox et al., 2004; Marchant et al., 2008; Nam et al., 2009; Oguchi et al., 2006; Shi et al., 2005; Yoon et al., 2003). However, it is also reported that the increment in specific recombinant protein productivity at low temperature does not result in volumetric productivity of the recombinant protein because the enhancement in productivity could not be balanced by the decrease in the cell growth rate. In order to overcome the limitation in the cell growth rate a pre-adaptation of CHO cells at low temperature can be performed (Mohan et al., 2008; Sunley et al., 2008).

Strategies for optimization of gene expression

DHFR destabilization: a weakening strategy, but not a strong improvement

Vector engineering aiming the destabilization of the selection marker is an attractive alternative to increase the selective pressure avoiding the increment of drug concentration. We decided to evaluate the effects of this strategy on our cell line, based on the hypothesis, supported by previous observations, that a few DHFR copies resulted sufficient to confer stability in clones generated with the DHFR knockout cell line after MTX treatment. In our work, different elements, which affect the DHFR expression pathway, were tested in the transposon to evaluate the effects on the selection process and the gene amplification. These elements include a DHFR deoptimization, a less efficient IRES (FMDV IRES) and PEST sequences. In not treated stable pools, we observed that the PEST sequences, the deoptimized DHFR and their association did not result in an AMH protein increment compared to the control. In contrast, the insertion of the FMDV IRES in the transposons showed an AMH protein expression increase. The insertion of each other element alone or in combination in the transposon carrying the FMDV IRES did not significantly enhance the protein expression. Therefore, the enhanced AMH protein expression, before the MTX

treatment, was ascribable to only the FMDV IRES. After 500 nM MTX treatment, the AMH protein expression was increased in all evaluable conditions. The pools containing the most intense constrains did not survive because the destabilization in the DHFR pathway was probably too detrimental to tolerate the MTX treatment. Also in this case, the PEST sequences, the deoptimized DHFR and their association did not give an advantage compared to the control. On the other hand, the FMDV IRES resulted to be more efficient also after the gene amplification process. Our observations are partially in line with the work of Dr. Chin and collaborators. In their analysis, before the MTX treatment, PEST sequences and IRES attenuation did not improve the protein expression, while it was observed after MTX amplification. They suggested that the constrains could probably negatively affect the recombinant protein production in the initial phase; however, after MTX treatment, the protein titers increase thanks to the increment in gene copy number induced by the high selective pressure. We supposed that the 500 nM MTX concentration resulted in an excessive selective pressure causing an opposite effect on the gene amplification. Indeed, there is not a direct correlation between increased MTX selection pressure and increased clone productivity because enhancing the MTX selective pressure could lead to display mechanisms other than DHFR-mediated gene amplification, such as altered MTX transport properties or an altered MTX-resistant DHFR enzyme amplification (Cacciatore et al., 2010; Chusainow et al., 2009; Voronina et al., 2016). In our case, in pool 5 FMDV DHFR, which carried the less efficient IRES (FMDV IRES), a treatment with 50 nM MTX concentration resulted in a similar improvement in protein expression compared to the 500 nM treatment. However, the treatment with 50 nM MTX concentration was efficient in pools carrying the most intense constrains obtaining an increase of 2-3 fold. No advantages were obtained treating the same pools at 250 nM MTX concentration. Even if pools containing the FMDV IRES survived, an impairment in viability was observed that prevented a correct protein expression evaluation. Therefore, we supposed that there is a critical MTX concentration that sustains the maximum gene amplification that pools can tolerate. Exceeding this concentration does not give advantage in term of protein expression, on the contrary, it could result in an opposite effect. Despite the positive results obtained with the 50 nM treatment on FMDV IRES pools, the protein production increase was less than our expectation based on data reported in literature (Chin et al., 2015; Lucas et al., 1996; Westwood et al., 2010; Wurm, 2004). We observed also that, even with the high selective pressure, the deoptimized DHFR gene insertion resulted in the lowest protein

expression both pre- and post- MTX treatment with different MTX concentration (i.e. 50 nM and 500 nM). This consideration is in accordance with the study of Dr. Chin et al., but it disagrees with the work of Dr. Westwood and collaborators. The discrepancy should not be attributed to a different deoptimization of DHFR. Indeed, the codon adaptation index (CAI) of the deoptimized DHFR gene that we inserted in our transposons was 0.58, which resulted similar to those employed in work of Dr. Chin et al. and Dr. Westwood and collaborators. Dr. Chin et al. suggested that this could be caused by codon usage changes in cells as a mechanism of survival. This modifications could, in turn, affect the transgene protein expression that is optimized for a cell-specific codon usage. We think it could be possible and it could be included in the defense strategies that the cells act to survive. This, inevitably, impacts on the recombinant protein production. To sum up, if the DHFR expression is excessively weakened, high difficulties on balancing the MTX treatment are encountered because, besides a slow cell growth, uncontrollable cell response mechanisms can occur that, in turn, affect the recombinant protein expression.

Epigenetic regulatory elements: the supremacy of UCOE

We decided to shift the focus on epigenetic regulatory elements in order to force the transcriptional efficiency. cHS4, UCOE, MAR were evaluated in a transposon vector (pSF). The UCOE potentiality emerged among all other epigenetic regulatory elements, increasing the AMH yield of 6-fold and 3-fold compared to the basal pSF and PB pools (pools derived from pSF and PB transposon), respectively. There were no significant difference in protein expression exploiting one or two UCOEs, but an impairment in viability and growth rate in the pool transfected with transposon containing the two UCOE sequences was observed. Our work resulted in accordance with that of Dr. Saunders et al. Indeed, they also demonstrated the supremacy of the UCOE in antibody expression among other regulatory elements with an increase of 6.5-fold in stable pools compared to the control, that was very close to what we obtained. Unexpectedly, no effect was obtained in the pools transfected with constructs containing the MAR contrary to several works reporting the beneficial effects of this element (Girod et al., 2007, 2005; Kim et al., 2004; Zahn-Zabal et al., 2001). In particular, our results were in contrast with work of Dr. Otte et al. They showed that neither the cHS4 insulator nor the UCOE had beneficial effect on the gene expression, but the MAR resulted in increase in the protein expression (Otte et al., 2007). We observed exactly the opposite because the MAR showed a comparable effect to the cSH4 and

resulted totally inefficient compared to the UCOE. Probably, these discrepancies were due to the position in which the elements were inserted. Indeed, Dr. Otte and collaborators placed the cSH4 and the MAR downstream the expression cassette. However, they inserted the UCOE upstream, as we did. We could affirm that many factors could influence the regulatory element activity including vector engineering, cell lines and selection method. The MAR elements were also evaluated in combination with PB transposon reporting a positive effect (Ley et al., 2013). Comparing our results with the work of Dr. Ley and collaborators, we could confirm that the transposon with the MAR 1-68 element positioned at the 5' of the expression cassette did not show the positive effect because it impaired the transposition due to the topological constrains. However, the reason of inefficiency of the MAR 1-68 element positioned in the middle of expression cassette observed in our work remained unclear. Despite the MAR elements resulted to be promoter independent, Dr. Ho and collaborators demonstrated that it showed its efficiency if associated to SV40 promoter and not with CMV promoter (Ho et al., 2015). Probably this could be one of the factor that have impacted on the MAR activity in our transposon which was driven by CMV promoter.

None of the cell pools showed an improvement in the protein yield after treatment with 50 nM MTX concentration. This was partially observed also after 250 nM MTX concentration treatment on pSF, pSF_U and PB pools (where psF_U pool is the pool derived from pSF transposon carrying an UCOE sequence). A mild gene amplification was obtained in pSF and PB pools, but no effect was observed in pSF_U pool. A possible limitation in the gene amplification process could be the increased size of the transposable fragment due to the presence of the epigenetic regulatory element. This could happen in a more evident manner after 250 nM MTX treatment, since pSF_U pool did not show increase in the protein expression in contrast to pSF and PB pools.

Despite the ineffectiveness of MTX treatment on the AMH protein expression in pSF_U pool, we performed a cloning step on not treated pSF_U pool in order to isolate high-producing clones. Data showed that the frequency of high-producing clones derived from pSF transposon resulted similar or even better than that obtained with PB transposon in CHO-K1 adherent cells showing a peak of clones falling in the medium category and a good percentage in high and very high category. Impressive results were obtained in clones derived from pSF_U transposon, in which productivity was completely shifted towards the highest producing category accounting for 65% of clones in the high and very high category. This was a great achievement because the frequency of high and very high producing clones

derived from pSF_U pool resulted better than that obtained from PB pool treated with 2 µM MTX concentration. This means reducing efforts and time in cloning and screening steps necessary to isolate high producing clones. Therefore, the UCOE was confirmed to be efficient in enhancing protein expression.

Finally, our aim was to compare the productivity of clones derived from pSF_U pool to that of clones obtained with the PB transposon system in previous experiments and more importantly to study if the UCOE can assure expression and genetic stability over time, avoiding problems encountered with PB transposon. We observed that pSF and PB transposon clones showed 5-fold and 2-fold more transgene integrated copies, respectively, than pSF_U ones. We supposed that the transposition in pSF_U transposon was hampered by the high size for the presence of the UCOE. However, no correlation between high transgene copy number and productivity was observed in pSF clones. We ascribed this event to the presence of the two promoters, one for the AMH gene and one for the DHFR gene, which could display the transcriptional interference. This well-known phenomenon causes a perturbation of one transcription unit by another (Curtin et al., 2008; Eszterhas et al., 2002; Uchida et al., 2013). Otherwise, the distance between AMH and DHFR genes in pSF, which is longer than that in PB transposon (which carries IRES sequence), could allow a more accessibility and sensibility of the AMH gene to silencing effects. Anyhow, it is important to highlight that, despite a low transgene copy number, pSF_U clones showed their supremacy in productivity with a 4-fold increment, in average, compared to both pSF and PB ones. Moreover, the expression and genetic stability were maintained over time in all clones with a remarkably consistency in pSF_U ones.

Our results supported the potentiality of the UCOE sequence in combination with the transposon system to substantially enhance the frequency of high producing clones and their protein productivity sustaining expression and genetic stability over time.

In conclusion, in this work a new generated DHFR knockout cell line was exploited in association with the *piggyBac* transposon system to achieve high expression of recombinant proteins. The PB transposon system was confirmed to be the best method to obtain high expression of the AMH protein allowing higher performances in gene copy number integration and genetic stability compared to standard transfection method. In order to enhance even more the protein productivity, several strategies based on transposon vector engineering were implemented to target various steps of the protein expression pathway or by acting on the epigenetic regulation to further increase the clone

productivity. A mild increment was obtained using strategies that reduce the DHFR activity acting on its expression pathway. On the other hand, by employing epigenetic regulatory elements, such as the UCOE, not only a substantial increase in the AMH protein yield was achieved, but also the expression and genetic stability were strictly conserved over time.

References

$315.9 Million Protein Therapeutics Market Analysis and Trends Report 2016-2025: Recent Innovations in Recombinant Protein drugs have Changed the Landscape - Research and Markets [WWW Document], 2017. AboutPharma. URL http://www.aboutpharma.com/blog/2017/01/27/315-9-million-protein-therapeutics-market-analysis-and-trends-report-2016-2025-recent-innovations-in-recombinant-protein-drugs-have-changed-the-landscape-research-and-markets/ (accessed 3.5.17).

Ackermann, M., Lachmann, N., Hartung, S., Eggenschwiler, R., Pfaff, N., Happle, C., Mucci, A., Göhring, G., Niemann, H., Hansen, G., Schambach, A., Cantz, T., Zweigerdt, R., Moritz, T., 2014. Promoter and lineage independent anti-silencing activity of the A2 ubiquitous chromatin opening element for optimized human pluripotent stem cell-based gene therapy. Biomaterials 35, 1531–1542. https://doi.org/10.1016/j.biomaterials.2013.11.024

Aga, M., Yamano, N., Kumamoto, T., Frank, J., Onitsuka, M., Omasa, T., 2015. Construction of a gene knockout CHO cell line using a simple gene targeting method. BMC Proc. 9, P2. https://doi.org/10.1186/1753-6561-9-S9-P2

Aker, M., Tubb, J., Groth, A.C., Bukovsky, A.A., Bell, A.C., Felsenfeld, G., Kiem, H.-P., Stamatoyannopoulos, G., Emery, D.W., 2007. Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. Hum. Gene Ther. 18, 333–343. https://doi.org/10.1089/hum.2007.021

Anderson, D.D., Quintero, C.M., Stover, P.J., 2011. Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. Proc. Natl. Acad. Sci. U. S. A. 108, 15163–15168. https://doi.org/10.1073/pnas.1103623108

Anderson, R.A., Nelson, S.M., Wallace, W.H.B., 2012. Measuring anti-Müllerian hormone for the assessment of ovarian reserve: when and for whom is it indicated? Maturitas 71, 28–33. https://doi.org/10.1016/j.maturitas.2011.11.008

Antoniou, M., Harland, L., Mustoe, T., Williams, S., Holdstock, J., Yague, E., Mulcahy, T., Griffiths, M., Edwards, S., Ioannou, P.A., Mountain, A., Crombie, R., 2003. Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics 82, 269–279.

Aruffo, A., 2001. Transient expression of proteins using COS cells. Curr. Protoc. Neurosci. Chapter 4, Unit 4.7. https://doi.org/10.1002/0471142301.ns0407s02

Backliwal, G., Hildinger, M., Chenuet, S., Wulhfard, S., De Jesus, M., Wurm, F.M., 2008. Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. Nucleic Acids Res. 36, e96. https://doi.org/10.1093/nar/gkn423

Bandaranayake, A.D., Almo, S.C., 2014. Recent advances in mammalian protein production. FEBS Lett. 588, 253–260. https://doi.org/10.1016/j.febslet.2013.11.035

Belville, C., Van Vlijmen, H., Ehrenfels, C., Pepinsky, B., Rezaie, A.R., Picard, J.-Y., Josso, N., di Clemente, N., Cate, R.L., 2004. Mutations of the anti-mullerian hormone gene in patients with persistent mullerian duct syndrome: biosynthesis, secretion, and processing of the abnormal proteins and analysis using a three-dimensional model. Mol. Endocrinol. Baltim. Md 18, 708–721. https://doi.org/10.1210/me.2003-0358

Benton, T., Chen, T., McEntee, M., Fox, B., King, D., Crombie, R., Thomas, T.C., Bebbington, C., 2002. The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein. Cytotechnology 38, 43–46. https://doi.org/10.1023/A:1021141712344

Berrios, J., Díaz-Barrera, A., Bazán, C., Altamirano, C., 2009. Relationship between tissue plasminogen activator production and specific growth rate in Chinese hamster ovary cells cultured in mannose at low temperature. Biotechnol. Lett. 31, 1493–1497. https://doi.org/10.1007/s10529-009-0050-1

BHK-21 [C-13] ATCC ® CCL-10™ Mesocricetus auratus kidney nor [WWW Document], n.d. URL https://www.lgcstandards-atcc.org/Products/All/CCL-10.aspx?geo_country=it (accessed 9.5.17).

Bochkov, Y.A., Palmenberg, A.C., 2006. Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. BioTechniques 41, 283–284, 286, 288 passim.

Boscolo, S., Mion, F., Licciulli, M., Macor, P., De Maso, L., Brce, M., Antoniou, M.N., Marzari, R., Santoro, C., Sblattero, D., 2012. Simple scale-up of recombinant antibody production using an UCOE containing vector. New Biotechnol. 29, 477–484. https://doi.org/10.1016/j.nbt.2011.12.005

Bouard, D., Alazard-Dany, N., Cosset, F.-L., 2009. Viral vectors: from virology to transgene expression. Br. J. Pharmacol. 157, 153–165. https://doi.org/10.1038/bjp.2008.349

Brown, A., Woods, W.S., Perez-Pinera, P., 2016. Multiplexed Targeted Genome Engineering Using a Universal Nuclease-Assisted Vector Integration System. ACS Synth. Biol. 5, 582–588. https://doi.org/10.1021/acssynbio.6b00056

Cacciatore, J.J., Chasin, L.A., Leonard, E.F., 2010. Gene amplification and vector engineering to achieve rapid and high-level therapeutic protein production using the Dhfr-based CHO cell selection system. Biotechnol. Adv. 28, 673–681. https://doi.org/10.1016/j.biotechadv.2010.04.003

Carter, P.J., 2011. Introduction to current and future protein therapeutics: A protein engineering perspective. Exp. Cell Res. 317, 1261–1269. https://doi.org/10.1016/j.yexcr.2011.02.013

Chen, C.Y., Shyu, A.B., 1995. AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. 20, 465–470.

Chin, C.L., Chin, H.K., Chin, C.S.H., Lai, E.T., Ng, S.K., 2015. Engineering selection stringency on expression vector for the production of recombinant human alpha1-antitrypsin using Chinese Hamster ovary cells. BMC Biotechnol. 15, 44. https://doi.org/10.1186/s12896-015-0145-9

Chusainow, J., Yang, Y.S., Yeo, J.H.M., Toh, P.C., Asvadi, P., Wong, N.S.C., Yap, M.G.S., 2009. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? Biotechnol. Bioeng. 102, 1182–1196. https://doi.org/10.1002/bit.22158

Clark, K.J., Carlson, D.F., Leaver, M.J., Foster, L.K., Fahrenkrug, S.C., 2009. Passport, a native Tc1 transposon from flatfish, is functionally active in vertebrate cells. Nucleic Acids Res. 37, 1239–1247. https://doi.org/10.1093/nar/gkn1025

Clark, K.J.R., Chaplin, F.W.R., Harcum, S.W., 2004. Temperature effects on product-quality-related enzymes in batch CHO cell cultures producing recombinant tPA. Biotechnol. Prog. 20, 1888–1892. https://doi.org/10.1021/bp049951x

Cohen, S.N., Chang, A.C.Y., Boyer, H.W., Helling, R.B., 1973. Construction of Biologically Functional Bacterial Plasmids In Vitro. Proc. Natl. Acad. Sci. U. S. A. 70, 3240–3244.

Curtin, J.A., Dane, A.P., Swanson, A., Alexander, I.E., Ginn, S.L., 2008. Bidirectional promoter interference between two widely used internal heterologous promoters in a late-generation lentiviral construct. Gene Ther. 15, 384–390. https://doi.org/10.1038/sj.gt.3303105

Dalsgaard, T., Moldt, B., Sharma, N., Wolf, G., Schmitz, A., Pedersen, F.S., Mikkelsen, J.G., 2009. Shielding of Sleeping Beauty DNA Transposon-delivered Transgene Cassettes

by Heterologous Insulators in Early Embryonal Cells. Mol. Ther. J. Am. Soc. Gene Ther. 17, 121–130. https://doi.org/10.1038/mt.2008.224

Dewailly, D., Andersen, C.Y., Balen, A., Broekmans, F., Dilaver, N., Fanchin, R., Griesinger, G., Kelsey, T.W., La Marca, A., Lambalk, C., Mason, H., Nelson, S.M., Visser, J.A., Wallace, W.H., Anderson, R.A., 2014. The physiology and clinical utility of anti-Mullerian hormone in women. Hum. Reprod. Update 20, 370–385. https://doi.org/10.1093/humupd/dmt062

di Clemente, N., Jamin, S.P., Lugovskoy, A., Carmillo, P., Ehrenfels, C., Picard, J.-Y., Whitty, A., Josso, N., Pepinsky, R.B., Cate, R.L., 2010. Processing of anti-mullerian hormone regulates receptor activation by a mechanism distinct from TGF-beta. Mol. Endocrinol. Baltim. Md 24, 2193–2206. https://doi.org/10.1210/me.2010-0273

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., Xu, T., 2005. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122, 473–483. https://doi.org/10.1016/j.cell.2005.07.013

Dorai, H., Corisdeo, S., Ellis, D., Kinney, C., Chomo, M., Hawley-Nelson, P., Moore, G., Betenbaugh, M.J., Ganguly, S., 2012. Early prediction of instability of Chinese hamster ovary cell lines expressing recombinant antibodies and antibody-fusion proteins. Biotechnol. Bioeng. 109, 1016–1030. https://doi.org/10.1002/bit.24367

DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H., Calos, M.P., 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. 7, 379–387.

Dumont, A., Robin, G., Catteau-Jonard, S., Dewailly, D., 2015. Role of Anti-Müllerian Hormone in pathophysiology, diagnosis and treatment of Polycystic Ovary Syndrome: a review. Reprod. Biol. Endocrinol. RBE 13. https://doi.org/10.1186/s12958-015-0134-9

Dumont, J., Euwart, D., Mei, B., Estes, S., Kshirsagar, R., 2016. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. Crit. Rev. Biotechnol. 36, 1110–1122. https://doi.org/10.3109/07388551.2015.1084266

Durocher, Y., Butler, M., 2009. Expression systems for therapeutic glycoprotein production. Curr. Opin. Biotechnol. 20, 700–707. https://doi.org/10.1016/j.copbio.2009.10.008

Estes, S., Melville, M., 2014. Mammalian cell line developments in speed and efficiency. Adv. Biochem. Eng. Biotechnol. 139, 11–33. https://doi.org/10.1007/10_2013_260

Eszterhas, S.K., Bouhassira, E.E., Martin, D.I.K., Fiering, S., 2002. Transcriptional Interference by Independently Regulated Genes Occurs in Any Relative Arrangement of the Genes and Is Influenced by Chromosomal Integration Position. Mol. Cell. Biol. 22, 469–479. https://doi.org/10.1128/MCB.22.2.469-479.2002

Fallaux, F.J., Bout, A., van der Velde, I., van den Wollenberg, D.J., Hehir, K.M., Keegan, J., Auger, C., Cramer, S.J., van Ormondt, H., van der Eb, A.J., Valerio, D., Hoeben, R.C., 1998. New helper cells and matched early region 1-deleted adenovirus vectors prevent generation of replication-competent adenoviruses. Hum. Gene Ther. 9, 1909–1917. https://doi.org/10.1089/hum.1998.9.13-1909

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J.L., Vázquez, E., Villaverde, A., 2009. Microbial factories for recombinant pharmaceuticals. Microb. Cell Factories 8, 17. https://doi.org/10.1186/1475-2859-8-17

Feyereisen, E., Méndez Lozano, D.H., Taieb, J., Hesters, L., Frydman, R., Fanchin, R., 2006. Anti-Müllerian hormone: clinical insights into a promising biomarker of ovarian follicular status. Reprod. Biomed. Online 12, 695–703.

Fischer, S.E., Wienholds, E., Plasterk, R.H., 2001. Regulated transposition of a fish transposon in the mouse germ line. Proc. Natl. Acad. Sci. U. S. A. 98, 6759–6764. https://doi.org/10.1073/pnas.121569298

Flintoff, W.F., Davidson, S.V., Siminovitch, L., 1976a. Isolation and partial characterization of three methotrexate-resistant phenotypes from Chinese hamster ovary cells. Somatic Cell Genet. 2, 245–261.

Flintoff, W.F., Spindler, S.M., Siminovitch, L., 1976b. Genetic characterization of methotrexate-resistant chinese hamster ovary cells. In Vitro 12, 749–757.

Fogolín, M.B., Wagner, R., Etcheverrigaray, M., Kratje, R., 2004. Impact of temperature reduction and expression of yeast pyruvate carboxylase on hGM-CSF-producing CHO cells. J. Biotechnol. 109, 179–191. https://doi.org/10.1016/j.jbiotec.2003.10.035

Fox, S.R., Patel, U.A., Yap, M.G.S., Wang, D.I.C., 2004. Maximizing interferon-gamma production by Chinese hamster ovary cells through temperature shift optimization: experimental and modeling. Biotechnol. Bioeng. 85, 177–184. https://doi.org/10.1002/bit.10861

Fu, J., Li, D., Xia, S., Song, H., Dong, Z., Chen, F., Sun, X., Tang, Z., 2009. Absolute quantification of plasmid DNA by real-time PCR with genomic DNA as external standard and its application to a biodistribution study of an HIV DNA vaccine. Anal. Sci. Int. J. Jpn. Soc. Anal. Chem. 25, 675–680.

Gaj, T., Gersbach, C.A., Barbas, C.F., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31, 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004

Gaj, T., Guo, J., Kato, Y., Sirk, S.J., Iii, C.F.B., 2012. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat. Methods 9, 805–807. https://doi.org/10.1038/nmeth.2030

Galbraith, D.J., Tait, A.S., Racher, A.J., Birch, J.R., James, D.C., 2006. Control of culture environment for improved polyethylenimine-mediated transient production of recombinant monoclonal antibodies by CHO cells. Biotechnol. Prog. 22, 753–762. https://doi.org/10.1021/bp050339v

Galfrè, G., Milstein, C., 1981. Preparation of monoclonal antibodies: strategies and procedures. Methods Enzymol. 73, 3–46.

Gassner, D., Jung, R., 2014. First fully automated immunoassay for anti-Müllerian hormone. Clin. Chem. Lab. Med. 52, 1143–1152. https://doi.org/10.1515/cclm-2014-0022

Ghaderi, D., Zhang, M., Hurtado-Ziola, N., Varki, A., 2012. Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. Biotechnol. Genet. Eng. Rev. 28, 147–175.

Girod, P.-A., Nguyen, D.-Q., Calabrese, D., Puttini, S., Grandjean, M., Martinet, D., Regamey, A., Saugy, D., Beckmann, J.S., Bucher, P., Mermod, N., 2007. Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat. Methods 4, 747–753. https://doi.org/10.1038/nmeth1076

Girod, P.-A., Zahn-Zabal, M., Mermod, N., 2005. Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines. Biotechnol. Bioeng. 91, 1–11. https://doi.org/10.1002/bit.20563

Gluzman, Y., 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell 23, 175–182.

Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., Riggs, A.D., 1979. Expression in Escherichia coli of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. U. S. A. 76, 106–110.

Grabenhorst, E., Schlenke, P., Pohl, S., Nimtz, M., Conradt, H.S., 1999. Genetic engineering of recombinant glycoproteins and the glycosylation pathway in mammalian host cells. Glycoconj. J. 16, 81–97.

Grabundzija, I., Irgang, M., Mátés, L., Belay, E., Matrai, J., Gogol-Döring, A., Kawakami, K., Chen, W., Ruiz, P., Chuah, M.K.L., VandenDriessche, T., Izsvák, Z., Ivics, Z., 2010. Comparative analysis of transposable element vector systems in human cells. Mol. Ther. J. Am. Soc. Gene Ther. 18, 1200–1209. https://doi.org/10.1038/mt.2010.47

Graham, F.L., Smiley, J., Russell, W.C., Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36, 59–74. https://doi.org/10.1099/0022-1317-36-1-59

Grandjean, M., Girod, P.-A., Calabrese, D., Kostyrko, K., Wicht, M., Yerly, F., Mazza, C., Beckmann, J.S., Martinet, D., Mermod, N., 2011. High-level transgene expression by homologous recombination-mediated gene transfer. Nucleic Acids Res. 39, e104. https://doi.org/10.1093/nar/gkr436

Graumann, K., Premstaller, A., 2006. Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnol. J. 1, 164–186. https://doi.org/10.1002/biot.200500051

Grynnerup, A.G.-A., Lindhard, A., Sørensen, S., 2012. The role of anti-Müllerian hormone in female fertility and infertility - an overview. Acta Obstet. Gynecol. Scand. 91, 1252–1260. https://doi.org/10.1111/j.1600-0412.2012.01471.x

Gustafsson, E., 2010. Evaluation of gene amplification for development of high producing biopharmaceutical cell lines. Recipharm, Cobra Bio.

Harraghy, N., Regamey, A., Girod, P.-A., Mermod, N., 2011. Identification of a potent MAR element from the mouse genome and assessment of its activity in stable and transient transfections. J. Biotechnol. 154, 11–20. https://doi.org/10.1016/j.jbiotec.2011.04.004

Hauschild-Quintern, J., Petersen, B., Cost, G.J., Niemann, H., 2013. Gene knockout and knockin by zinc-finger nucleases: current status and perspectives. Cell. Mol. Life Sci. CMLS 70, 2969–2983. https://doi.org/10.1007/s00018-012-1204-1

Hehenkamp, W.J.K., Looman, C.W.N., Themmen, A.P.N., de Jong, F.H., Te Velde, E.R., Broekmans, F.J.M., 2006. Anti-Müllerian hormone levels in the spontaneous menstrual cycle do not show substantial fluctuation. J. Clin. Endocrinol. Metab. 91, 4057–4063. https://doi.org/10.1210/jc.2006-0331

Hendrick, V., Winnepenninckx, P., Abdelkafi, C., Vandeputte, O., Cherlet, M., Marique, T., Renemann, G., Loa, A., Kretzmer, G., Werenne, J., 2001. Increased productivity of recombinant tissular plasminogen activator (t-PA) by butyrate and shift of temperature: a cell cycle phases analysis. Cytotechnology 36, 71–83. https://doi.org/10.1023/A:1014088919546

Ho, S., Wah Tong, Y., Yang, Y., 2013. Generation of monoclonal antibody-producing mammalian cell lines. Pharm. Bioprocess. 1, 71–87. https://doi.org/10.4155/pbp.13.8

Ho, S.C.L., Bardor, M., Feng, H., Mariati, null, Tong, Y.W., Song, Z., Yap, M.G.S., Yang, Y., 2012. IRES-mediated Tricistronic vectors for enhancing generation of high monoclonal antibody expressing CHO cell lines. J. Biotechnol. 157, 130–139. https://doi.org/10.1016/j.jbiotec.2011.09.023

Ho, S.C.L., Mariati, null, Yeo, J.H.M., Fang, S.G., Yang, Y., 2015. Impact of using different promoters and matrix attachment regions on recombinant protein expression level and stability in stably transfected CHO cells. Mol. Biotechnol. 57, 138–144. https://doi.org/10.1007/s12033-014-9809-2

Hou, J.J.C., Hughes, B.S., Smede, M., Leung, K.M., Levine, K., Rigby, S., Gray, P.P., Munro, T.P., 2014. High-throughput ClonePix FL analysis of mAb-expressing clones using the UCOE expression system. New Biotechnol. 31, 214–220. https://doi.org/10.1016/j.nbt.2014.02.002

Hudson, P.L., Dougas, I., Donahoe, P.K., Cate, R.L., Epstein, J., Pepinsky, R.B., MacLaughlin, D.T., 1990. An immunoassay to detect human müllerian inhibiting substance in males and females during normal development. J. Clin. Endocrinol. Metab. 70, 16–22. https://doi.org/10.1210/jcem-70-1-16

Huye, L.E., Nakazawa, Y., Patel, M.P., Yvon, E., Sun, J., Savoldo, B., Wilson, M.H., Dotti, G., Rooney, C.M., 2011. Combining mTor inhibitors with rapamycin-resistant T cells: a two-pronged approach to tumor elimination. Mol. Ther. J. Am. Soc. Gene Ther. 19, 2239–2248. https://doi.org/10.1038/mt.2011.179

Ivics, Z., Hackett, P.B., Plasterk, R.H., Izsvák, Z., 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91, 501–510.

Ivics, Z., Izsvák, Z., 2010. The expanding universe of transposon technologies for gene and cell engineering. Mob. DNA 1, 25. https://doi.org/10.1186/1759-8753-1-25

Izsvák, Z., Ivics, Z., Plasterk, R.H., 2000. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J. Mol. Biol. 302, 93–102. https://doi.org/10.1006/jmbi.2000.4047

Jayapal, Wlaschin, Hu, 2007. Recombinant Protein Therapeutics from CHO Cells - 20 Years and Counting [WWW Document]. ResearchGate. URL https://www.researchgate.net/publication/279541139_Recombinant_Protein_Therapeutics_from_CHO_Cells_-_20_Years_and_Counting (accessed 9.4.17).

Josso, N., Belville, C., di Clemente, N., Picard, J.-Y., 2005. AMH and AMH receptor defects in persistent Müllerian duct syndrome. Hum. Reprod. Update 11, 351–356. https://doi.org/10.1093/humupd/dmi014

Josso, N., Picard, J.Y., Imbeaud, S., di Clemente, N., Rey, R., 1997. Clinical aspects and molecular genetics of the persistent müllerian duct syndrome. Clin. Endocrinol. (Oxf.) 47, 137–144.

Josso, N., Rey, R.A., Picard, J.-Y., 2013. Anti-müllerian hormone: a valuable addition to the toolbox of the pediatric endocrinologist. Int. J. Endocrinol. 2013, 674105. https://doi.org/10.1155/2013/674105

Kaas, C.S., Kristensen, C., Betenbaugh, M.J., Andersen, M.R., 2015. Sequencing the CHO DXB11 genome reveals regional variations in genomic stability and haploidy. BMC Genomics 16. https://doi.org/10.1186/s12864-015-1391-x

Kacherovsky, N., Liu, G.W., Jensen, M.C., Pun, S.H., 2015. Multiplexed gene transfer to a human T-cell line by combining Sleeping Beauty transposon system with methotrexate selection. Biotechnol. Bioeng. 112, 1429–1436. https://doi.org/10.1002/bit.25538

Kahlig, K.M., Saridey, S.K., Kaja, A., Daniels, M.A., George, A.L., Wilson, M.H., 2010. Multiplexed transposon-mediated stable gene transfer in human cells. Proc. Natl. Acad. Sci. U. S. A. 107, 1343–1348. https://doi.org/10.1073/pnas.0910383107

Kantardjieff, A., Zhou, W., 2013. Mammalian Cell Cultures for Biologics Manufacturing, in: Zhou, W., Kantardjieff, A. (Eds.), Mammalian Cell Cultures for Biologics Manufacturing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–9. https://doi.org/10.1007/10_2013_255

Kao, F.T., Puck, T.T., 1968. Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells. Proc. Natl. Acad. Sci. U. S. A. 60, 1275–1281.

Kaufmann, H., Mazur, X., Fussenegger, M., Bailey, J.E., 1999. Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. Biotechnol. Bioeng. 63, 573–582.

Khan, K.H., 2013. Gene Expression in Mammalian Cells and its Applications. Adv. Pharm. Bull. 3, 257–263. https://doi.org/10.5681/apb.2013.042

Kilbey, B.J., Legator, M., Nicholson, W., Ramel, C., 2012. Handbook of Mutagenicity Test Procedures. Elsevier.

Kim, J.-M., Kim, J.-S., Park, D.-H., Kang, H.S., Yoon, J., Baek, K., Yoon, Y., 2004. Improved recombinant gene expression in CHO cells using matrix attachment regions. J. Biotechnol. 107, 95–105.

Kim, J.Y., Kim, Y.-G., Lee, G.M., 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl. Microbiol. Biotechnol. 93, 917–930. https://doi.org/10.1007/s00253-011-3758-5

Kim, N.S., Byun, T.H., Lee, G.M., 2001. Key determinants in the occurrence of clonal variation in humanized antibody expression of cho cells during dihydrofolate reductase mediated gene amplification. Biotechnol. Prog. 17, 69–75. https://doi.org/10.1021/bp000144h

Kim, S.J., Lee, G.M., 1999. Cytogenetic analysis of chimeric antibody-producing CHO cells in the course of dihydrofolate reductase-mediated gene amplification and their stability in the absence of selective pressure. Biotechnol. Bioeng. 64, 741–749.

Kim, T.K., Eberwine, J.H., 2010. Mammalian cell transfection: the present and the future. Anal. Bioanal. Chem. 397, 3173–3178. https://doi.org/10.1007/s00216-010-3821-6

Kingston, R.E., Kaufman, R.J., Bebbington, C.R., Rolfe, M.R., 2002. Amplification using CHO cell expression vectors. Curr. Protoc. Mol. Biol. Chapter 16, Unit 16.23. https://doi.org/10.1002/0471142727.mb1623s60

Kost, T.A., Condreay, J.P., Jarvis, D.L., 2005. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nat. Biotechnol. 23, 567–575. https://doi.org/10.1038/nbt1095

Kotani, T., Nagayoshi, S., Urasaki, A., Kawakami, K., 2006. Transposon-mediated gene trapping in zebrafish. Methods San Diego Calif 39, 199–206. https://doi.org/10.1016/j.ymeth.2005.12.006

Kumar, A., Kalra, B., Patel, A., McDavid, L., Roudebush, W.E., 2010. Development of a second generation anti-Müllerian hormone (AMH) ELISA. J. Immunol. Methods 362, 51–59. https://doi.org/10.1016/j.jim.2010.08.011

Kumar, N., Gammell, P., Meleady, P., Henry, M.S., Clynes, M., 2008. Differential protein expression following low temperature culture of suspension CHO-K1 cells. BMC Biotechnol. https://doi.org/10.1186/1472-6750-8-42

Kunert, R., Reinhart, D., 2016. Advances in recombinant antibody manufacturing. Appl. Microbiol. Biotechnol. 100, 3451–3461. https://doi.org/10.1007/s00253-016-7388-9

Kwaks, T.H.J., Barnett, P., Hemrika, W., Siersma, T., Sewalt, R.G.A.B., Satijn, D.P.E., Brons, J.F., van Blokland, R., Kwakman, P., Kruckeberg, A.L., Kelder, A., Otte, A.P., 2003. Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. Nat. Biotechnol. 21, 553–558. https://doi.org/10.1038/nbt814

Kwaks, T.H.J., Otte, A.P., 2006. Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. Trends Biotechnol. 24, 137–142. https://doi.org/10.1016/j.tibtech.2006.01.007

La Marca, A., Stabile, G., Artenisio, A.C., Volpe, A., 2006. Serum anti-Mullerian hormone throughout the human menstrual cycle. Hum. Reprod. Oxf. Engl. 21, 3103–3107. https://doi.org/10.1093/humrep/del291

Lai, T., Yang, Y., Ng, S.K., 2013. Advances in Mammalian Cell Line Development Technologies for Recombinant Protein Production. Pharmaceuticals 6, 579–603. https://doi.org/10.3390/ph6050579

Levings, P.P., Bungert, J., 2002. The human beta-globin locus control region. Eur. J. Biochem. 269, 1589–1599.

Lewis, N.E., Liu, X., Li, Y., Nagarajan, H., Yerganian, G., O'Brien, E., Bordbar, A., Roth, A.M., Rosenbloom, J., Bian, C., Xie, M., Chen, W., Li, N., Baycin-Hizal, D., Latif, H., Forster, J., Betenbaugh, M.J., Famili, I., Xu, X., Wang, J., Palsson, B.O., 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. Nat. Biotechnol. 31, 759–765. https://doi.org/10.1038/nbt.2624

Ley, D., Harraghy, N., Le Fourn, V., Bire, S., Girod, P.-A., Regamey, A., Rouleux-Bonnin, F., Bigot, Y., Mermod, N., 2013. MAR Elements and Transposons for Improved Transgene Integration and Expression. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0062784

Li, M.A., Pettitt, S.J., Eckert, S., Ning, Z., Rice, S., Cadiñanos, J., Yusa, K., Conte, N., Bradley, A., 2013. The piggyBac transposon displays local and distant reintegration preferences and can cause mutations at noncanonical integration sites. Mol. Cell. Biol. 33, 1317–1330. https://doi.org/10.1128/MCB.00670-12

Li, M.A., Turner, D.J., Ning, Z., Yusa, K., Liang, Q., Eckert, S., Rad, L., Fitzgerald, T.W., Craig, N.L., Bradley, A., 2011. Mobilization of giant piggyBac transposons in the mouse genome. Nucleic Acids Res. 39, e148. https://doi.org/10.1093/nar/gkr764

Li, R., Zhuang, Y., Han, M., Xu, T., Wu, X., 2013. piggyBac as a high-capacity transgenesis and gene-therapy vector in human cells and mice. Dis. Model. Mech. 6, 828–833. https://doi.org/10.1242/dmm.010827

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., Kain, S.R., 1998. Generation of destabilized green fluorescent protein as a transcription reporter. J. Biol. Chem. 273, 34970–34975.

Li, Z., Michael, I.P., Zhou, D., Nagy, A., Rini, J.M., 2013. Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. Proc. Natl. Acad. Sci. U. S. A. 110, 5004–5009. https://doi.org/10.1073/pnas.1218620110

Liang, Q., Kong, J., Stalker, J., Bradley, A., 2009. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. Genes. N. Y. N 2000 47, 404–408. https://doi.org/10.1002/dvg.20508

Lindahl Allen, M., Antoniou, M., 2007. Correlation of DNA methylation with histone modifications across the HNRPA2B1-CBX3 ubiquitously-acting chromatin open element (UCOE). Epigenetics 2, 227–236.

Liu, C., Dalby, B., Chen, W., Kilzer, J.M., Chiou, H.C., 2008. Transient transfection factors for high-level recombinant protein production in suspension cultured mammalian cells. Mol. Biotechnol. 39, 141–153. https://doi.org/10.1007/s12033-008-9051-x

Liu, M., Maurano, M.T., Wang, H., Qi, H., Song, C.-Z., Navas, P.A., Emery, D.W., Stamatoyannopoulos, J.A., Stamatoyannopoulos, G., 2015. Genomic discovery of potent chromatin insulators for human gene therapy. Nat. Biotechnol. 33, 198–203. https://doi.org/10.1038/nbt.3062

Liu, P.-Q., Chan, E.M., Cost, G.J., Zhang, L., Wang, J., Miller, J.C., Guschin, D.Y., Reik, A., Holmes, M.C., Mott, J.E., Collingwood, T.N., Gregory, P.D., 2010. Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. Biotechnol. Bioeng. 106, 97–105. https://doi.org/10.1002/bit.22654

Lucas, B.K., Giere, L.M., DeMarco, R.A., Shen, A., Chisholm, V., Crowley, C.W., 1996. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. Nucleic Acids Res. 24, 1774–1779.

Ludwig, D.L., 2006. Mammalian Expression Cassette Engineering for High-Level Protein Production. BioProcess Int 4, S14–S23.

Majocchi, S., Aritonovska, E., Mermod, N., 2014. Epigenetic regulatory elements associate with specific histone modifications to prevent silencing of telomeric genes. Nucleic Acids Res. 42, 193–204. https://doi.org/10.1093/nar/gkt880

Marchant, R.J., Al-Fageeh, M.B., Underhill, M.F., Racher, A.J., Smales, C.M., 2008. Metabolic rates, growth phase, and mRNA levels influence cell-specific antibody production levels from in vitro-cultured mammalian cells at sub-physiological temperatures. Mol. Biotechnol. 39, 69–77. https://doi.org/10.1007/s12033-008-9032-0

Martin, U., n.d. New record of biologics sales in 2015 exceeding the USD 150 bln threshold [WWW Document]. URL http://www.lamerie.com/2013-10-02-18-25-07/item/21-new-record-of-biologics-sales-in-2015-exceeding-the-usd-150-bln-threshold (accessed 3.5.17).

Masters, J.R., 2002. TIMELINEHeLa cells 50 years on: the good, the bad and the ugly. Nat. Rev. Cancer 2, 315–319. https://doi.org/10.1038/nrc775

Matasci, M., Baldi, L., Hacker, D.L., Wurm, F.M., 2011. The PiggyBac transposon enhances the frequency of CHO stable cell line generation and yields recombinant lines with superior productivity and stability. Biotechnol. Bioeng. 108, 2141–2150. https://doi.org/10.1002/bit.23167

Mátés, L., Chuah, M.K.L., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., Grzela, D.P., Schmitt, A., Becker, K., Matrai, J., Ma, L., Samara-Kuko, E., Gysemans, C., Pryputniewicz, D., Miskey, C., Fletcher, B., VandenDriessche, T., Ivics, Z., Izsvák, Z., 2009. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat. Genet. 41, 753–761. https://doi.org/10.1038/ng.343

Matuszczak, E., Hermanowicz, A., Komarowska, M., Debek, W., 2013. Serum AMH in Physiology and Pathology of Male Gonads. Int. J. Endocrinol. 2013, 128907. https://doi.org/10.1155/2013/128907

Matzuk, M.M., Lamb, D.J., 2008. The biology of infertility: research advances and clinical challenges. Nat. Med. 14, 1197–1213. https://doi.org/10.1038/nm.f.1895

McBurney, M.W., Mai, T., Yang, X., Jardine, K., 2002. Evidence for repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes. Exp. Cell Res. 274, 1–8. https://doi.org/10.1006/excr.2001.5443

Meir, Y.-J.J., Weirauch, M.T., Yang, H.-S., Chung, P.-C., Yu, R.K., Wu, S.C.-Y., 2011. Genome-wide target profiling of piggyBac and Tol2 in HEK 293: pros and cons for gene discovery and gene therapy. BMC Biotechnol. 11, 28. https://doi.org/10.1186/1472-6750-11-28

Meissner, P., Pick, H., Kulangara, A., Chatellard, P., Friedrich, K., Wurm, F.M., 2001. Transient gene expression: recombinant protein production with suspension-adapted HEK293-EBNA cells. Biotechnol. Bioeng. 75, 197–203.

Milstein, C., 1999. The hybridoma revolution: an offshoot of basic research. BioEssays 21, 966–973. https://doi.org/10.1002/(SICI)1521-1878(199911)21:11<966::AID-BIES9>3.0.CO;2-Z

Miskey, C., Izsvák, Z., Plasterk, R.H., Ivics, Z., 2003. The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. Nucleic Acids Res. 31, 6873–6881. https://doi.org/10.1093/nar/gkg910

Miskey, C., Papp, B., Mátés, L., Sinzelle, L., Keller, H., Izsvák, Z., Ivics, Z., 2007. The ancient mariner sails again: transposition of the human Hsmar1 element by a reconstructed transposase and activities of the SETMAR protein on transposon ends. Mol. Cell. Biol. 27, 4589–4600. https://doi.org/10.1128/MCB.02027-06

Mohan, C., Kim, Y.-G., Koo, J., Lee, G.M., 2008. Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells. Biotechnol. J. 3, 624–630. https://doi.org/10.1002/biot.200700249

Müller, Johannes Peter nell'Enciclopedia Treccani [WWW Document], n.d. URL http://www.treccani.it//enciclopedia/johannes-peter-muller (accessed 11.5.17).

Müller-Kuller, U., Ackermann, M., Kolodziej, S., Brendel, C., Fritsch, J., Lachmann, N., Kunkel, H., Lausen, J., Schambach, A., Moritz, T., Grez, M., 2015. A minimal ubiquitous chromatin opening element (UCOE) effectively prevents silencing of juxtaposed heterologous promoters by epigenetic remodeling in multipotent and pluripotent stem cells. Nucleic Acids Res. 43, 1577–1592. https://doi.org/10.1093/nar/gkv019

Muñoz-López, M., García-Pérez, J.L., 2010. DNA transposons: nature and applications in genomics. Curr. Genomics 11, 115–128. https://doi.org/10.2174/138920210790886871

Nachtigal, M.W., Ingraham, H.A., 1996. Bioactivation of Müllerian inhibiting substance during gonadal development by a kex2/subtilisin-like endoprotease. Proc. Natl. Acad. Sci. U. S. A. 93, 7711–7716.

Nair, A.R., Jinger, X., Hermiston, T.W., 2011. Effect of different UCOE-promoter combinations in creation of engineered cell lines for the production of Factor VIII. BMC Res. Notes 4, 178. https://doi.org/10.1186/1756-0500-4-178

Nam, J.H., Ermonval, M., Sharfstein, S.T., 2009. The effects of microcarrier culture on recombinant CHO cells under biphasic hypothermic culture conditions. Cytotechnology 59, 81–91. https://doi.org/10.1007/s10616-009-9196-x

Narayanavari, S.A., Chilkunda, S.S., Ivics, Z., Izsvák, Z., 2017. Sleeping Beauty transposition: from biology to applications. Crit. Rev. Biochem. Mol. Biol. 52, 18–44. https://doi.org/10.1080/10409238.2016.1237935

Neville, J.J., Orlando, J., Mann, K., McCloskey, B., Antoniou, M.N., 2017. Ubiquitous Chromatin-opening Elements (UCOEs): Applications in biomanufacturing and gene therapy. Biotechnol. Adv. 35, 557–564. https://doi.org/10.1016/j.biotechadv.2017.05.004

Ng, S.K., Tan, T.R.M., Wang, Y., Ng, D., Goh, L.-T., Bardor, M., Wong, V.V.T., Lam, K.P., 2012. Production of functional soluble Dectin-1 glycoprotein using an IRES-linked destabilized-dihydrofolate reductase expression vector. PloS One 7, e52785. https://doi.org/10.1371/journal.pone.0052785

Ng, S.K., Wang, D.I.C., Yap, M.G.S., 2007. Application of destabilizing sequences on selection marker for improved recombinant protein productivity in CHO-DG44. Metab. Eng. 9, 304–316. https://doi.org/10.1016/j.ymben.2007.01.001

Nielsen, L.S., Baer, A., Müller, C., Gregersen, K., Mønster, N.T., Rasmussen, S.K., Weilguny, D., Tolstrup, A.B., 2010. Single-batch production of recombinant human polyclonal antibodies. Mol. Biotechnol. 45, 257–266. https://doi.org/10.1007/s12033-010-9270-9

Niwa, H., Yamamura, K., Miyazaki, J., 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene 108, 193–199.

Oguchi, S., Saito, H., Tsukahara, M., Tsumura, H., 2006. pH Condition in temperature shift cultivation enhances cell longevity and specific hMab productivity in CHO culture. Cytotechnology 52, 199–207. https://doi.org/10.1007/s10616-007-9059-2

Omasa, T., 2002. Gene amplification and its application in cell and tissue engineering. J. Biosci. Bioeng. 94, 600–605.

Other Plasmid Components [WWW Document], n.d. URL https://www.oxfordgenetics.com/SiteContent/TeamResources/Other_Plasmid_Co

mponents?complex=True&opened=internal-ribosome-entry-site-ires-information (accessed 12.1.17).

Otte, A.P., Kwaks, T.H.J., van Blokland, R.J.M., Sewalt, R.G.A.B., Verhees, J., Klaren, V.N.A., Siersma, T.K., Korse, H.W.M., Teunissen, N.C., Botschuijver, S., van Mer, C., Man, S.Y., 2007. Various expression-augmenting DNA elements benefit from STAR-Select, a novel high stringency selection system for protein expression. Biotechnol. Prog. 23, 801–807. https://doi.org/10.1021/bp070107r

Papakostas, T.D., Pieretti-Vanmarcke, R., Nicolaou, F., Thanos, A., Trichonas, G., Koufomichali, X., Anago, K., Donahoe, P.K., Teixeira, J., MacLaughlin, D.T., Vavvas, D., 2010. Development of an efficiently cleaved, bioactive, highly pure FLAG-tagged recombinant human Mullerian Inhibiting Substance. Protein Expr. Purif. 70, 32–38. https://doi.org/10.1016/j.pep.2009.09.004

Pau, M.G., Ophorst, C., Koldijk, M.H., Schouten, G., Mehtali, M., Uytdehaag, F., 2001. The human cell line PER.C6 provides a new manufacturing system for the production of influenza vaccines. Vaccine 19, 2716–2721.

Pépin, D., Hoang, M., Nicolaou, F., Hendren, K., Benedict, L.A., Al-Moujahed, A., Sosulski, A., Marmalidou, A., Vavvas, D., Donahoe, P.K., 2013. An albumin leader sequence coupled with a cleavage site modification enhances the yield of recombinant C-terminal Mullerian Inhibiting Substance. Technol. Elmsford N 1, 63–71. https://doi.org/10.1142/S2339547813500076

Pepinsky, R.B., Sinclair, L.K., Chow, E.P., Mattaliano, R.J., Manganaro, T.F., Donahoe, P.K., Cate, R.L., 1988. Proteolytic processing of mullerian inhibiting substance produces a transforming growth factor-beta-like fragment. J. Biol. Chem. 263, 18961–18964.

Petricciani, J., Sheets, R., 2008. An overview of animal cell substrates for biological products. Biol. J. Int. Assoc. Biol. Stand. 36, 359–362. https://doi.org/10.1016/j.biologicals.2008.06.004

Picanco-Castro, V., Biaggio, R.T., Cova, D.T., Swiech, K., 2013. Production of recombinant therapeutic proteins in human cells: current achievements and future perspectives. Protein Pept. Lett. 20, 1373–1381.

Pritham, E.J., 2009. Transposable Elements and Factors Influencing their Success in Eukaryotes. J. Hered. 100, 648–655. https://doi.org/10.1093/jhered/esp065

Puck, T.T., Cieciura, S.J., Robinson, A., 1958. GENETICS OF SOMATIC MAMMALIAN CELLS. J. Exp. Med. 108, 945–956.

Reilly, M.T., Faulkner, G.J., Dubnau, J., Ponomarev, I., Gage, F.H., 2013. The role of transposable elements in health and diseases of the central nervous system. J. Neurosci. Off. J. Soc. Neurosci. 33, 17577–17586. https://doi.org/10.1523/JNEUROSCI.3369-13.2013

Rey, R., 2005. Anti-Müllerian hormone in disorders of sex determination and differentiation. Arq. Bras. Endocrinol. Metabol. 49, 26–36. https://doi.org//S0004-27302005000100005

Rey, R., Josso, N., Racine, C., 2016. Sexual Differentiation. MDText.com, Inc.

Rivella, S., Callegari, J.A., May, C., Tan, C.W., Sadelain, M., 2000. The cHS4 Insulator Increases the Probability of Retroviral Expression at Random Chromosomal Integration Sites. J. Virol. 74, 4679–4687.

(Rob) Aggarwal, S., 2014. What's fueling the biotech engine—2012 to 2013. Nat. Biotechnol. 32, 32–39. https://doi.org/10.1038/nbt.2794

Rosenblum, G., Cooperman, B.S., 2014. Engine out of the chassis: cell-free protein synthesis and its uses. FEBS Lett. 588, 261–268. https://doi.org/10.1016/j.febslet.2013.10.016

Sanders, B.P., Edo-Matas, D., Custers, J.H.H.V., Koldijk, M.H., Klaren, V., Turk, M., Luitjens, A., Bakker, W.A.M., Uytdehaag, F., Goudsmit, J., Lewis, J.A., Schuitemaker, H., 2013. PER.C6(®) cells as a serum-free suspension cell platform for the production of high titer poliovirus: a potential low cost of goods option for world supply of inactivated poliovirus vaccine. Vaccine 31, 850–856. https://doi.org/10.1016/j.vaccine.2012.10.070

Santiago, Y., Chan, E., Liu, P.-Q., Orlando, S., Zhang, L., Urnov, F.D., Holmes, M.C., Guschin, D., Waite, A., Miller, J.C., Rebar, E.J., Gregory, P.D., Klug, A., Collingwood, T.N., 2008. Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc. Natl. Acad. Sci. 105, 5809–5814. https://doi.org/10.1073/pnas.0800940105

Saunders, F., Sweeney, B., Antoniou, M.N., Stephens, P., Cain, K., 2015. Chromatin Function Modifying Elements in an Industrial Antibody Production Platform - Comparison of UCOE, MAR, STAR and cHS4 Elements. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0120096

Schatz, S.M., Kerschbaumer, R.J., Gerstenbauer, G., Kral, M., Dorner, F., Scheiflinger, F., 2003. Higher expression of Fab antibody fragments in a CHO cell line at reduced temperature. Biotechnol. Bioeng. 84, 433–438. https://doi.org/10.1002/bit.10793

Schmidt, F.R., 2004. Recombinant expression systems in the pharmaceutical industry. Appl. Microbiol. Biotechnol. 65, 363–372. https://doi.org/10.1007/s00253-004-1656-9

Sharma, N., Hollensen, A.K., Bak, R.O., Staunstrup, N.H., Schrøder, L.D., Mikkelsen, J.G., 2012. The Impact of cHS4 Insulators on DNA Transposon Vector Mobilization and Silencing in Retinal Pigment Epithelium Cells. PLoS ONE 7. https://doi.org/10.1371/journal.pone.0048421

Sheets, 2000. History and Characterization of the Vero Cell Line [WWW Document]. URL https://www.fda.gov/ohrms/dockets/ac/00/backgrd/3616b1a.pdf (accessed.5.17).

Shi, M., Xie, Z., Yu, M., Shen, B., Guo, N., 2005. Controlled growth of Chinese hamster ovary cells and high expression of antibody-IL-2 fusion proteins by temperature manipulation. Biotechnol. Lett. 27, 1879–1884. https://doi.org/10.1007/s10529-005-3897-9

Skipper, K.A., Andersen, P.R., Sharma, N., Mikkelsen, J.G., 2013. DNA transposon-based gene vehicles - scenes from an evolutionary drive. J. Biomed. Sci. 20, 92. https://doi.org/10.1186/1423-0127-20-92

Sp2/0-Ag14 ATCC ® CRL-1581™ Mus musculus (B cell); Mus muscu [WWW Document], n.d. URL https://www.lgcstandards-atcc.org/products/all/CRL-1581.aspx?geo_country=it#generalinformation (accessed 9.5.17).

Stanley, P., Siminovitch, L., 1976. Selection and characterization of Chinese hamster ovary cells resistant to the cytotoxicity of lectins. In Vitro 12, 208–215.

Sun, Q.-L., Zhao, C.-P., Chen, S.-N., Wang, L., Wang, T.-Y., 2016. Molecular characterization of a human matrix attachment region that improves transgene expression in CHO cells. Gene 582, 168–172. https://doi.org/10.1016/j.gene.2016.02.009

Sunley, K., Tharmalingam, T., Butler, M., 2008. CHO cells adapted to hypothermic growth produce high yields of recombinant β-interferon. Biotechnol. Prog. 24, 898–906. https://doi.org/10.1002/btpr.9

Swiech, K., Picanço-Castro, V., Covas, D.T., 2012. Human cells: new platform for recombinant therapeutic protein production. Protein Expr. Purif. 84, 147–153. https://doi.org/10.1016/j.pep.2012.04.023

Thomas, C.E., Ehrhardt, A., Kay, M.A., 2003. Progress and problems with the use of viral vectors for gene therapy. Nat. Rev. Genet. 4, 346–358. https://doi.org/10.1038/nrg1066

Thomas, P., Smart, T.G., 2005. HEK293 cell line: a vehicle for the expression of recombinant proteins. J. Pharmacol. Toxicol. Methods 51, 187–200. https://doi.org/10.1016/j.vascn.2004.08.014

Turchiano, G., Latella, M.C., Gogol-Döring, A., Cattoglio, C., Mavilio, F., Izsvák, Z., Ivics, Z., Recchia, A., 2014. Genomic analysis of Sleeping Beauty transposon integration in human somatic cells. PloS One 9, e112712. https://doi.org/10.1371/journal.pone.0112712

Uchida, N., Hanawa, H., Yamamoto, M., Shimada, T., 2013. The chicken hypersensitivity site 4 core insulator blocks promoter interference in lentiviral vectors. Hum. Gene Ther. Methods 24, 117–124. https://doi.org/10.1089/hgtb.2012.152

Underhill, M.F., Marchant, R.J., Carden, M.J., James, D.C., Smales, C.M., 2006. On the effect of transient expression of mutated eIF2alpha and eIF4E eukaryotic translation initiation factors on reporter gene expression in mammalian cells upon cold-shock. Mol. Biotechnol. 34, 141–149.

Urlaub, G., Chasin, L.A., 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. Proc. Natl. Acad. Sci. U. S. A. 77, 4216–4220.

Urlaub, G., Käs, E., Carothers, A.M., Chasin, L.A., 1983. Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. Cell 33, 405–412.

van Helden, J., Weiskirchen, R., 2015. Performance of the two new fully automated anti-Müllerian hormone immunoassays compared with the clinical standard assay. Hum. Reprod. Oxf. Engl. 30, 1918–1926. https://doi.org/10.1093/humrep/dev127

Visser, J.A., Themmen, A.P.N., 2005. Anti-Müllerian hormone and folliculogenesis. Mol. Cell. Endocrinol. 234, 81–86. https://doi.org/10.1016/j.mce.2004.09.008

Voronina, E.V., Seregin, Y.A., Litvinova, N.A., Shvets, V.I., Shukurov, R.R., 2016. Design of a stable cell line producing a recombinant monoclonal anti-TNFα antibody based on a CHO cell line. SpringerPlus 5. https://doi.org/10.1186/s40064-016-3213-2

Wang, F., Wang, T.-Y., Tang, Y.-Y., Zhang, J.-H., Yang, X.-J., 2012. Different matrix attachment regions flanking a transgene effectively enhance gene expression in stably transfected Chinese hamster ovary cells. Gene 500, 59–62. https://doi.org/10.1016/j.gene.2012.03.049

Wang, T.-Y., Zhang, J.-H., Jing, C.-Q., Yang, X.-J., Lin, J.-T., 2010. Positional effects of the matrix attachment region on transgene expression in stably transfected CHO cells. Cell Biol. Int. 34, 141–145. https://doi.org/10.1042/CBI20090017

Westwood, A.D., Rowe, D.A., Clarke, H.R.G., 2010. Improved recombinant protein yield using a codon deoptimized DHFR selectable marker in a CHEF1 expression plasmid. Biotechnol. Prog. 26, 1558–1566. https://doi.org/10.1002/btpr.491

Williams, S., Mustoe, T., Mulcahy, T., Griffiths, M., Simpson, D., Antoniou, M., Irvine, A., Mountain, A., Crombie, R., 2005. CpG-island fragments from the HNRPA2B1/CBX3genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol. 5, 17. https://doi.org/10.1186/1472-6750-5-17

Wulhfard, S., 2009. Transient recombinant protein expression in mammalian cells. https://doi.org/10.5075/epfl-thesis-4393, urn:nbn:ch:bel-epfl-thesis4393-7

Wulhfard, S., Tissot, S., Bouchet, S., Cevey, J., De Jesus, M., Hacker, D.L., Wurm, F.M., 2008. Mild hypothermia improves transient gene expression yields several fold in Chinese hamster ovary cells. Biotechnol. Prog. 24, 458–465. https://doi.org/10.1021/bp070286c

Wurm, F., 2013. CHO Quasispecies—Implications for Manufacturing Processes. Processes 1, 296–311. https://doi.org/10.3390/pr1030296

Wurm, F.M., 2014. 2.2 CHO History, CHO Evolution and CHO Genomics – an Unsolvable Enigma?, in: Animal Cell Biotechnology In Biologics Production. De Gruyter, Berlin, Boston. https://doi.org/10.1515/9783110278965.38

Wurm, F.M., 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol. 22, 1393–1398. https://doi.org/10.1038/nbt1026

Wurm, F.M., Hacker, D., 2011. First CHO genome. Nat. Biotechnol. 29, 718–720. https://doi.org/10.1038/nbt.1943

Xu, X., Nagarajan, H., Lewis, N.E., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M.R., Neff, N., Passarelli, B., Koh, W., Fan, H.C., Wang, J., Gui, Y., Lee, K.H., Betenbaugh, M.J., Quake, S.R., Famili, I., Palsson, B.O., Wang, J., 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat. Biotechnol. 29, 735–741. https://doi.org/10.1038/nbt.1932

Yoon, S.K., Song, J.Y., Lee, G.M., 2003. Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in Chinese hamster ovary cells. Biotechnol. Bioeng. 82, 289–298. https://doi.org/10.1002/bit.10566

Yuan, Y.-W., Wessler, S.R., 2011. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. Proc. Natl. Acad. Sci. U. S. A. 108, 7884–7889. https://doi.org/10.1073/pnas.1104208108

Yusa, K., Zhou, L., Li, M.A., Bradley, A., Craig, N.L., 2011. A hyperactive piggyBac transposase for mammalian applications. Proc. Natl. Acad. Sci. U. S. A. 108, 1531–1536. https://doi.org/10.1073/pnas.1008322108

Zahn-Zabal, M., Kobr, M., Girod, P.A., Imhof, M., Chatellard, P., de Jesus, M., Wurm, F., Mermod, N., 2001. Development of stable cell lines for production or regulated expression using matrix attachment regions. J. Biotechnol. 87, 29–42.

Zhang, F., Frost, A.R., Blundell, M.P., Bales, O., Antoniou, M.N., Thrasher, A.J., 2010. A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. Mol. Ther. J. Am. Soc. Gene Ther. 18, 1640–1649. https://doi.org/10.1038/mt.2010.132

Zhao, C.-P., Guo, X., Chen, S.-J., Li, C.-Z., Yang, Y., Zhang, J.-H., Chen, S.-N., Jia, Y.-L., Wang, T.-Y., 2017. Matrix attachment region combinations increase transgene expression in transfected Chinese hamster ovary cells. Sci. Rep. 7. https://doi.org/10.1038/srep42805

Zhu, J., 2012. Mammalian cell protein expression for biopharmaceutical production. Biotechnol. Adv. 30, 1158–1170. https://doi.org/10.1016/j.biotechadv.2011.08.022